Methods for the Synchronous Acquisition of 3D Shape and Material Appearance

A dissertation presented to the faculty of the University of Virginia
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

Michael Joseph Holroyd
University of Virginia
Department of Computer Science

May 2011

APPROVAL SHEET

This dissertation is submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Michael Holroyd, author

Jason Lawrence, advisor

Worthy Martin, chair

Kevin Skadron

William Walker

Todd Zickler

Accepted for the School of Engineering and Applied Science:

James Aylor, dean

May, 2011

Abstract

Measuring the 3-dimensional shape and material appearance of physical objects is a challenging task, but has a wide range of applications in areas including cultural heritage preservation, architecture, law enforcement, industry, and entertainment. This thesis extends existing research on acquisition pipelines to enable synchronous measurement of shape and appearance. Previous work that measures each independently suffers from errors due to the difficult alignment and merging problem. We overcome the difficulties of synchronous acquisition by exploiting surface reflectance symmetries and a novel optical design.

We demonstrate the first photometric approach for measuring the shape and appearance of objects with anisotropic surface reflectance. For such materials, both the surface normal as well as tangent direction at each surface point must be recovered, unlike in the case of isotropic materials where the tangent direction can be ignored. We present an optimization strategy to locate vectors of maximum reflectance symmetry using a dense set of images taken under variable point lighting, resulting in a normal and tangent vector recovered at each pixel along with a dense 2D slice of the surface reflectance function.

We present an analysis of sinusoidal illumination in the context of shape and appearance measurement, which shows how to simultaneously perform optical descattering, recover 3D geometry, and acquire reflectance measurements from images taken under phase-shifted sinusoidal lighting. In addition, we show that accurately measuring the geometry of translucent objects using sinusoidal illumination is possible for a wide range of materials and frequencies. These two theoretical results provide a basis for the design of practical scanning systems.

Finally, a full pipeline for acquiring the shape and appearance of opaque objects is presented. We combine sinusoidal illumination with a novel optical design to create

coaxial devices, which greatly simplify the joint measurement task compared to prior work. Using the key result from our analysis of sinusoidal illumination and a novel multiview stereo algorithm, we demonstrate the capture of complete 360° surface meshes with spatially varying reflectance. Our evaluation compares the geometry obtained by our system with state-of-the-art laser scans; the reflectance with traditional reference gonioreflectometers; and the final model with reference photographs.

Contents

List of Figures

Chapter 1

Introduction

The automated digitization of real-world objects offers a practical and more accurate alternative to the tedious and difficult process of manual modeling. For example, accurate digital representations of important statues, paintings, and artifacts give museums new opportunities for analyzing, disseminating, and preserving these important artifacts. Architects with access to digital models of building interiors can view photorealistic images of proposed modifications. Scientific simulations and animations require highly accurate depictions of real world objects and how they interact with their environment. Computer vision tasks such as automated quality control and reverse engineering are only enabled by accurate models of the geometry and materials in question. The entertainment industry also has a great interest in digitizing physical objects to improve the realism of movies and video games.

We show an approach for measuring high-resolution normal and tangent fields along with the surface reflectance of objects. By acquiring many photographs of an object under different controlled lighting conditions, and exploiting symmetries in the way light reflects off object surfaces, we estimate shape in the form of a local surface orientation at each pixel. This enables the measurement, rendering, and editing of *anisotropic ma-*

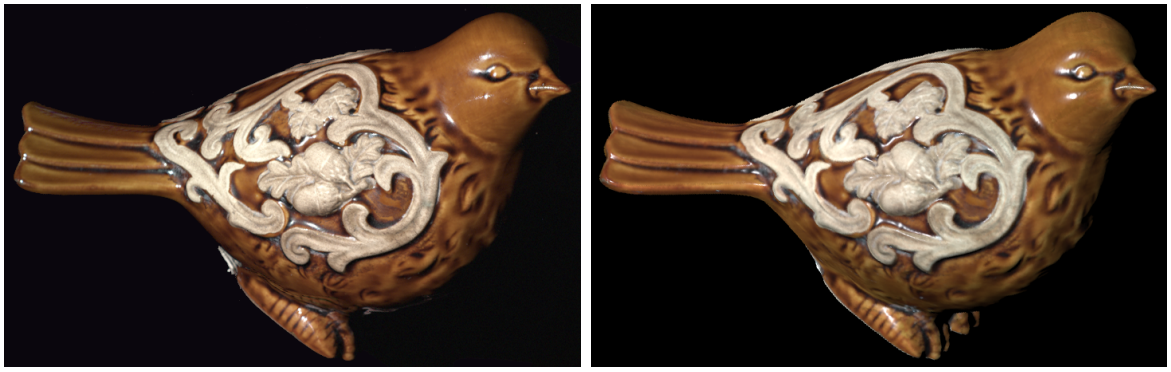


Figure 1.1: A digital replica of a porcelain and stone bird statue, acquired using our coaxial scanner system. **(left)** Photograph of the original object. **(right)** Rendering of the digital model.

terials such as brushed metal and satin, which are asymmetric with respect to rotations around the surface normal.

Sinusoidal illumination has been used successfully in many 3D scanning applications. This illumination pattern also enables computationally separating the contribution of single scattering from multiple scattering events such as subsurface scattering and inter-reflections by assuming only direct reflections retain the high-frequency content of the original sinusoid. We provide a detailed analysis of optical descattering using sinusoidal illumination, which shows that the single scattering component is strongly affected by the choice of sinusoid frequency as well as the geometric configuration of camera and lightsource. Fortunately, this effect can be corrected for using our analysis, enabling accurate joint measurement of 3D shape and surface reflectance. We also present an extension of this analysis to the more complex case of translucent objects, which allow light to scatter beneath the surface of the object. In this case, we investigate the relationship between material parameters and bias in the estimated surface geometry. Both of these results have broad implications for image-based measurement systems, and give insight into how such systems should be designed for different materials.

Finally, we introduce a full pipeline for joint measurement of opaque objects that

results in complete 3D meshes with surface reflectance measurements over their surface (see figure 1.1). This pipeline exploits a novel optical setup that utilizes two "coaxial devices" – a camera and lightsource aligned to share a common optical axis and focal point – with high-frequency sinusoidal lightsources. Our approach is enabled by a novel multiview stereo algorithm as well as the previous analysis of optical descattering. We compare both the recovered geometry and reflectance data against ground truth measurements, and show rendered images of our final models that closely match reference photographs.

Chapter 2

Background and Related Work

Systems designed to measure the of shape and appearance of physical objects draw on knowledge from computer vision, radiometry, and reflectometry. This chapter reviews the core concepts and related work from these areas.

2.1 Shape

The *shape* of many objects is well described by a 2D surface associated with the external boundary of the object. Other objects such as pillars of smoke or inhomogeneous liquids may require a 3D volumetric description, but will not be considered here. Each point on a smooth surface also has an implicit *surface normal* that points perpendicular to the local tangent plane as shown in figure 2.1.

The most common digital representation for a 2D surface is a triangle mesh, one natural piece-wise linear discrete representation of a 2D manifold. Triangle meshes have a long history in computer graphics, and current graphics hardware is heavily optimized toward their use. One alternative is to consider a freeform point-cloud, such as seen in figure 2.2, either with or without associated surface normals at each point. This repre-

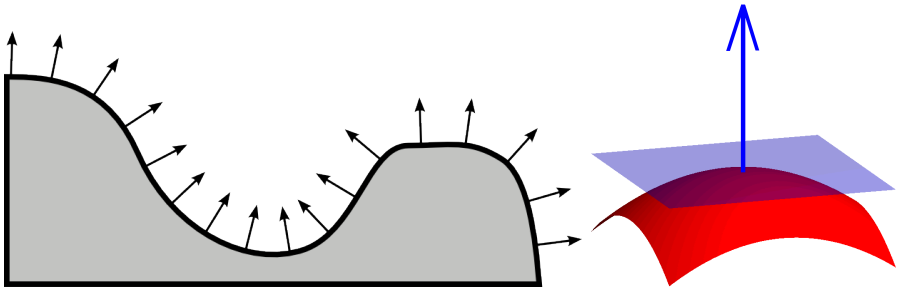


Figure 2.1: Surface normals represented in 2D and 3D.

sentation is disadvantaged by having less hardware rendering support; however, point clouds are the output of many 3D scanning techniques as will be described in section 2.3. Reducing a mesh to a point cloud is straightforward, and techniques also exist for converting point clouds to meshes, such as the ball-pivoting algorithm [Bernardini et al., 1999] or Poisson surface reconstruction [Kazhdan et al., 2006].

Figure 2.2: The 2D surface of a small statue represented with a connected triangle mesh and with point samples.

2.2 Appearance

The *appearance* of an object is determined by how incoming light interacts with the object to produce outgoing light. The amount of light traveling along a ray is measured

by the radiance, L, $\frac{\mathrm{W}}{\mathrm{srm}^2}$. One approach is to ignore the underlying surface's geometry and view this interaction as a black box, known as the reflectance field [Debevec et al., 2000]. The reflectance field is an 8-dimensional function defined as the ratio of incoming radiance to outgoing radiance, for every possible incoming and outgoing ray of light. One way to see that the reflectance field is 8-dimensional is to consider indexing all rays based on their location on some bounding sphere that encloses the object. In this case each ray is defined by its intersection with the sphere (2D) and its direction of travel (2D). We use the notation $R(x_i, \omega_i, x_o, \omega_o)$, indicating the points x and directions ω. The additional dependence of this function on wavelength is usually ignored, and instead computations are carried out independently in the red, green, and blue color channels. Defining the appearance of an object independently from its geometry has the advantage of disentangling the geometric complexity from the representation. Measurement systems have been proposed for restricted slices of the reflectance field, such as considering fixed lighting [Levoy and Hanrahan, 1996; Gortler et al., 1996] or distant lighting "non-local reflectance fields" [Debevec et al., 2000].

In contrast, an accurate model of the surface geometry enables tasks such as animation, shape analysis, smoothing algorithms, collision detection, and so on. Representing appearance as reflectance at the surface also makes editing object appearance easier, for example using 3D painting interfaces. Following nomenclature introduced by Nicodemus [1977], the 4D *bidirectional reflectance distribution function* (BRDF) written $f_r(\omega_i, \omega_o)$ defines the ratio of outgoing radiance along ω_o to the irradiance ($\frac{\mathrm{W}}{\mathrm{m}^2}$) at a surface point from ω_i. The 6D *spatially varying BRDF* (SVBRDF), $f_r(x, \omega_i, \omega_o)$ defines a BRDF at each point x on the surface. For non-opaque objects (such as figure 2.3) in which light entering at a point x_i might exit the surface at a different point x_o, the more general 8D *bidirectional surface scattering reflectance distribution function* (BSSRDF) is similar to the full reflectance field, but defined on the object's surface, written

$S(x_i, \omega_i, x_o, \omega_o)$.

Figure 2.3: Example of an object rendered with an opaque BRDF (left), and a translucent BSSRDF (right).

Surface scattering functions are traditionally defined relative to the local coordinate system at the surface point x – the surface normal $\mathbf{n}$ and tangent direction $\mathbf{t}$. For isotropic BRDFs, which are symmetric under rotations around the surface normal, the tangent direction can be left undefined. For this special case, the BRDF can be reduced to a 3-dimensional function of $(\theta_i, \theta_o, \phi_i - \phi_o)$ – see figure 2.5. In chapter 3 we explore the importance of anisotropic BRDFs and describe an approach for jointly measuring these functions along with their local coordinate system using photographs under controlled lighting.

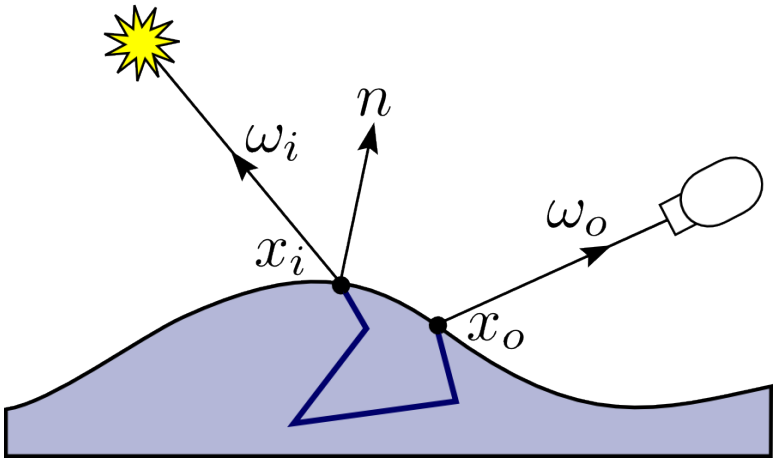


Figure 2.4: Notation used for defining the bidirectional surface scattering reflectance distribution function (BSSRDF), $S(x_i, \omega_i, x_o, \omega_o)$.

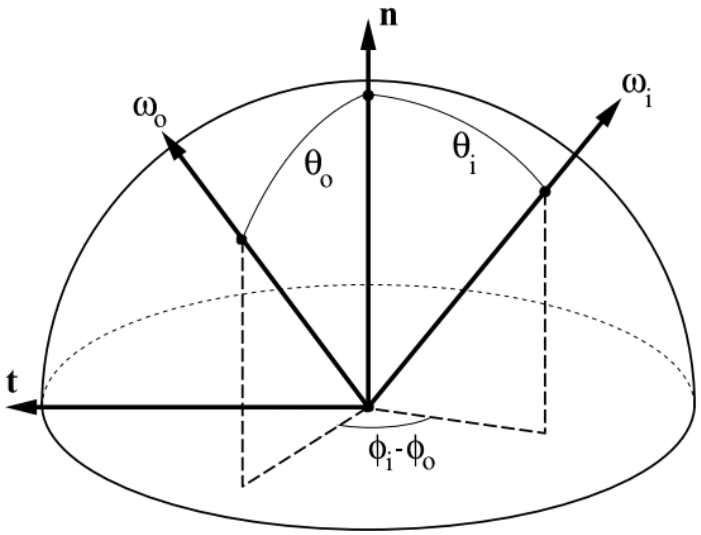


Figure 2.5: Isotropic BRDFs can be reduced to 3-dimendional functions of $(\theta_i, \theta_o, \phi_i - \phi_o)$.

2.3 Shape Acquisition

Methods for estimating the 3D geometry of objects can be broadly classified as active or passive, based on whether they introduce energy into the scene (laser, light, sound, physical contact, etc.) to help acquire shape. Passive techniques include classic stereo triangulation [Marr and Poggio, 1976], shape from motion [Tomasi and Kanade, 1992], shape from shading [Horn, 1975], and many extensions. Passive techniques often perform poorly compared to their active counterparts, because most cues for shape can be ambiguous in natural settings, especially in textureless regions. We instead focus on active techniques.

A variety of active techniques are available, such as those in figure 2.6. Touch probes are the standard tool for industrial inspection as they achieve measurement accuracies exceeding 25 micrometers and are virtually noise free. Unfortunately, touch probes can only take several measurements per second due to the mechanical readjustment required, making them impractical for acquiring densely sampled geometry. In addition, making physical contact with the object is often undesirable, especially in law enforcement or cultural heritage applications. Volumetric techniques such as computer tomography (CT) and magnetic-resonance imaging (MRI) offer another method for acquiring geometry, but place restrictions on the materials that can be scanned (non-metallic) and are very expensive. Time of flight systems record the length of time a signal takes to reach and return from the object being scanned. Time of flight laser scanners in particular have grown in speed and accuracy with recent models able to achieve submillimeter accuracy at distances up to 10 meters [Boehler et al., 2003]. The primary disadvantage of such systems is their bulkiness, which makes them difficult to reposition accurately and automatically. Such scanners have difficulty with shiny objects that reflect most of the signal away from the detector. Triangulation based

systems include structured light scanners that required calibrated lighting[Valkenburg and McIvor, 1998], as well as special cases of spacetime stereo [Zhang et al., 2003; Davis et al., 2005] such as commonly used laser/light stripe scanners and simple uncalibrated projector time-coding. The appearance of objects is most naturally measured using cameras, and thus image-based approaches are desirable for the joint measurement of shape and appearance. In addition to simplifying system design by requiring no extra parts, it also reduces calibration requirements and avoids the difficult process of merging data from different sensors.

Figure 2.6: A variety of scanning technologies. **(left)** Touch probe. **(center)** Time of flight laser scanner. **(right)** Magnetic resonance imaging machine.

In addition to measuring shape directly as 3D geometry in the scene, *photometric techniques* exist that recover surface normals by observing a fixed scene under different lighting conditions. Classical photometric stereo [Woodham, 1980] estimates the surface normal and albedo of a perfectly diffuse (Lambertian) object as seen from each pixel in a fixed view using three or more input images under varying point lighting. Although real-world materials are never perfectly diffuse, the simplicity and stability of the resulting linear optimization are undeniable, and this simple technique can be made more robust using outlier rejection to reduce the effects of highlights (see figure 2.7). Additionally,

color-based analysis [Mallick et al., 2005] and/or intensity-based analysis [Coleman and Jain, 1982] can be used to isolate the diffuse component of materials with more complex reflectance provided that a significant diffuse component exists.

Numerous improvements to this basic idea have been proposed, most of which extend photometric stereo to non-Lambertian materials by fitting measurements to low-dimensional BRDF models while simultaneously estimating surface normals. Tagare et al. [1991] extended the Lambertian case to a simple multi-lobe reflectance model. Georghiades [2003] estimates surface orientation and reflectance (as well as light source positions) by assuming the surface reflectance function can be represented by a simplified Torrance-Sparrow BRDF [Torrance and Sparrow, 1967] with a constant specular component over the surface. Goldman et al. [2005] allow more variation over the surface by simultaneously estimating a set of basis materials modeled as isotropic Ward BRDFs [1992]. A similar method was recently proposed by Chung et al. [2008] which focuses on materials with wide specular lobes. The disadvantage of fitting data to parametric models is that it can be difficult to predict when materials deviate from these models, and the resulting errors in the final surface normals are difficult to characterize.

A few techniques avoid relying on parametric models. Hertzmann and Seitz [2003] use homogeneous reference objects of known shape in the scene and model the surface as a spatially-varying mixture of these known basis materials. Unfortunately, reference objects are not practical to create for many scenes and materials. Another example is the work of Alldrin et al. [2008], which represents isotropic reflectance using a linear basis of general non-paramateric bivariate functions. Although such an approach holds promise, jointly estimating the surface BRDFs and surface normals leads to a fragile and expensive optimization that must integrate reflectance data across the entire object. A key benefit of the approach we present in chapter 3 is that it processes each pixel independently and does so without imposing any particular analytic model on the data.

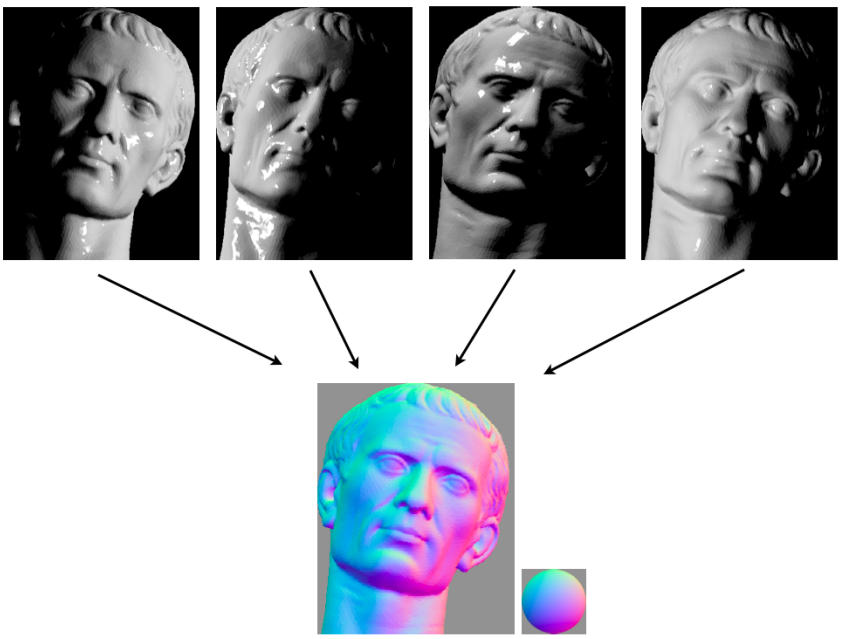


Figure 2.7: Example of a normal field (bottom) computed from a set of input images under varying point lighting (top) using a robust variant of photometric stereo [Wu et al., 2010]. We use this type of RGB visualization for normal fields throughout this thesis.

Similar to our approach are methods that exploit symmetries in real-world BRDFs. Zickler et al. [2002] exploit Helmholtz reciprocity by capturing pairs of images in which the position of the camera and light source are interchanged. This enables the recovery of both depth and surface normals independent of reflectance but requires finding corresponding points in images taken from different viewpoints. In contrast, our single-view approach permits a simpler acquisition and allows the direct recovery of surface normals at the resolution of the camera. Isotropy is another form of symmetry that has been exploited for photometric stereo by Alldrin and Kriegman [2007]. For isotropic BRDFs, they show that one component of the surface normal can be recovered by exploiting the reflective symmetry across the view-normal plane. In contrast, the symmetries we consider apply to both isotropic and anisotropic materials and allow measuring the complete surface normal and tangent directions.

2.3.1 Appearance Acquisition

The BRDFs of specific materials were traditionally acquired using a gonioreflectometer – large expensive devices designed to take measurements from a sample of the material at individual view and light directions. Dense angular measurements are impractical with this design, however high-accuracy measurements are available at sparse sample points for many materials [Touloukian, 1970].

Image-based techniques for measuring view and light dependent reflectance have flourished with the increased availability of high resolution digital cameras. Many designs have been proposed that use the high spatial resolution of camera CCDs to densely sample the angular domain of the BRDF. Ward proposed one of the earliest such designs, using a half-mirrored dome to measure the outgoing light field around a small planar patch of homogeneous material in a single exposure [1992]. Dana [2001] proposed a design utilizing a parabolic mirror, which only requires the lightsource to move in a plane. Many devices require sample objects of cylindrical or spherical geometry with homogeneous material to allow sampling an isotropic BRDF from a fixed camera position (see figure 2.8) [Marschner et al., 2000; Matusik et al., 2003]. The design of Marschner et al. [1999] also assume a homogeneous isotropic BRDF, but use a range scanner to acquire object geometry and align reflectance measurements using fiducials in the room, attached to the sample, and on the lightsource. Ngan et al. [2005] wrap strips of anisotropic material at different orientations around a cylinder to allow dense sampling in the angular domain and coarse sampling of rotations around the surface normal. Dana et al. [1999] have defined and acquired the *bidirectional texture function* (BTF), a non-local reflectance field defined over a small patch near the surface. Their measurements can be averaged spatially to estimate BRDF measurements of the surface if viewed from a distance. They use a robotic arm to automatically position a planar sample at different orientations, which allows them to leave the lightsource fixed and reposition

only the camera several times. Han and Perlin [2003] take similar measurements more quickly using a kaleidoscope instead of moving the camera and lightsource. McAllister [2002] uses a robotic arm to position planar samples, but measures the SVBRDF by holding the camera fixed and mechanically moving the lightsource. Marschner et al. [2005] estimate the parameters to a scattering model of finished wood using planar samples. They use a computer-controlled spherical gantry to accurately position a camera and lightsource on a sphere surrounding the object. This substantially reduces the amount of manual positioning required and makes a much denser sampling of the scene possible. Zickler et al. [2005] assume known geometry and estimate the SVBRDF with a small number of images by assuming the BRDF varies smoothly over the surface. Peers et. al [2006] estimate the heterogeneous BSSRDF of planar samples from a fixed view and light direction (ignoring directional dependency). Ghosh et al. [2007] recently demonstrated an acquisition device for isotropic materials that reduces capture time by requiring no moving parts, and take measurements of a planar sample directly under zonal basis lighting. Romeiro and Zickler [2010] simultaneously estimate unknown reflectance and lighting from a single image of an object with known shape using a prior on the statistics of real-world lighting environments. The BRDF is represented in a 2D bivariate domain [Romeiro et al., 2008]. See the survey paper by Weyrich et. al [2007] for more details on measurement systems for surface scattering functions.

2.3.2 Joint Acquisition of Shape and Appearance

Fewer systems have been designed to measure the spatially varying reflectance of objects with complex geometry. Sato et al. [1997] acquire geometry using a light-stripe scanner and take sparse reflectance measurements with the same video camera. The main disadvantage of this setup is that stripe based scanners must take many images per scan, which results in either long acquisition times or, in the case of a video camera, very noisy

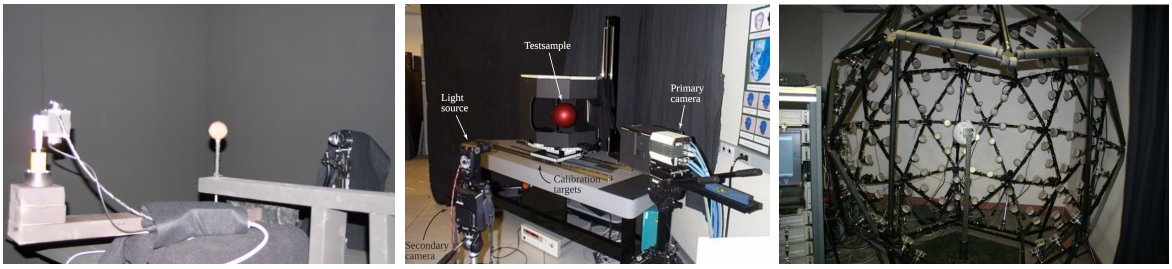


Figure 2.8: Three systems for measuring surface reflectance. **(left,center)** A camera photographs spherical samples under varying point lighting [Matusik et al. 2003; Marschner et al. 2000]. **(right)** A dome of lights accelerates the process by removing moving parts from the system [Gu et al. 2006]

low-quality images. Lensch et al. [2003a] manually adjust cameras and light sources and use chrome spheres to calibrate the lightsource relative to the camera. 3D geometry is acquired with a CT scan and is aligned to each image taken under point lighting using silhouette matching. Aside from the shortcomings of a silhouette matching technique, slight camera miscalibration inevitably results in misalignment between images and geometry. This results in misprojected reflectance measurements, especially near depth discontinuities. To alleviate these problems, discontinuities are detected explicitly and adjacent samples are discarded. Goesele et al. [2004] propose a design for measuring translucent objects, but must coat the objects with diffuse powder before acquiring geometry with a laser scanner. They use the same silhouette matching algorithm to register the resulting 3D model to 2D images. Goldman et al. [2005] avoid explicitly recovering the geometry and instead compute per-pixel normals by fitting reflectance measurements to a linear combination of Ward model BRDFs. Their setup requires calibrating each new light position using a chrome sphere to estimate the lightsource direction and a diffuse sphere to estimate intensity. Alldrin et al. [2008] take a similar approach, but assume a bivariate isotropic BRDF instead of fitting to the Ward model. Unfortunately, integrating normals to recover an explicit surface is prone to low-frequency bias and is only applicable to objects with no depth discontinuities. Muller et al. [2005] use an

elaborate 151-camera dome to quickly measure many view and light combinations. Despite the efficiency of their design, geometry is acquired using a convex-hull algorithm and results are only shown for a nearly spherical object. Weyrich et al. [2006] have introduced a larger dome with the goal of measuring spatially-varying surface reflectance and subsurface scattering parameters of human faces. They acquire geometry with a commercial structured light scanner. This reintroduces the misprojection problems from Lensch et al., but is less noticeable in the case of human faces that have smoothly varying reflectance properties. In chapter 5 we present a new optical device and algorithms for synchronous measurement of shape and appearance that overcome these limitations.

Chapter 3

A Photometric Approach for Estimating Normals and Tangents

First presented in Holroyd et al. [2008] at SIGGRAPH Asia 2008.

Photometric stereo uses images of a scene under different lighting conditions to infer information about the surface orientation and reflectance. As discussed in section 2.3, prior work on photometric stereo has largely relied on strong assumptions about the observed surface reflectance, such as requiring it to obey a certain analytic function [Woodham, 1980; Tagare and DeFigueiredo, 1991; Georghiades, 2003; Goldman et al., 2005] or be isotropic [Alldrin et al., 2008]. Anisotropic BRDFs are those with a preferred scattering direction around the surface normal, in contrast to isotropic BRDFs which are rotationally symmetric (see figure 3.1). Common anisotropic materials include brushed metal, satin, grass, wood, hair, and velvet. Previously, no photometric technique has been designed to measure both the surface normal as well as tangent directions of anisotropic materials.

We present a new photometric technique that overcomes these limitations. As with existing approaches, the input to our algorithm consists of a sequence of images taken

under varying point lighting, and the output includes an estimate of the surface normal for each pixel in a reference view. Unlike previous techniques, however, our approach does not require estimating a model of the surface reflectance, and it performs well for diffuse, glossy, shiny, metallic, dielectric, isotropic, anisotropic, homogeneous and spatially-varying surfaces.

Our basic strategy is to identify symmetries in the 2D slice of the BRDF obtained by fixing the local view direction. In particular, we consider the reflective symmetry of the halfway vector (the bisector of the local view and light directions) across the normal-tangent and normal-binormal planes. Although these symmetries are not required to be satisfied for all physically-valid BRDFs, we show that when light positions are properly restricted during acquisition, one can reliably estimate normals for a wide range of BRDFs by localizing their planes of symmetry.

In addition to providing a means for recovering surface normals, our symmetry-based approach has the important advantage of delivering per-pixel tangent vectors as well. For anisotropic surfaces, the tangent field rivals the normal field in its importance for reflectometry and rendering, but to date, its recovery has received very little attention.

This chapter presents empirical data for analytic and measured BRDFs demonstrating that our approach produces normals and tangents that are accurate to within a few degrees. We also show normal and tangent fields measured from real-world objects and demonstrate appearance editing applications.

3.1 Theoretical Framework

The goal of this work is to estimate per-pixel normal and tangent vectors $(\mathbf{n}, \mathbf{t})$ from a sequence of images recorded at different light positions. For a fixed view vector $\mathbf{v}$, let $\beta_{\mathbf{v}}(\theta_h, \phi_h)$ denote the 2D slice of the BRDF sampled at each pixel, parameterized by the

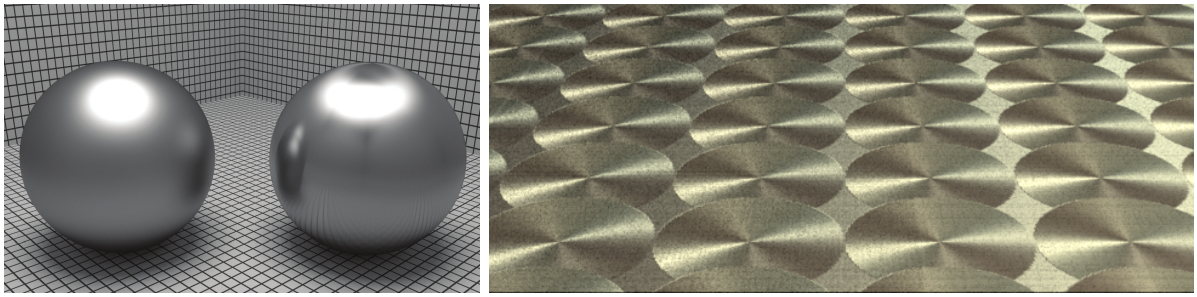


Figure 3.1: **(left)** An isotropic metallic BRDF, next to an anisotropic BRDF of brushed metal. **(right)** A plane rendered using measured anisotropic brushed aluminium and a procedurally generated tangent field. The BRDF is constant over the plane, only the tangent direction is changing.

elevation and azimuthal angles of the halfway vector expressed in the local coordinate system defined by $(\mathbf{n}, \mathbf{t})$. As illustrated in Figure 3.2, we explore the following symmetries in these functions:

- Reflection of the halfway vector across the **nt** plane:

  $\beta_{\mathbf{v}}(\theta_h, \phi_h) = \beta_{\mathbf{v}}(\theta_h, -\phi_h)$

- Reflection of the halfway vector across the **nb** plane:

  $\beta_{\mathbf{v}}(\theta_h, \phi_h) = \beta_{\mathbf{v}}(\theta_h, -\phi_h + \pi)$

- Together, these two symmetries imply reflection across the surface normal:

  $\beta_{\mathbf{v}}(\theta_h, \phi_h) = \beta_{\mathbf{v}}(\theta_h, \phi_h + \pi)$

These symmetries are different from those typically associated with BRDFs such as isotropy, which implies rotation of the light and view around the surface normal; we are considering BRDF values at light/view positions corresponding to reflections of *their associated halfway vectors*. Although such symmetries are not necessary for a physically valid BRDF, we will show that they are present for a wide range of materials and under certain light/view configurations.

Although many simple analytic BRDFs (e.g. Blinn-Phong [1977]) exhibit these symmetries perfectly, it is more informative to consider physically-based models that apply

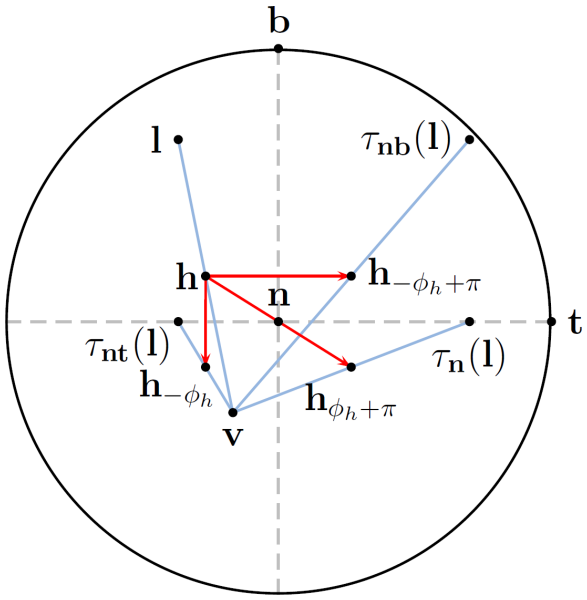


Figure 3.2: Orthographic projection of the hemisphere surrounding the surface normal **n** annotated with the tangent **t** and binormal **b** along with the three types of symmetry explored in this paper. For a fixed view **v**, we expect the value of the BRDF at light position **l** to be equal to that at light positions corresponding to reflections of the halfway vector **h** across the **nt** plane, the **nb** plane, and the surface normal.

to a much wider class of materials. In particular, microfacet models treat the surface as a collection of microscopic Fresnel mirrors and have been shown to accurately represent a wide variety of both isotropic and anisotropic materials [Cook and Torrance, 1981; Ashikhmin et al., 2000; Ngan et al., 2005]. The specular component of a microfacet-based BRDF has the form [Torrance and Sparrow, 1967; Ashikhmin et al., 2000]

$$f_r(\mathbf{v},\mathbf{l}) \propto p(\mathbf{h})F(\mathbf{l}\cdot\mathbf{h})\frac{S(\mathbf{v},\mathbf{l},\mathbf{h})}{(\mathbf{v}\cdot\mathbf{n})(\mathbf{l}\cdot\mathbf{n})}, \tag{3.1}$$

where $p(\mathbf{h})$ captures the distribution of microfacet orientations, F is the Fresnel term, and $S(\mathbf{v},\mathbf{l},\mathbf{h})$ models shadowing and masking effects at the micro scale. The distribution $p(\mathbf{h})$ is commonly modeled with an analytic function such as an elliptical Gaussian [Ward, 1992] or the Beckmann distribution [Cook and Torrance, 1981]. Alterna-

tively, Ngan et al. [2005] have shown that it can be fit to measured data using Ashikhmin et al.'s [2000] approximation to Equation 3.1.

In the context of the symmetries described above, one can make two observations regarding the microfacet model: 1) the distribution term $p(\mathbf{h})$ has a dominating influence on appearance for a broad range of view and light directions; and 2) it generally exhibits the symmetries that we seek to exploit. The latter observation comes from existing analytical and measured distributions, almost all of which satisfy the symmetry conditions. While exceptions do exist (see Section 3.4), we assume they are rare. The first observation is based on both physical and empirical evidence. Empirically, while the right-most term in Equation 3.1 can be quite complex, it is typically smooth and can be simplified tremendously without significant loss of visual accuracy [Ashikhmin et al., 2000; Ngan et al., 2005]. The other term to consider—the Fresnel term—is approximately proportional to $(1 - (\mathbf{l} \cdot \mathbf{h}))^5$ [Schlick, 1994], and is substantial only at grazing angles. By restricting the set of light positions to a limited cone of directions about the view vector, the influence of this term can be made small as well.

In the following, we present an acquisition system that samples the BRDF inside the cone of light directions for which $\cos^{-1}(\mathbf{l} \cdot \mathbf{h}) < \theta_{d_{\max}}$ (the notation θ_d is due to Rusinkiewicz [1998]; it refers to the elevation angle of the *difference* vector). Since θ_d depends only on the light and view directions, this bound can be achieved by simply restricting the light source positions during acquisition without any prior knowledge of the surface normal. By bounding the light directions in this way, we ensure the dominance of the microfacet distribution. Our results show that it is then possible to robustly identify the planes of reflective symmetry from the partial slice of the BRDF for a wide range of surface orientations.

3.1.1 A Measure of BRDF Symmetry

In this section, we define a precise measure of the symmetry previously discussed. For a fixed view, let $\mathbf{l}$ and $\mathbf{l}'$ be two light positions with corresponding halfway vectors at symmetric positions across some plane. The measured intensities at these positions will be $I_{\mathbf{l}} \propto f_r(\mathbf{v},\mathbf{l})(\mathbf{n}\cdot\mathbf{l})$ and $I_{\mathbf{l}'} \propto f_r(\mathbf{v},\mathbf{l}')(\mathbf{n}\cdot\mathbf{l}')$. We ignore the intensity of the light which is the same in all measurements and therefore does not affect our analysis. Note that even if the BRDF is equal at these positions, the product of the BRDF and the cosine term need not be. This can easily be corrected by cross-multiplying the cosine terms. In other words, if the BRDF is indeed symmetric across the chosen plane then $(\mathbf{n}\cdot\mathbf{l}')I_{\mathbf{l}} = (\mathbf{n}\cdot\mathbf{l})I_{\mathbf{l}'}$.

Now let $\tau(\mathbf{l})$ be an arbitrary transformation on lighting positions. For some hypothesized normal $\mathbf{n}$ and tangent $\mathbf{t}$, we define the *symmetry distance* of the BRDF under this transformation as:

$$\mathrm{SD}_\tau(\mathbf{n},\mathbf{t}) = \frac{\int_{\Omega_\tau} ||(\mathbf{n}\cdot\tau(\mathbf{l}))I_{\mathbf{l}} - (\mathbf{n}\cdot\mathbf{l})I_{\tau(\mathbf{l})}||^2\, d\omega_l}{\int_{\Omega_\tau} ||(\mathbf{n}\cdot\tau(\mathbf{l}))I_{\mathbf{l}}||^2\, d\omega_l},$$

where the domain of integration Ω_τ is constrained to include only light positions $\mathbf{l}$ and $\tau(\mathbf{l})$ that lie in the upper hemisphere with respect to $\mathbf{n}$ and correspond to halfway vectors between 0 and $\theta_{d_{\max}}$. As shown in Figure 3.2, we define $\tau_{\mathbf{nt}}(\mathbf{l})$ to reflect the halfway vector associated with $\mathbf{l}$ across the $\mathbf{nt}$ plane, $\tau_{\mathbf{nb}(\mathbf{l})}$ to reflect the halfway vector across the $\mathbf{nb}$ plane, and $\tau_{\mathbf{n}}$ to reflect the halfway vector across the normal. Note also that the transformation $\tau_{\mathbf{n}}$ is independent of the chosen tangent direction. Finally, we define a combined symmetry measure as

$$\mathrm{SD}(\mathbf{n},\mathbf{t}) = \mathrm{SD}_{\tau_{\mathbf{nt}}}(\mathbf{n},\mathbf{t}) + \mathrm{SD}_{\tau_{\mathbf{nb}}}(\mathbf{n},\mathbf{t}) + \mathrm{SD}_{\tau_{\mathbf{n}}}(\mathbf{n}). \tag{3.2}$$

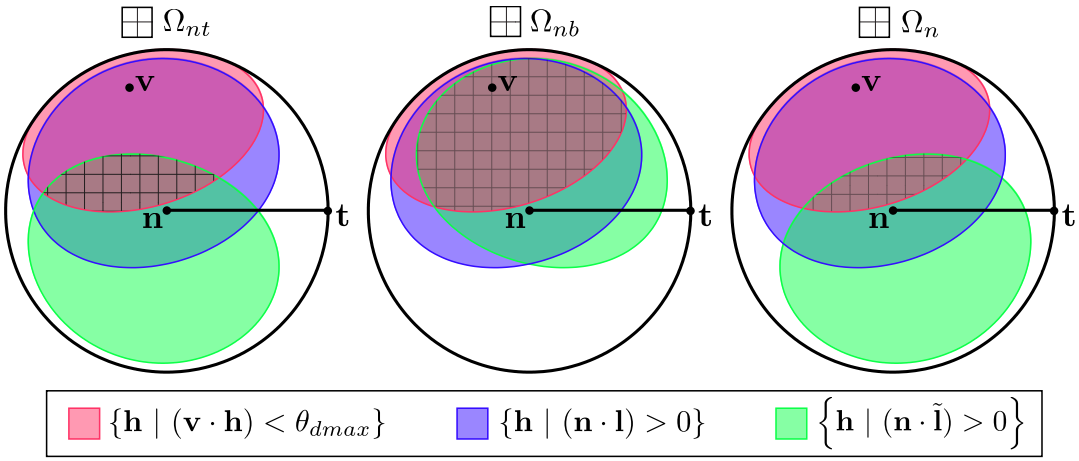


Figure 3.3: Integration regions for the three different transformations we consider shown over halfway vectors. These are defined as the intersection of the sets of halfway vectors for which $(\mathbf{n}\cdot\mathbf{l}) > 0$ and $(\mathbf{n}\cdot\mathbf{l}') > 0$ that also have $\cos^{-1}(\mathbf{v}\cdot\mathbf{h})$ within $\theta_{d_{\max}}$. These regions can form complex shapes and can even become empty for normals that lie beyond $\theta_{d_{\max}}$.

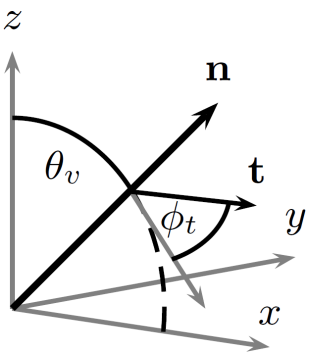
When $\mathbf{n}$ and $\mathbf{t}$ coincide with the true surface frame, we expect SD to be minimized.

As illustrated in Figure 3.3, the integration regions $\Omega_{\mathbf{nt}}$, $\Omega_{\mathbf{nb}}$ and $\Omega_{\mathbf{n}}$ typically form complex shapes and are even empty for certain combinations of $\mathbf{n}$, $\mathbf{t}$, and $\theta_{d_{\max}}$. Therefore, care must be taken when evaluating these integrals as discussed in the following section.

3.1.2 Validation

We performed numerous simulations designed to verify our key assumption that the symmetry distance is minimized at the correct surface frame and to evaluate the effect of $\theta_{d_{\max}}$. This was done using a standard Monte Carlo algorithm to estimate Equation 3.2 at a dense set of normal and tangent vectors. We compared the point of maximum symmetry to the correct surface frame for a number of analytic and measured BRDFs and for different values of $\theta_{d_{\max}}$ and surface orientations.

We first considered anisotropic materials using the elliptical Gaussian Ward model [Ward, 1992] along with three of the measured anisotropic samples acquired by Ngan et al. [2005].



For the measured data, we followed their reported procedure for fitting the tabulated microfacet distribution, Fresnel parameter, and shadowing and masking terms of the Ashikhmin model [2000], and these fits were used for our experiments. For our simulations, we fixed the view vector to be the z-axis, chose a normal vector within the zx plane, and fixed the tangent at an angle $\phi_t = 25°$ off this plane. We estimated the normal and tangent vectors that minimize Equation 3.2 and compared these to the correct surface frame.

Figure 3.4 reports the errors in these estimates for each material over a range of values of $\theta_{d_{\max}}$ (varying along the columns in each false-color error visualization) and a range of surface orientations (varying along the rows). These visualizations show two important trends. First, our ability to identify the correct normal and tangent diminishes as $\theta_{d_{\max}}$ diminishes. For very restrictive sets of light positions, most candidate normals result in a very small region of integration (see Figure 3.3), making it difficult to reliably identify the minimum. Therefore, we disregard estimates at pixels where we identify any of these domains to be empty. This limits the range of recovered normals to those less than $\theta_{d_{\max}}$ since $\Omega_{\mathbf{n}}$ is always empty beyond this point.

The second important trend is that the accuracy of our approach steadily increases as $\theta_{d_{\max}}$ reaches around 70° degrees and then falls off for larger values. We attribute this to the presence of Fresnel effects and the influence of shadowing and masking components in these regions of the BRDF as previously discussed. Based on this analysis, we chose $\theta_{d_{\max}} = 65°$ as the "sweet spot" (indicated by the white boundary in Figure 3.4) and used this for all the datasets we captured. Note that for these materials our approach typically identified the correct normal and tangent to within a single degree and only at extreme orientations did we observe errors of up to four degrees. These results are

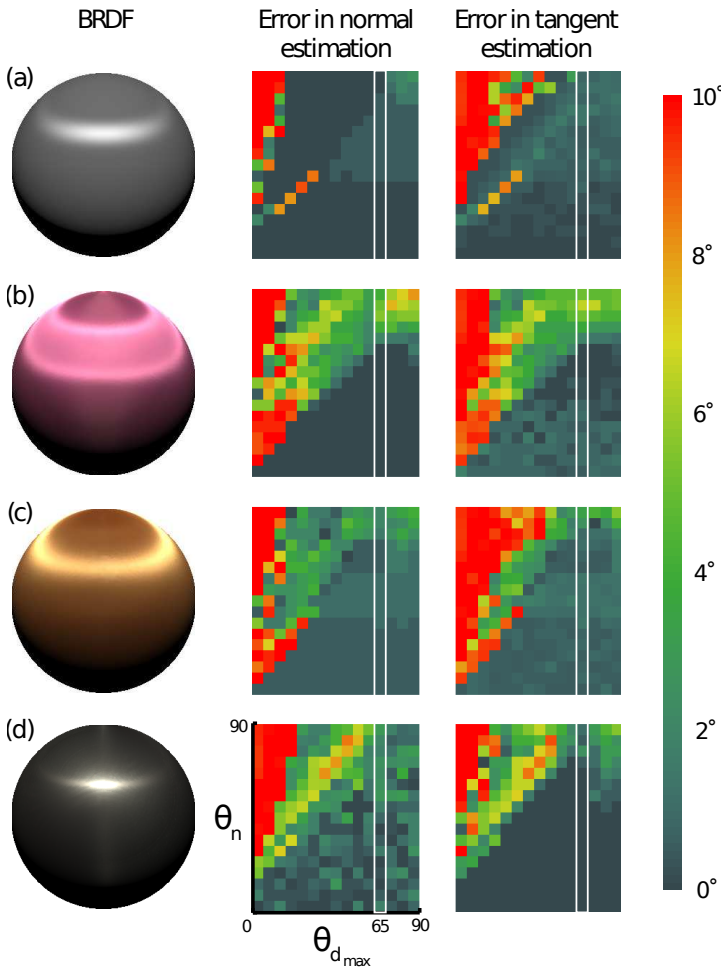


Figure 3.4: Error analysis using the Ward anisotropic BRDF and three measured samples [Ngan et al., 2005]. The errors in the estimated normal and tangent vectors are reported in degrees over a range of reflectance angles θ_n (the angle formed by the view and normal) which vary across the rows and values of $\theta_{d_{\max}}$ which vary across the columns. The materials include (a) an elliptical Ward with parameters $k_d = 0.5$, $k_s = 0.5$, $\alpha_x = 0.1$, and $\alpha_y = 0.5$.; (b) measured purple satin; (c) measured yellow satin; and (d) measured brushed aluminum.

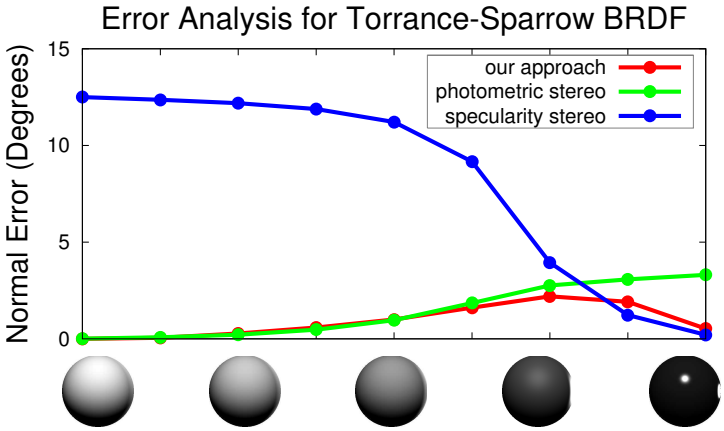


Figure 3.5: Average error computed for Lambertian photometric stereo with outlier rejection, our approach, and a method that associates the direction of maximum reflectance with the normal for different materials ranging from perfectly diffuse to dark and shiny modeled with the Torrance-Sparrow BRDF [1967].

representative of those we observed at different tangent vectors and we chose to present those for $\phi_t = 25°$ only for the sake of clarity.

We also experimented with isotropic materials using the Torrance-Sparrow BRDF [Torrance and Sparrow, 1967] with parameters ranging from perfectly diffuse to dark and shiny. The results are shown in Figure 3.5, where the average error was computed over normal orientations ranging from 0° to 60° and for $\theta_{d_{\max}} = 65°$ (recall that the value of $\theta_{d_{\max}}$ places a restriction on the range of normals that our method can reliably estimate). As a baseline comparison, we also report the average error for Lambertian photometric stereo [Woodham, 1980] and specularity stereo [Wang and Dana, 2006; Chen et al., 2006; Ma et al., 2007; Francken et al., 2008], which assumes the maximum reflectance occurs when the halfway vector is coincident to the normal. We applied photometric stereo to a simulated dataset consisting of 90×90 light positions sampled uniformly over the sphere and rejected data greater than two standard deviations away from the mean in order to remove areas in shadow or highlight. For the specularity method, we found the maximum of the product of the BRDF and cosine term using a non-linear search; this gives an upper-bound on the performance of these techniques since the angular sampling

rate of any practical acquisition device would further reduce accuracy. For all parameter values, our method performs well, and the surface normal is recovered to within a few degrees. Better performance is achieved for very specular surfaces, in which case the distribution term clearly dominates, and for very diffuse surfaces, in which case the $(\mathbf{n}\cdot\mathbf{l})$ terms cancel out only at the correct surface frame. While traditional photometric stereo and specularity-based methods perform well for diffuse and shiny materials respectively, our approach performs well in both cases. We found these trends and absolute errors to be consistent with other analytic models including Cook-Torrance [1981], Ward [1992], and He-Torrance-Sillion-Greenberg [1991].

3.2 Implementation

Our acquisition setup consists of a QImaging Retiga 4000R digital camera and a tungsten-halogen lightsource mounted to a computer-controlled spherical gantry [Levoy and Curless, 2002]. We assume distant lighting and an orthographic camera model. Note that a less precisely calibrated system would be possible if used in conjunction with methods for estimating the position of a hand-held light source, including free-form capture [Masselus et al., 2002; Chen et al., 2006; Toler-Franklin et al., 2007].

At each pixel in the reference view we reconstruct the BRDF slice $\beta_{\mathbf{v}}(\theta_h, \phi_h)$ using spherical barycentric interpolation over the measurements. As a pre-process, we compute the Delaunay triangulation [Shewchuk, 1996] of the halfway vectors corresponding to measurement locations projected onto the unit disc. Figure 3.6 (middle) visualizes the reconstructed BRDF slices at a single pixel in two anisotropic datasets. These slices are functions defined over the hemisphere shown here using a parabolic projection [Heidrich and Seidel, 1999] onto the plane.

In practice, the accuracy of this reconstruction is dependent on the frequency content

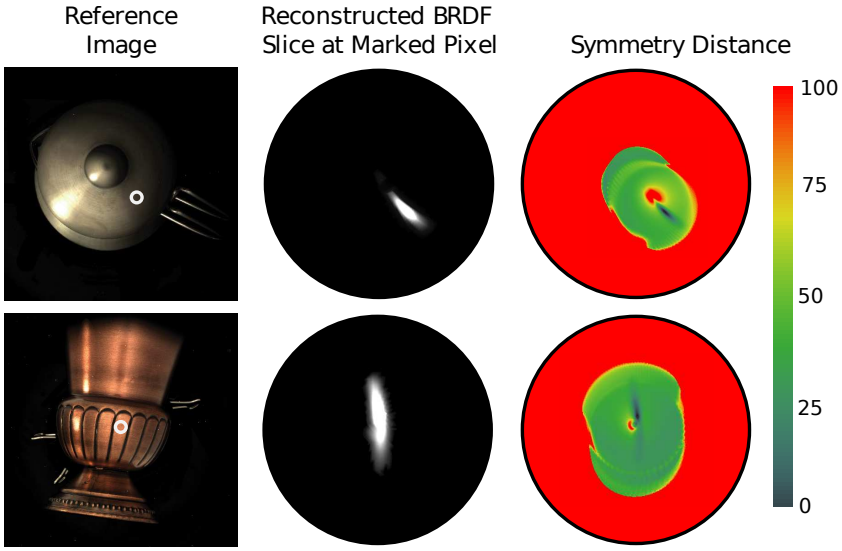


Figure 3.6: Typical reconstructions of the BRDF slice shown for two anisotropic datasets along with a visualization of the symmetry distance at a dense set of normal vectors. The BRDF slices and the symmetry values are functions defined over the hemisphere which are shown here using a parabolic projection onto the plane. Symmetry distances are normalized to occupy the range $[0, 100]$.

of the underlying BRDF, the density and pattern of light source positions, and the interpolation method. Each of the four datasets we acquired consists of approximately $1,500$ high dynamic range images recorded at light positions sampled uniformly over the cone of directions surrounding the view for which $\theta_d < 65°$. This acquisition process requires approximately 45 minutes. The impact of lower sampling densities is examined in section 3.4.

Given $\beta_{\mathbf{v}}(\theta_h, \phi_h)$, we compute the normal and tangent vectors in two steps. First, we find the normal vector that minimizes $\mathrm{SD}_{\tau_{\mathbf{n}}}$. Recall that this distance is independent of the choice of tangent direction even for anisotropic materials. This involves performing a non-linear optimization over a 2D domain for which we use the Nelder-Mead simplex algorithm [1965]. We initialize this search at the direction where we observed the maximum reflectance. Figure 3.6 (right) visualizes our estimate of $\mathrm{SD}_{\tau_{\mathbf{n}}}$ over a dense range of possible normal directions. The distance at each candidate normal $\mathbf{n}$ is computed

as a discrete sum over the sampled light positions. For each light position, we reflect its halfway vector across $\mathbf{n}$ and evaluate our reconstructed $\beta_{\mathbf{v}}(\theta_h, \phi_h)$ at that position. Reflected positions that fall outside $\Omega_{\mathbf{n}}$ are discounted. In cases where all of the light positions are reflected to areas without data, we assign an arbitrarily large error value (visible as the constant red area along the boundary in Figure 3.6). These visualizations clearly illustrate the stability of our optimization; the correct minimum is clearly defined and the error space is largely free of poor local minima. These properties result in a robust search that converges quickly, requiring only 60 evaluations of the error function on average.

In the second step, we hold the normal vector fixed and search for the tangent vector that minimizes $\mathrm{SD}_{\tau_{\mathbf{nt}}} + \mathrm{SD}_{\tau_{\mathbf{nb}}}$. Again, we use the Nelder-Mead algorithm to perform this 1D search. Because there exists a natural ambiguity between the binormal and tangent, we use a simple heuristic that assigns the tangent to the direction along which the specular highlight is widest.

We also found it was important to assign a confidence to each pixel based on the residual error and the magnitude of the denominator in Equation 3.2 at the computed frame. We consistently observed low confidence values in areas where the normal either lies beyond 60° from the view or where the reflectance data is corrupted by interreflections or cast shadows. We cull any estimate whose confidence is below a user-set threshold, which is determined manually for each dataset. Since each pixel may be processed independently, we distributed this optimization over a cluster of forty-two machines with Dual 1.6 GHz Opteron252 CPUs and 2GB of memory. Each dataset consists of $1,500$ images with 1024×1024 resolution which consume 2.3 GB of disk space. The optimization required approximately 10 minutes to complete or the equivalent of 7 hours on a single machine.

3.3 Results

Figure 3.7 shows the normal fields we estimated for two isotropic datasets. The first dataset is of a vase with an intricate spatially-varying design and a shiny surface finish. The second is of a small wooden figurine in the shape of a frog also with a spatially-varying appearance, but a more matte finish. We have masked out regions with confidence values below our chosen threshold (shown as a dotted line in the reference image). The few errors we observe are primarily caused by scene interreflections (e.g., near the top of the vase in Figure 3.7). Despite these errors, our algorithm is able to capture the overall shape of these objects even though they exhibit drastically different surface reflectance. Furthermore, our technique recovers fine geometric details such as the embossed pattern along the vase's surface and small indentations in the frog as seen in the cutouts.

Figure 3.8 shows normal and tangent fields computed from two objects with anisotropic reflectance. The first dataset is of the top of a french press coffeemaker made from aluminum brushed in a circular pattern. The second dataset is a bronze vase which has been brushed in a constant direction around its axis of symmetry. The tangent fields computed with our algorithm correctly capture these patterns, although there are regions where the material is not brushed and therefore the tangent direction is undefined (e.g., near the ridges in the middle of the vase). Additionally, errors in the normals often lead to errors in the tangents as seen near the bottom of the bronze vase.

3.3.1 Application: Appearance Editing

Our method provides separate estimates of the 3D shape of an object and its surface reflectance due to synchronous acquisition, allowing for independent manipulation of either component. Figure 3.9 shows an example in which the surface reflectance is

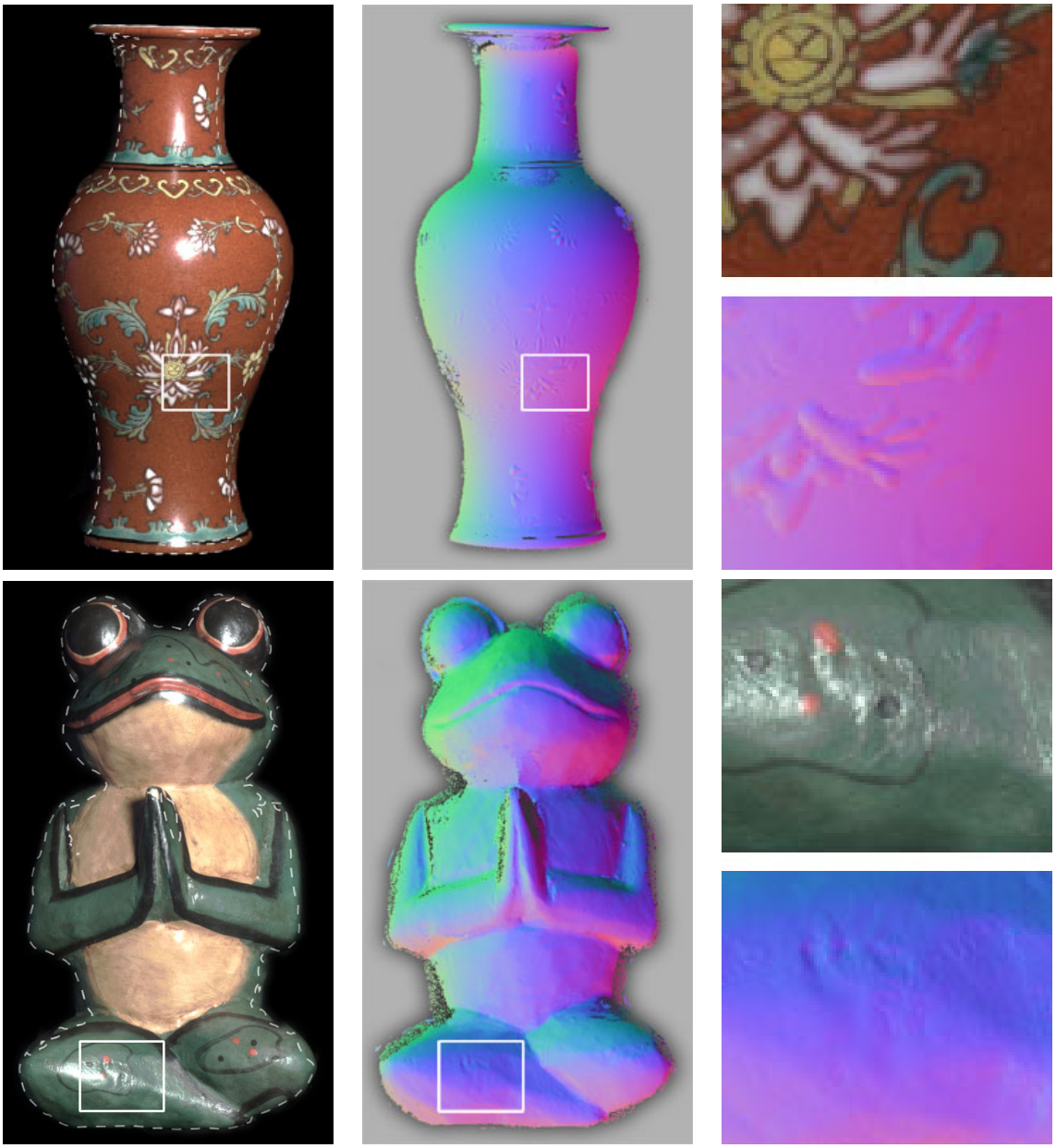


Figure 3.7: Normals computed from two objects with complex spatially-varying isotropic reflectance. Normals are visualized according to $(r, g, b) = (\frac{n_x+1}{2}, \frac{n_y+1}{2}, n_z)$. The dashed line in each reference image indicates the region shown at right.

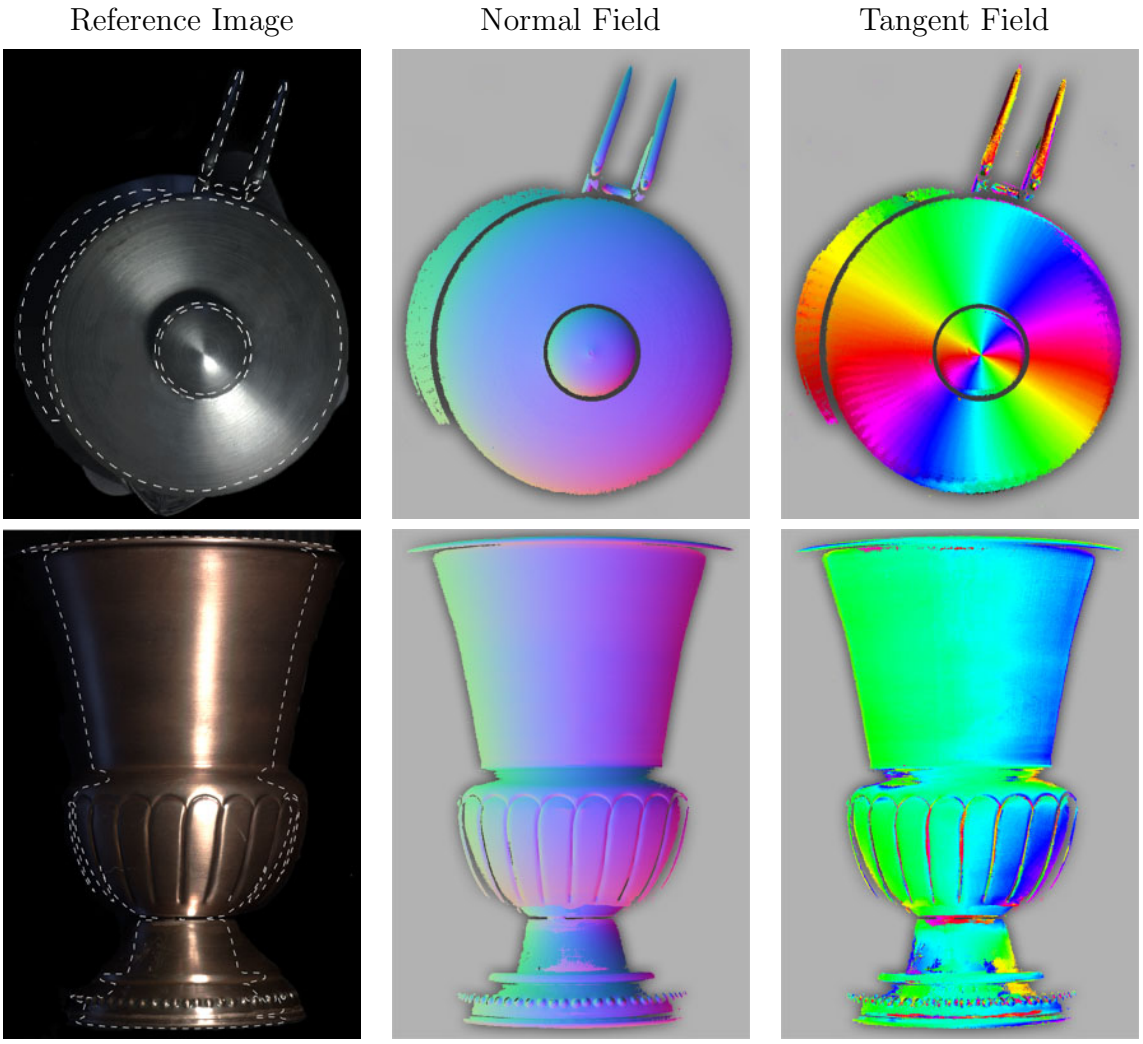


Figure 3.8: Normals and tangents computed from two objects with complex anisotropic reflectance. Tangent directions are visualized according to $(h, s, v) = (2\phi, 1, 1)$ where the angle ϕ is defined with respect to the x-axis.

edited to enhance surface gloss. In this example, we enable control over the reflectance by independently fitting an analytic Ward BRDF to the reflectance measurements available at each surface position. A comparison of this fit to a reference image suggests that the Ward model provides a reasonable description of this surface's reflectance, at least for this particular slice of the BRDF domain. The right of this figure shows an editing example in which we have scaled the specular and diffuse components at each pixel to give the finish a glossier appearance.

Since we use a parametric BRDF model for editing, this approach is conceptually similar to what could be achieved using the method of Goldman et al. [2005]. An important difference, however, is that our approach recovers geometry in a manner that is separate from reflectance. This provides the ability to choose editing tools that are most appropriate for the reflectance of a particular object. These tools can be based on a parametric model appropriate to the object—as we have demonstrated here—or on data-driven representations similar to those described by Lawrence et al. [2006].

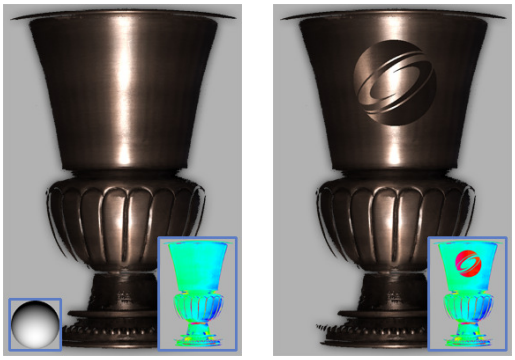
It is also possible to manipulate the underlying normal and tangent fields of a surface while keeping the original reflectance intact. Figure 3.10 shows a relighting result that uses an edited version of the bronze vase's tangent field to display the SIGGRAPH logo. This edit was accomplished by rotating each tangent vector within the edited region by 90° about its corresponding surface normal; Fisher et al. [2007] have recently introduced tools that enable more sophisticated manipulations of tangent fields.

3.4 Discussion

While we have shown that our strategy of associating planes of symmetry in the BRDF with the correct surface frame is valid for many different materials, we did observe cases where this approach fails. A good example is the red velvet sample acquired by

Reference Image Ward Fits Edited Fits

Figure 3.9: Edited surface reflectance combined with measured geometry. **Left:** original HDR image not included in training set. **Middle:** reconstruction from fits of the Ward BRDF model computed at each pixel. **Right:** result of editing the specular component to give the surface a glossier appearance.



Original Edited Tangents

Figure 3.10: Edited tangent field combined with measured reflectance. **Left:** original HDR image not included in training set. **Right:** image generated by editing the tangent field to display the SIGGRAPH logo.

Ngan et al. [2005]. The deduced microfacet distribution does not exhibit the expected symmetries and we calculated that our technique would produce errors of up to ten degrees. Intuitively, materials with microgeometries symmetric around the surface normal (grooves in brushed metal, threads in satin) satisfy our underlying assumption, whereas materials with asymmetric microgeometries such as velvet (composed of tiny cylinders systematically brushed to a fixed angle off the normal) do not. Retroreflective materials would also present challenges for our approach. These BRDFs exhibit backscattering

lobes that are substantial when the light and view directions are close to one another. However, we expect it would be possible to handle these cases by placing a lower-bound on $\theta_{d_{\max}}$ during acquisition to avoid sampling these regions.

Accurately identifying the planes of symmetry in $\beta_{\mathbf{v}}(\theta_h, \phi_h)$ depends strongly on our ability to reconstruct this function from our sampled data. Figure 3.11 shows the effect on the resulting normal field of using a decreasing number of light source positions. As expected, with fewer source positions the quality of these reconstructions decrease and the resulting normal field exhibits a faceted appearance as the optimization more often returns a poor local minimum. Using a more sophisticated interpolation technique such as radial basis functions could result in better reconstructions, and is an interesting direction of future work.

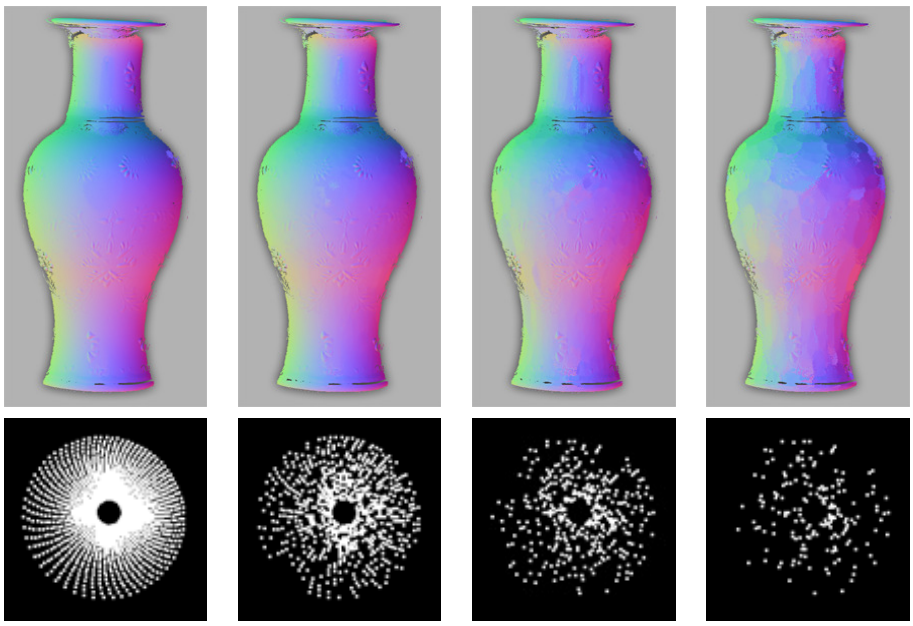


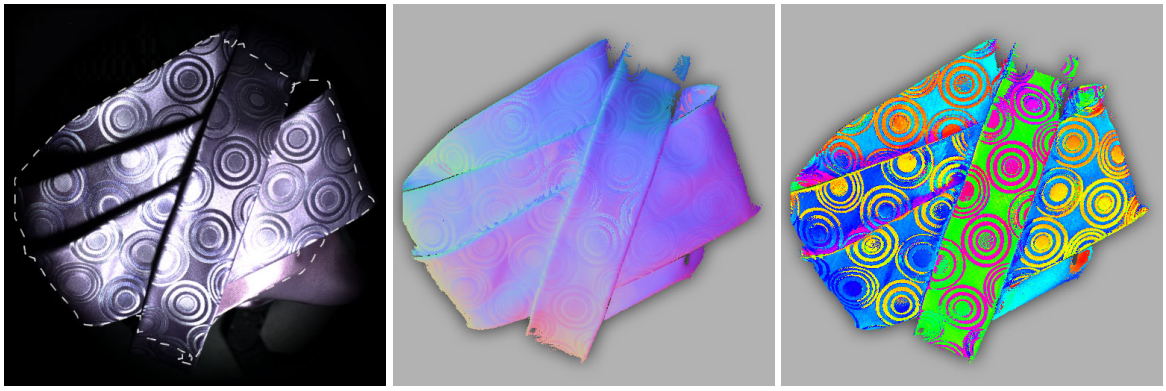
Figure 3.11: Normal maps of a porcelain vase computed using different numbers of light positions. The images along the bottom visualize the source positions showing the corresponding half-way vector projected onto the x-y plane of a view-centered coordinate system.

Another limitation of our technique is that it does not account for scene interreflec-

tions or subsurface scattering. Figure 3.12 visualizes the normal and tangent map of a silk necktie computed using our approach. Note that this scene exhibits strong discontinuities in the tangent field and many depth boundaries. The majority of errors are caused by interreflections which occur near the overlapping regions. Future research might consider combining our approach with techniques for isolating light that is directly scattered from the object surface such as that presented by Nayar et al. [2006]. We should also note that in some places our heuristic for choosing between the tangent and the binormal is unreliable (e.g., the abrupt blue-to-yellow transition in the tangent map visualization in the bullseye pattern near the bottom left). This is due to the fact that the highlights for this material are relatively wide and sparsely sampled near grazing angles. Despite these few inconsistencies, however, the majority of this challenging scene is accurately captured by our method.

It is also worth emphasizing that there are a number of other photometric stereo methods that could be used in conjunction with our approach. For example, in the isotropic case, symmetries along intensity profiles from view-centered circles of light directions provide an additional constraint on the surface normal [Alldrin and Kriegman, 2007]. Rather than explore the benefits of combining complimentary approaches with our symmetry-based technique, we instead explored what can be recovered using these symmetries alone. This allows us to consider isotropic, anisotropic, diffuse, specular and hybrid surfaces in a unified manner.

Finally, there are many applications for which partial and even imprecise surface geometry is sufficient, such as the applications described in Section 3.3. We believe that the simplicity of the acquisition required for our technique along with its applicability to a wide range of materials will have broad applicability in similar systems.



Reference Image Normal Field Tangent Field

Figure 3.12: A folded necktie with complex reflectance, depth boundaries and discontinuities in the tangent field presents a particularly challenging case.

3.5 Conclusion

This chapter described a new photometric approach for estimating the surface normals and, for anisotropic materials, the tangent vectors over an object's surface as seen from a fixed viewpoint. The input to our algorithm is a set of images taken under varying point lighting constrained to lie within a cone of directions centered at the view. We show that the size of this cone can be chosen to allow reliable detection of the planes of reflective symmetry in the microfacet distribution of the measured BRDF. We demonstrated an optimization that computes the surface frame at each pixel by identifying these planes of symmetry and showed that our approach is accurate to within a few degrees over a range of analytic and measured BRDF data. Finally, we presented normal and tangent fields computed with our technique from four real-world objects that encompassed a range of complex isotropic, anisotropic, and spatially-varying reflectance.

Chapter 4

Sinusoidal Illumination for Shape and Appearance Measurement

First presented in Holroyd et al. [2010a; 2010b] at SIGGRAPH 2010, and Holroyd and Lawrence [2011] at CVPR 2011.

Sinusoidal illumination is used as an active lighting pattern for recovering geometry in active stereo triangulation systems. In contrast to discrete patterns such as binary stripes, the continuous nature of sinusoidal patterns allows for subpixel precision when recovering geometry. More recently, it was discovered that high-frequency illumination patterns can be used for *optical descattering*, which computationally separates local surface reflections that retain the high-frequency pattern from diffuse interreflections and subsurface-scattering that do not. Nayar et. al [2006] observed that a high-frequency sinusoid could be used to simultaneously perform optical descattering while recovering 3D geometry. The advantage of this approach is that it removes interreflections, which are traditionally difficult to detect and can create produce errors during triangulation.

In this chapter we show that sinusoidal illumination patterns are also useful for *appearance capture*, because they simultaneously separate local and non-local reflections

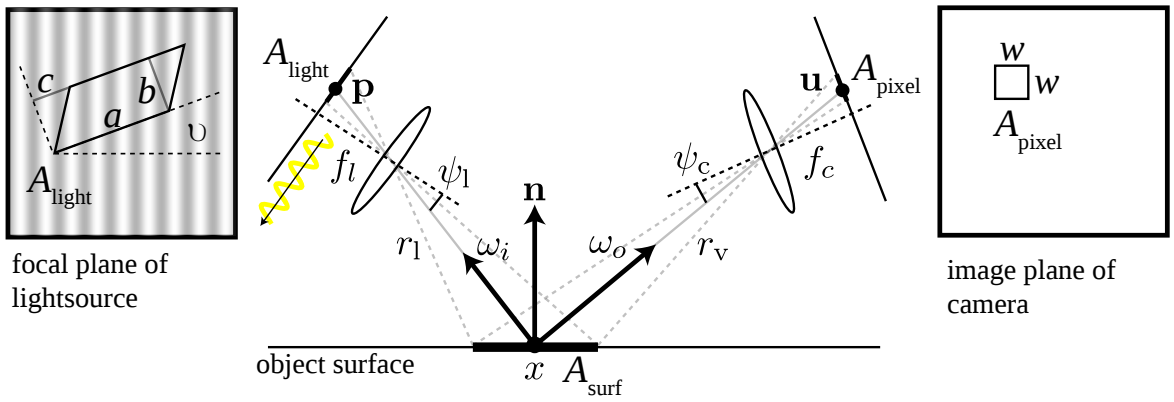


Figure 4.1: A square pixel in the camera projects, via a planar surface patch, to an oriented parallelogram in the focal plane of a focused light source. We are interested in the relationship between the amplitude of the translating sinusoidal radiance pattern on the focal plane of the light source and the apparent amplitude at the camera's pixel. Note that the regions A_{light} and A_{pixel} are not drawn to scale.

and allow the synchronous recovery of surface geometry. We derive a closed-form expression describing the relationship between the BRDF and the resulting sinusoid's apparent amplitude, validate it experimentally, and discuss its implications.

In addition, using optical triangulation methods to measure the shape of translucent objects has previously been difficult because subsurface scattering contaminates measurements of the direct reflection at the surface. Optical descattering allows isolating this direct component, which in turn enables accurate estimation of the shape of translucent objects. Despite encouraging results in recent work, there is currently no rigorous mathematical analysis of the expected error in the measured surface as it relates to the parameters of these systems: the frequency of the projected sinusoid, the geometric configuration of the source and camera, and the optical properties of the target object. We present the first such analysis, which confirms earlier empirical results and provides a much needed tool for designing 3D scanners for translucent objects.

Table 4.1: Summary of notation used in this chapter

x	scene point
$\mathbf{u} = (u, v)$	point on image plane of camera
$\mathbf{p} = (p, q)$	point on focal plane of lightsource
$d\mathbf{u}, d\mathbf{p}, dS$	differential areas on the image plane, source plane, and surface
$\mathbf{n}$	surface normal
ω_i, ω_o	local light and view directions
$\rho(\omega_i, \omega_o)$	BRDF
$\sigma(\mathbf{u})$	camera sensitivity function
$l(x)$	lightsource emission function
f_l, f_c	focal lengths of lightsource and camera
ψ_l, ψ_c	angles from optical axes of lightsource and camera
r_l, r_c	distances from scene point to lightsource and camera centers
$A_\mathrm{pixel} = w^2$	area of square camera pixel
A_surf	area of corresponding region on surface
$A_\mathrm{light} = ab$	area of corresponding parallelogram on source focal plane
c	signed distance that characterizes the degree of skew in parallelogram
A_l, A_c	areas of lightsource and camera apertures
Ω_l, Ω_c	solid angles subtended by lightsource and camera as seen by surface
$E_s(\cdot), E_c(\cdot)$	irradiance at the surface, and image plane of camera
$L(\cdot), L_s(\cdot)$	radiance emitted at the lightsource, and surface
$\Phi(\cdot)$	power (flux)

4.1 Preliminaries

Consider the system in Fig. 4.1. A thin-lens camera observes a planar surface patch that is illuminated by a custom light assembly, and this light assembly consists of a planar Lambertian area source placed at the focal plane of another thin lens. The area source in this light assembly produces radiance patterns that are shifted horizontal sinusoids with fixed frequency f, amplitude A, and DC offset G. The shifts are represented by a discrete set of phase values: $\{\phi_k\}_{k=1\ldots M}$, so we can write the radiance at a point $\mathbf{p} = (p, q)$ on the focal plane of the source as

$$L^k(p, q) = A\cos(2\pi f(p + \phi_k)) + G, \quad k = 1 \ldots M. \tag{4.1}$$

Illumination from the source is focused at a point x on a planar surface patch, and this patch is observed by a thin-lens camera, which is also focused at x. A pixel (or any square region) on the image plane that is centered at the projection of x and has dimensions $w \times w$ measures flux due to the radiance from a neighborhood of the point x on the surface, and assuming that the camera is a linear device, the intensity recorded at the pixel is proportional to this flux. Under the sinusoidal illumination of equation 4.1, the pixel response can be written in the form,

$$I^k(\mathbf{u}) = \alpha(\mathbf{u}) \cos(\gamma t_k + \phi(\mathbf{u}))) + \beta(\mathbf{u}), \quad k = 1 \ldots M, \tag{4.2}$$

where $\mathbf{u} = \Pi_c(x)$ is the projection of x (the center of the pixel), and $\gamma = 2\pi f(t_{k+1} - t_k)$, the product of the spatial frequency and the displacement of the sinusoid between consecutive shifts. This relation plays a central role in phase mapping techniques (e.g. [Srinivasan et al., 1985]), since the *apparent phase* $\phi(\mathbf{u})$ provides information about the depth of the surface along the ray that is back-projected from pixel $\mathbf{u}$.

Presently, we are interested in the *apparent amplitude* $\alpha(\mathbf{u})$ since, as we will show, it provides information about the local surface reflectance (the BRDF at x) and can be used for reflectometry. We show that, in addition to the BRDF, this expression depends on the intrinsic parameters of the lightsource and camera, as well as their positions and orientations relative to the surface.

4.2 Relating the projected and observed sinusoid intensity

We obtain flux incident on the image plane from radiance emitted from the focal plane of the source using three basic relations.

Image irradiance from surface radiance. The image irradiance $E_c(\mathbf{u})$ that is due to radiance $L_s(x, \omega_o)$ at surface point x in direction ω_o is given by the familiar thin-lens equation,

$$E_c(\mathbf{u}) = \sigma(\mathbf{u}) L_s(x, \omega_o). \tag{4.3}$$

Here, $\sigma(\mathbf{u})$ models optical fall-off in the camera, which for a thin lens depends on the area of the camera's aperture, its focal length, and the angle ψ_c between its optical axis and the ray through $\mathbf{u}$: $\sigma(\mathbf{u}) = \sigma(\psi_c) = A_c \cos^4(\psi_c)/f_c^2$. (See Fig. 4.1 and Table 4.1 for summaries of notation.) More generally, this sensitivity function can be measured during a radiometric calibration procedure, and we assume it to be known.

Emitted radiance from incident irradiance at the surface. This relation simply follows from the definition of the bi-directional reflectance distribution function, or BRDF [Nicodemus et al., 1977]:

$$L_s(x, \omega_o) = \rho(\omega_i, \omega_o) E_s(x). \tag{4.4}$$

We assume that the scattering properties are statistically uniform over the small area A_{surf} observed by a single pixel, and that following the isolation of local reflections [Nayar et al., 2006], light transport within the material occurs over distances that are small relative to this area.

Surface irradiance from radiance on the source focal plane. This relation is less familiar, so we provide a derivation. The basic idea is to compute the power emitted from the focal plane of the source toward its lens, and then divide this by the differential surface area δS to obtain surface irradiance. The underlying assumption is that δS is in focus, so that all of the power that reaches the lens arrives at δS.

Let L be the emitted radiance at the center of a differential patch $\delta\mathbf{p}$ on the focal plane of the source assembly. The power received by the lens is this radiance multiplied by the

differential area foreshortened in the direction of travel and the solid angle subtended by the lens as seen by this small area:

$$\Phi = L\delta\mathbf{p} \cdot \cos\psi_{\mathrm{l}} \cdot \frac{A_{\mathrm{l}} \cos\psi_{\mathrm{l}}}{(f_l^2/\cos^2\psi_{\mathrm{l}})} = L\delta\mathbf{p} \cdot \frac{A_{\mathrm{l}} \cos^4\psi_{\mathrm{l}}}{f_l^2}.$$

Since it is in focus, all of this power arrives at a differential area on the surface δS, and the surface irradiance is $E_s = \Phi/\delta S$. The ratio of areas can be obtained by equating the solid angles subtended by $\delta\mathbf{p}$ and δS as seen by the center of the lens,

$$\frac{\delta\mathbf{p}}{\delta S} = \frac{f_l^2}{r_l^2} \frac{(\mathbf{n} \cdot \omega_i)}{\cos^3\psi_{\mathrm{l}}}, \tag{4.5}$$

and combining these expressions yields the desired relationship:

$$E_s = A_{\mathrm{l}} \cos\psi_{\mathrm{l}} \cdot L \cdot \frac{(\mathbf{n} \cdot \omega_i)}{r_l^2}.$$

Analogous to the camera sensitivity function described above, in practice we generalize this expression to

$$E_s = l(x)L(\mathbf{n} \cdot \omega_i), \tag{4.6}$$

where $l(x)$ is an *emission function* that can be measured during a radiometric calibration process and is assumed to be known. In fact, as described in chapter 5, we find it beneficial to incorporate the r_l^2 term into this function and thus allow it to vary over the three-dimensional working volume: $l(x)$.

Equipped with equations 4.3, 4.4, and 4.6, we are ready to proceed. The power received by the finite pixel area A_{pixel} centered at image point $\mathbf{u}_o$ (and thus the response

of that pixel) is the integral of the image irradiance, which by equation 4.3 is

$$\Phi_{\text{pixel}} = \sigma(\mathbf{u}_o) \int_{A_{\text{pixel}}} L_s(\Pi_c^{-1}(\mathbf{u}))d\mathbf{u},$$

where $\Pi_c^{-1}(\cdot)$ is the back-projection of image point $\mathbf{u}$ onto the surface. (In this expression, we have assumed the sensitivity function to be constant over the pixel.) Using an expression analogous to equation 4.5, we change variables to integrate over the observed area (A_{surf}) on the surface instead, and this yields

$$\Phi_{\text{pixel}} = \sigma(\mathbf{u}_o) \frac{f_c^2}{\cos^3 \psi_{\text{c}}} \frac{(\mathbf{n} \cdot \omega_o)}{r_c^2} \int_{A_{\text{surf}}} L_s(x)dS,$$

when the pixel is small enough for the angles and distances (ψ_{c}, r_c^2, ω_o) to be constant over its extent.

Finally, the surface radiance is related to the radiance on the source focal plane through equations 4.4 and 4.6, and making these substitutions along with another change of variables, $dS \rightarrow d\mathbf{p}$, gives

$$\Phi_{\text{pixel}}^k = \sigma(\mathbf{u}_o) l(\mathbf{p}_o) \frac{f_c^2 \cos^3 \psi_{\text{l}}}{f_l^2 \cos^3 \psi_{\text{c}}} \frac{(\mathbf{n} \cdot \omega_o)}{r_c^2} \rho(\omega_i, \omega_o) \int_{A_{\text{light}}} L^k(\mathbf{p})d\mathbf{p}. \tag{4.7}$$

The last term in this expression is the integral of the sinusoidal radiance pattern (equation 4.1) over an area A_{light} that is obtained by projecting the square pixel A_{pixel} onto the planar surface and into the lightsource. The area A_{light} is a quadrilateral that, due to the small extent of a single pixel, is relatively unaffected by perspective distortion and is well-approximated by a parallelogram, as depicted in Fig. 4.1 and Fig. 4.2. Any such parallelogram is completely described by the lengths of its sides, a and b, it's orientation ξ relative to the horizontal axis of the source focal plane, the position $\mathbf{p}_o = (p_o, q_o)$ of its center, and it's skew, which we parameterize by the *signed* distance c shown in Fig. 4.2.

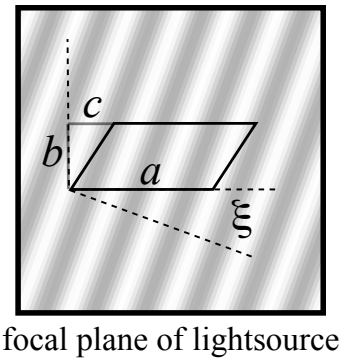


Figure 4.2: The source region A_{light} corresponding to a single camera pixel, depicted in a transformed coordinate system (p', q') that aligns one side of the region with a coordinate axis (compare to the left of Fig. 4.1). The shape of this parallelogram is determined by the lengths of its sides, a and b, and its skew, which is characterized by the *signed* distance c. These parameters, along with the position and orientation ξ relative to the horizontal axis of the source focal plane can be computed from the known surface geometry and the parameters of the lightsource and camera.

In the present case, all of these parameters are determined by the camera and source parameters and the surface position and orientation, and since all of these are known, we can compute the parallelogram parameters corresponding to any pixel in our camera.

In order to compute the integral over this region, we first change the coordinate system to be aligned with side b and have one corner as its origin as shown in Fig. 4.2. We use (p', q') for these new coordinates, which allow us to write the integral as

$$\int_{A_{\text{light}}} L^k(\mathbf{p}) d\mathbf{p} = \int_0^b \int_{\frac{c}{b}p}^{\frac{c}{b}p+a} G + A\cos\left(2\pi f(p'\cos\xi - q'\sin\xi + p_o + \phi_k)\right) dp' dq'.$$

By twice using the identity $\int \cos(sx+t)dx = sin(sx+t)/s$, performing trigonometric manipulations, and using the expression $\text{sinc}(x) = \sin(\pi x)/(\pi x)$, we obtain

$$ab(A\,\text{sinc}(af\cos\xi)\,\text{sinc}(bf\sin\xi + cf\cos\xi)\cos(\pi f(b\sin\xi + a\cos\xi + c\cos\xi + p_o + \phi_k)) + G).$$

Now, substituting this into equation 4.7, we obtain an expression for the flux Φ^k_{pixel}

(and thus the pixel response) that is in the desired form of equation 4.2. From this it follows that the apparent amplitude of the observed sinusoid in the camera satisfies

$$\boxed{\alpha(\mathbf{u}_o) \propto \sigma(\mathbf{u}_o) l(x) \rho(\omega_i, \omega_o)(\mathbf{n} \cdot \omega_i) A \ \mathrm{sinc}(af \cos\xi) \ \mathrm{sinc}(bf \sin\xi + cf \cos\xi),} \qquad (4.8)$$

where we have used the fact that the total area of the parallelogram $A_{\text{light}} = ab$ is given by

$$ab = w^2 \frac{(\mathbf{n} \cdot \omega_i) r_c^2 f_l^2 \cos^3 \psi_c}{(\mathbf{n} \cdot \omega_o) r_l^2 f_c^2 \cos^3 \psi_l}.$$

4.3 Implications for reflectometry

The result in equation 4.8 confirms that the apparent amplitude measured at each camera pixel $\alpha(\mathbf{u})$ (Section 5.2 in the main paper) is proportional to the product of the surface irradiance under point lighting $l(\mathbf{x})(\mathbf{n} \cdot \omega_i)$ (recall that $1/r_l^2$ is captured by $l(\mathbf{x})$), the amplitude of the projected sinusoidal illumination A, the BRDF $\rho(\omega_i, \omega_o)$, and the camera sensitivity $\sigma(\mathbf{u})$. In addition, it also predicts a less obvious effect that we call "amplitude loss" whereby the measured response is inversely proportional (via the sinc functions) to the product of the pixel width w and the frequency f of the source radiance pattern in addition to the relative orientations and distances between $\mathbf{x}$ and the camera and source. In words, if either f or w increase (holding everything else fixed) the measured amplitude will decrease at a rate predicted by the product of the sinc functions in Equation 4.8 and eventually reach zero — this corresponds to the point at which the sine pattern is no longer visible in the image. Similarly, as the camera approaches a grazing view of the surface in a direction perpendicular to that of the sine wave with an overhead source held fixed or the distance from $\mathbf{x}$ to the camera r_v increases, then the measured amplitude will similarly decrease. This makes equation 4.8 a useful analytic tool for estimating the lower bounds on f for a particular experimental

setup and, more importantly, it allows converting the amplitude measured at the camera into measurements of the surface BRDF.

The graphs in Fig. 4.3 confirm this effect and validate the analytic model derived above. They show the amplitude measured at a camera pixel as it moves toward the horizon $(\mathbf{n} \cdot \omega_o) \rightarrow 0$ with it's up vector in the epipolar plane and for a stationary overhead light $(\mathbf{n} \cdot \omega_i) = 1$. For this "in-plane" configuration, the parallelogram reduces to a rectangle ($c = 0$). The three graphs correspond to different orientations of the sine wave with respect to the plane of motion. These graphs also include predictions by a numerical simulation that agree with our analytic model exactly — we have verified that this is true beyond the in-plane case illustrated here. Note that this amplitude loss can be significant. In the case where $\xi = 0$, a roughly 20% decrease in the amplitude is observed at an elevation angle of 60 degrees which falls off to roughly 90% at 80 degrees.

In chapter 5, we use measured amplitudes $\alpha(\mathbf{u})$ for reflectometry, and correct for this effect by dividing by the terms on the right-hand side in equation 4.8 in order to isolate the BRDF $\rho(\omega_i, \omega_o)$.

4.4 Recovering the geometry of translucent objects

The most common approach for extending optical triangulation methods to handle translucent objects is to first isolate the direct reflection at the object surface before recovering geometry. One set of techniques is based on the idea that the specular component of the surface response is due to light that reflects directly off of the object surface while the diffuse component is produced by subsurface scattering. Chen et al. [Chen et al., 2006] apply a simple threshold to a histogram of the measured intensities in order to identify specular highlights, which are subsequently used to estimate a high-resolution normal field. A similar set of methods rely on color heuristics to estimate normal fields

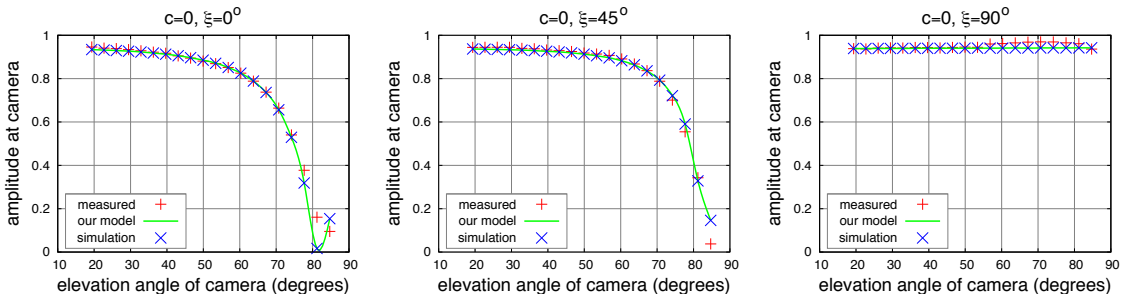


Figure 4.3: Measurements and simulation results confirming the "amplitude loss" phenomenon and validating our analytic model. Each graph shows the amplitude measured at a single camera pixel as the view direction approaches the horizon $(\mathbf{n} \cdot \omega_o) \to 0$ with the light source held fixed overhead $(\mathbf{n} \cdot \omega_i) = 1$. The camera's up vector remains in the epipolar plane, causing the parallelogram A_{light} to reduce to a rectangle ($c = 0$). For the measurements, we imaged a Spectralon board and corrected for deviations from a perfect Lambertian reflector. We used the parameters of our experimental setup (focal lengths, pixel size, stand-off distances, etc.) to perform a numerical simulation of the amplitude measured at the camera and to evaluate our model in equation 4.8. We observed very close agreement between measured data, our simulation, and our model.

of human faces [Weyrich et al., 2006] and repair errors in a structured light scan [Barsky and Petrou, 2001]. These methods generally suffer in the presence of glossy highlights (as opposed to sharp specular highlights), objects with a small specular response, or objects whose diffuse color is similar to that of the illuminant. Additionally, they require a very dense sampling of light directions in order to observe the specular highlight everywhere over the object surface.

If the light striking a surface is polarized, the portion that is directly reflected at the surface typically retains this polarization, whereas subsurface scattering acts to depolarize the light [Tyo et al., 1996] (see figure 4.4). Under this assumption, the direct component of the scene can be isolated by computing the difference between images taken behind parallel and perpendicular polarization filters. This approach has been used in combination with sinusoidal illumination for recovering the geometry of translucent objects [Chen et al., 2007]. Ma et al. [Ma et al., 2007] extend this idea to use circularly polarized spherical gradient illumination in order to recover dense normal fields

of translucent objects such as human faces. However, a limitation of these approaches is that, for certain materials, light that has been scattered multiple times may still retain a significant amount of polarization and other materials are known to depolarize light that is reflected at the surface [Chen et al., 2007].

Figure 4.4: Polarized light that reflects directly off a surface typically retains its polarization. A common photography technique is to apply a polarizing filter to remove direct reflections off bodies of water.

Another approach to the separation of direct and indirect illumination uses optical descattering. Light that scatters multiple times within the material is significantly diffused whereas light that scatters only once retains the projected pattern and can be isolated. Chen et al. [Chen et al., 2008] present a method that uses sinusoidal patterns both for the purpose of isolating the direct surface reflection and for recovering depth using a standard phase unwrapping technique, a special case of a broader class of phase profilometry methods [Srinivasan et al., 1985]. Gupta et al. [Gupta et al., 2009] analyze the effect of defocus on this descattering process and show how to both correct for it and leverage it for the purpose of depth recovery. More recent work uses the light transport equation to estimate each component of the indirect response through a recursive procedure [Mukaigawa et al., 2010].

We build on this prior work and our analysis in 4.2 by providing a radiometric analysis in the specific case of sinusoidal illumination. This analysis demonstrates the nature and degree of error in the estimated surface for stereo triangulation systems that use sinusoidal illumination patterns.

4.5 Depth estimation using active sinusoidal illumination

Active stereo triangulation systems that use sinusoidal lighting capture a sequence of images of a scene while it is illuminated by a high-frequency sinusoidal pattern. Between each exposure, the sinusoid is translated within the lightsource's focal plane by a known amount. A key observation about this approach is that summing together multiple sinusoids with the same frequency produces a sinusoid that also has that frequency:

$$\sum_i A_i \cos(ft + \theta_i) + G_i = A\cos(ft + \theta) + G. \tag{4.9}$$

Regardless of the properties of the scene, due to the linearity of light transport and the closure of sinusoids under addition, the intensities $\{I_i | i \in [1, n]\}$ measured at each camera pixel will trace a time-varying sinusoid, $I_i = A\cos(t_i + \phi_0) + G$, where t_i can be calculated from the spatial frequency of the projected sinusoid and the magnitude of the translation in the i^{th} image. This fact makes sinusoidal lighting especially desirable for scanning translucent objects: *the signal measured at the camera has a simple relationship to the signal produced by the source without having to explicitly account for global illumination*

in the scene. The per-pixel amplitude, phase, and DC offset can be computed by solving:

$$\begin{bmatrix} \cos(t_1) & -\sin(t_1) & 1 \\ & \vdots & \\ \cos(t_n) & -\sin(t_n) & 1 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} I_1 \\ \vdots \\ I_n \end{bmatrix} \tag{4.10}$$

$$A = \sqrt{a_1^2 + a_2^2} \qquad \phi_0 = \arctan(a_2/a_1) \qquad G = a_3$$

The initial phase offset at each pixel ϕ_0 can be used to identify corresponding rays between the source and camera and estimate geometry in a process known as phase profilometry [Srinivasan et al., 1985]. However, many pixels along each epipolar line will observe the same phase value. Resolving this "phase ambiguity" is a key problem and a number of solutions have been developed (a good review is provided by Salvi et. al [2010]). In the following, we will not address this aspect of phase profilometry and assume the ambiguity can be resolved using one of the available methods. Instead, our goal is to analyze the effect of subsurface scattering on the recovered phase value at each pixel and how this in turn biases the estimated depth.

4.6 Phase offet for translucent objects

As illustrated in Figure 4.5, we will assume the scene is composed of a homogeneous flat translucent surface, which ignores effects due to local curvature or thin shells. The translucent medium is characterized by the absorption coefficient σ_a, scattering coefficient σ_s, and extinction coefficient $\sigma_t = \sigma_a + \sigma_s$, as well as the medium's index of refraction η and phase function $p(\cdot, \cdot)$. A camera images the scene while the only source of illumination is a light source modulated to produce a spatially-varying sinusoid (a "projector"). Translating this projected pattern produces time-varying sinusoids along

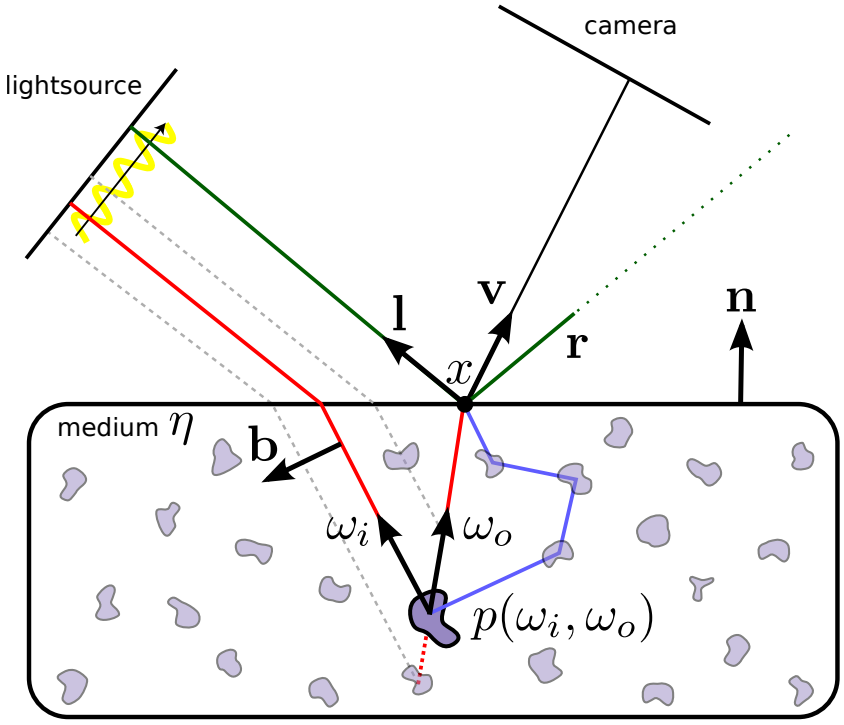


Figure 4.5: Setup and notation used in this section. The vector b is perpendicular to ω_i and in the direction of the sinusoid projected into the scene. The function $p(\omega_i, \omega_o)$ is the material's phase-function. Note that ω_i and ω_o refer to the incoming and outgoing vectors **inside the medium**, unlike in other sections.

individual rays as they leave the projector and strike the surface. Both the projector and camera are assumed to be orthographic.

Paths of light can reach the camera along the view direction $\mathbf{v}$ either due to surface reflection (green), single scattering within the material (red), or multiple scattering (blue). We ignore surface reflection by assuming the camera is not positioned along the mirror direction $\mathbf{r}$. At the surface, the Fresnel equations predict the amount of light that is transmitted into the medium and Snell's law predicts the direction it travels due to refraction, $-\omega_i$. Light that enters the object is eventually either absorbed or scattered until it exits the medium. Scattering occurs according to the material's phase function $p(\omega_i, \omega_o)$, which characterizes the angular distribution of light scattered when striking a

particle in the medium.

Some portion of the incident light scatters exactly once in the direction ω_o and exits the medium at point x traveling in the view direction $\mathbf{v}$ towards the camera. Another portion will scatter multiple times before exiting at x in this same direction. The key assumption in optical descattering [Nayar et al., 2006; Chen et al., 2007; 2008] is that the contribution made by multiple scattering to the exiting light traveling along $\mathbf{v}$ is unaffected (constant) under translations of the projected sinusoid. Therefore, it only affects the DC offset of the time-varying sinusoid measured along the camera ray and can be ignored (we describe a simulation we performed to validate this important assumption in Section 4.8.1). Building on this prior work, we will assume that only single scattering contributes to the amplitude and phase of the time-varying sinusoid measured along the camera ray.

Let L_o represent the response measured by the camera along this ray. An analytic expression for L_o can be obtained by integrating the contribution due to single scattering along the camera ray as it travels through the medium. Let F_η represent the percentage of light lost at the surface due to Fresnel effects while entering and exiting the medium, $f_{\sin}$ represent the spatial frequency of the sinusoid, and ϕ_0 represent the initial phase offset of the time-varying sinusoid at point x. We integrate along s, the total distance light travels through the medium – from the point of entrance, to the scattering event, and to the point of exit. Then, L_o can then be written as

$$F_\eta p(\omega_i, \omega_o) \int_{s=0}^{\infty} e^{-\sigma_t s} \cos(f_{\sin} \frac{(\mathbf{n} \cdot \mathbf{l})(b \cdot \omega_o)(\mathbf{n} \cdot \omega_i)}{(\mathbf{n} \cdot \omega_i) + (\mathbf{n} \cdot \omega_o)} s + \phi_0)$$

The rate at which the projected sinusoid oscillates *along the direction* ω_o depends both on the spatial frequency of the sinusoid $f_{\sin}$ as well as the sinusoid's direction of variation within the medium $\mathbf{b}$, which is always perpendicular to ω_i.

We use the relationship

$$\int_{s=0}^{\infty} e^{-\sigma s}\cos(As+\theta) = \frac{\cos(\theta+\arctan(A/\sigma))}{\sqrt{A^2+\sigma^2}},$$

to derive a closed-form expression for L_o:

$$\boxed{L_o = \frac{F_\eta p(\omega_i,\omega_o)}{\sqrt{A^2+\sigma_t^2}}\cos(\phi_0+\arctan(A/\sigma_t))} \tag{4.11}$$

where

$$\boxed{A = f_{\sin}\frac{(\mathbf{n}\cdot\mathbf{l})(\mathbf{b}\cdot\omega_o)(\mathbf{n}\cdot\omega_i)}{(\mathbf{n}\cdot\omega_i)+(\mathbf{n}\cdot\omega_o)}}.$$

Note that the "phase error" $\Delta\phi = \arctan(A/\sigma_t)$. If $\Delta\phi = 0$ then the same phase produced along the projector ray $\mathbf{l}$ will be precisely measured along the camera ray $\mathbf{v}$ that intersects it at point x. When this quantity is not zero, this triangulation will intersect at some point away from x, introducing an error in the estimated depth. Naturally, the phase error depends on the geometric setup of the projector and camera (through A), the extinction coefficient σ_t, and indirectly on the index of refraction η (which determines ω_i and ω_o).

4.7 Implications for recovering geometry

The phase error in equation 4.11 has a number of important implications for scanning translucent objects. The phase error will impact recovered geometry differently based on the specific experimental setup and reconstruction algorithm that is used. Perhaps the simplest and most common setup to consider uses one calibrated projector and one calibrated camera, often referred to as "structured lighting". In this case, the geometric error depends on the phase offset scaled by the period of the sinusoid along the camera

ray v. Specifically, the geometric error ϵ is equal to

$$\epsilon = \frac{\Delta\phi}{2\pi}\frac{(1 - \mathbf{v} \cdot \mathbf{l})}{f_{\sin}}. \tag{4.12}$$

This relationship allows one to establish upper bounds on ϵ as a function of the material properties and system parameters. For example, we can answer the question: how is ϵ affected by the spatial frequency of the sinusoid? Figure 4.7 visualizes this relationship by plotting ϵ, measured in millimeters, for five different materials over a range of sinusoid frequencies (the graph actually shows the period). Since the maximum error is bounded by one period of the sinusoid, and assuming the phase ambiguity can be resolved perfectly, ϵ tends to 0 as the frequency of the sinusoid approaches infinity. Of course, as the frequency increases the ability of the camera to resolve it also decreases (discussed in section 4.7.2). Additionally, note that for optically dense materials such as marble, lower-frequency sinusoids have less of an impact on ϵ, because light rays are not able to penetrate as far into the medium.

Another interesting question is: for a given projector location, where should the camera be placed to minimize ϵ? This could be used, for example, to optimize a specific setup or to drive a view planning algorithm. Figure 4.8 shows upper bounds on ϵ derived from equation 4.11 for a fixed projector located 45° off the surface normal, and view directions that vary in the plane formed by the surface normal and direction to the projector (i.e., $\mathbf{l}$ in figure 4.5). Note that as the camera and projector become coaxial, each camera ray integrates along a matching projector ray so that the phase error is zero; however, the ability to robustly triangulate two rays diminishes rapidly as the baseline becomes very small.

Finally, we visualize the manner in which ϵ depends on the material properties. Figure 4.9 plots ϵ for a specific camera/projector configuration and a sinusoidal frequency

Figure 4.6: Legend for the different materials we consider in section 4.7. Colors correspond to: $\sigma_t = 0.1\text{mm}^{-1}$ (ocean water), 0.2mm^{-1} (chicken broth), 0.5mm^{-1} (potato), 1.0mm^{-1} (skin), 2.0mm^{-1} (marble). These are approximate values based on measurements published by Jensen et al. [2001].

of $2\pi/5.0\,\text{mm}^{-1}$. Other frequencies follow a similar trend, although the absolute error varies as seen in figure 4.7.

4.7.1 Relation to phase unwrapping

A common method for solving the phase ambiguity problem is to project progressively lower-frequency sinusoidal patterns into the scene, until each point receives a unique phase offset [Ghiglia and Pritt, 1998]. For translucent objects, figure 4.7 gives some insight into when this "phase unwrapping" may fail. Note that modifying the sinusoid's frequency can easily produce a large difference in the phase error in certain cases. If care is not taken in choosing these progressively lower-frequency sinusoids, each iteration could make the problem of localizing the correct period more difficult. We expect equation 4.11 will help strengthen these techniques when they are applied to translucent surfaces.

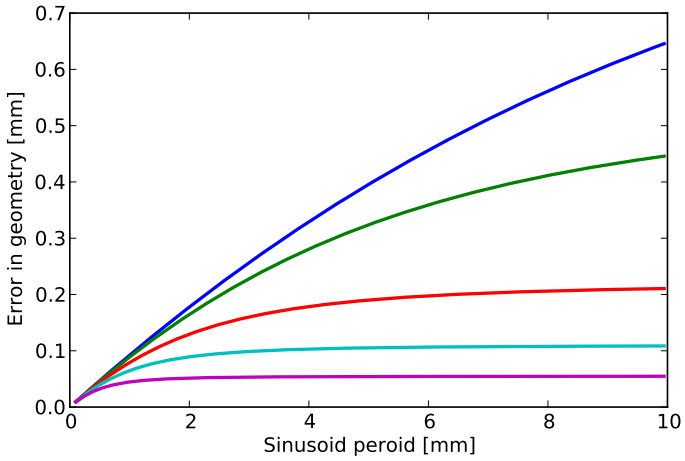


Figure 4.7: Geometric error ϵ for a structured lighting setup. Light and view are at $\pm 45°$ across the surface normal. The colored lines correspond to the materials in figure 4.6. Ocean water ($\sigma_t = 0.1\text{mm}^{-1}$) is the most translucent and has the largest phase offset error (blue).

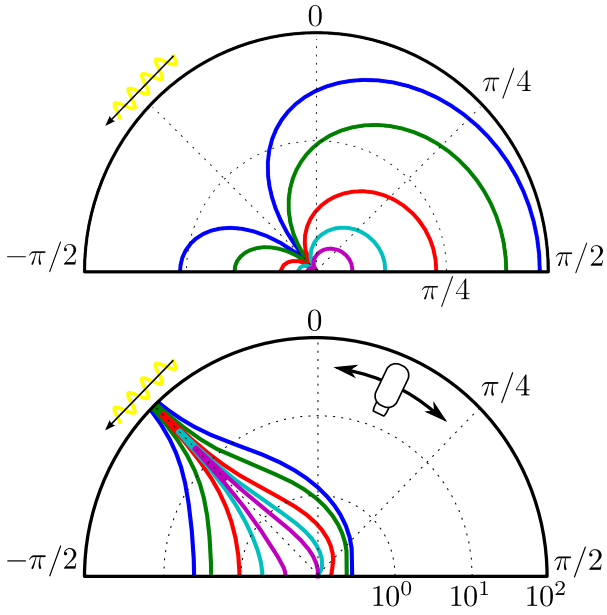


Figure 4.8: Phase error (top) and geometric error (bottom, in mm, log-scale) for a light at 45° elevation angle and view directions from $-90°$ to $90°$. The index of refraction $\eta = 1.55$, and the sinusoid frequency is $2\pi/5.0\,\text{mm}^{-1}$, modulated spatially in the plane formed by the light and camera (worst case). The colored lines correspond to the materials in figure 4.6.

4.7.2 Amplitude loss

Although it is possible to decrease the phase error arbitrarily by increasing the frequency of the sinusoid (figure 4.7), at some point this process breaks down. Each camera pixel integrates over a finite region of the surface. As shown in section 4.3, this can be modeled as a convolution of the scene radiance with a compact kernel, which has the effect of diminishing the measured amplitude of the time-varying sinusoid. At some point it is no longer possible to reliably measure the amplitude and, consequently, the phase. Combined with our previous result, this analysis provides a tool for choosing the optimal frequency in these types of systems: one that is as small as possible to reduce $\Delta\phi$ while not becoming impractical due to amplitude loss.

4.8 Validation

We validated equation 4.11 using a volumetric ray-tracer [Jakob, 2010] to simulate a structured light setup (one projector and one camera). This simulation included both single and multiple scattering, and was used to verify our analysis over a wide range of parameters including various indices of refraction η and phase functions $p(\omega_i, \omega_o)$. Figure 4.9 compares our simulated data to the predictions made by equation 4.11 for a range of extinction coefficients.

4.8.1 Multiple scattering

Although light that is scattered multiple times within a translucent object contributes significantly to its overall *appearance*, for many materials it has a negligible effect on the *phase error*. We performed an additional series of simulations to validate this assumption. We first computed the phase at one camera pixel based on single scattering only, followed by a full simulation that includes multiple scattering. The projector and cam-

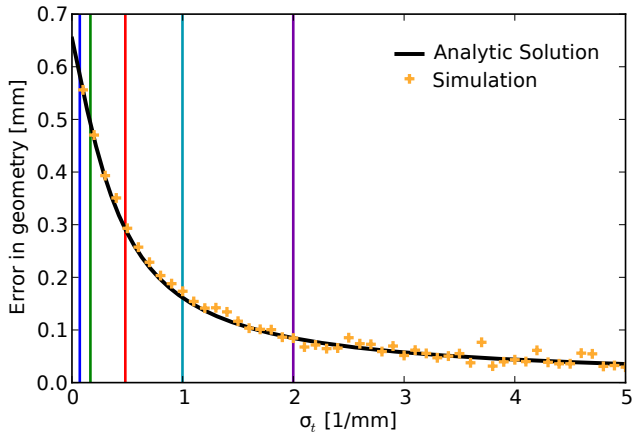


Figure 4.9: Geometric error ϵ as a function of the material extinction coefficient for a light $-30°$ and view $45°$ off of the surface normal and a sinusoidal frequency of $2\pi/5.0\,\mathrm{mm}^{-1}$. The analytic solution is derived from equation 4.11 and is compared to a simulation of this setup using a physically-based volumetric ray-tracer [Jakob, 2010].

era were both located $\pm 45°$ off the surface normal, and the phase function was assumed to be constant (i.e., the parameter $g = 0$ in the Henyey-Greenstein function [Henyey and Greenstein, 1940]). We ran these simulations over a range of extinction coefficients $\sigma_t \in [0.1, 2.0]$ and spatial frequencies $f_{\sin} \in [0.04, 2.0]$. In all of these cases, the difference between the phases produced for single scattering only and single+multiple scattering were on average 0.007 radians apart with a standard deviation of 0.005. This offers further support for our choice to assume that the phase is predominantly affected by single scattering.

Chapter 5

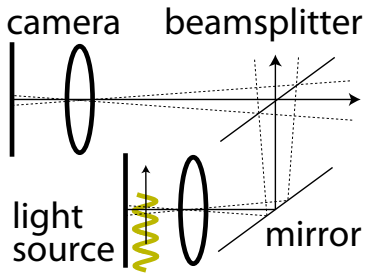
A Coaxial Optical Scanner for Synchronous Acquisition of 3D Geometry and Surface Reflectance

First presented in Holroyd et al. [2010a] at SIGGRAPH 2010.

Previously proposed systems for measuring 3D shape and reflectance suffer from a number of drawbacks that have limited the development of practical scanners. First, many systems make restrictive assumptions about the BRDF, and these lead to errors in the output that are difficult to detect and characterize. Second, systems often use separate sensors to measure reflectance and geometry, and this leads to a difficult and error-prone 2D-3D data registration problem that can cause a reduction in accuracy. Third, most existing systems involve a series of fragile calibration steps that limit usability and degrade the quality of the recovered model. Finally, previous designs often fail to correctly account for global illumination, such as interreflections and sub-surface scattering, which contaminate measurements of the desired *local* surface reflectance.

We address these issues with a simple optical setup and processing pipeline that build

on recent advancements in multiview stereo, phase-based profilometry, and light descattering. The basic building block in our design is a digital camera and a high-frequency, spatially-modulated sinusoidal light source aligned to share a common focal point and optical axis.



Using at least two of these assemblies, it is possible to capture a sequence of images of an object from different viewpoints under sinusoidal illumination originating from different locations. Combining a new active multiview stereo algorithm and the theoretical analysis of chapter 4; we show how these images allow recovering precise high-resolution estimates of object shape *and* local surface reflectance. Additionally, we show how several scans captured with our system can be aligned and merged into a single watertight model using existing techniques.

We analyze several models acquired by our scanner, including those of objects with challenging material properties such as very shiny spatially-varying surface reflectance. We show that even in these difficult cases, the measured geometry is accurate to within 50 microns at 200 micron resolution, and the BRDF measurements agree with reference data to within 5.5% over a wide range of angular configurations. Finally, we show that images rendered using the acquired models agree very well with reference images at view and light positions that are far away from those initially measured. We conclude with a discussion of the larger design space of scanners that may benefit from these coaxial devices.

5.1 System Overview

The basic building block of our system, which is depicted in Figure 5.1, is a coaxial camera/lightsource assembly. This produces dynamically-modulated lighting by coupling a stable source with a translating glass slide, and it effectively co-locates this modulated lighting with a camera through a beamsplitter. Our prototype system uses two of these assemblies, and it achieves view and lighting variation using a spherical gantry.

This design is the result of a unified consideration of the three desiderata of the previous section. Each of the two basic components—the slide-based high-frequency sinusoidal lighting and the coaxial camera/lightsource pairs—addresses them in multiple ways.

Coaxial cameras and illumination. Coaxial cameras and light sources have been previously used for related applications including 3D reconstruction [Lu and Little, 1999; Zickler et al., 2002; Zhang et al., 2006] and measuring BRDFs and general reflectance fields [Han and Perlin, 2003; Garg et al., 2006; Ghosh et al., 2007]. We use them in our design for a number of reasons.

First, these assemblies simplify calibration by allowing us to leverage recent feature-based camera calibration algorithms [Brown and Rusinkiewicz, 2005; Furukawa and Ponce, 2008] to automatically calibrate the geometric aspects of both the cameras *and the light sources*. Once the assembly is carefully manufactured, the sub-pixel camera calibration information that is obtained by these automatic methods gives the pose of the accompanying source without any additional effort. In particular, this eliminates the need for mirrored spheres or other scene fiducials of known size and position [Lensch et al., 2003b].

Second, by using two of these coaxial devices, we can easily capture a reciprocal image pair in which the center of projection of the camera and source are swapped [Zickler

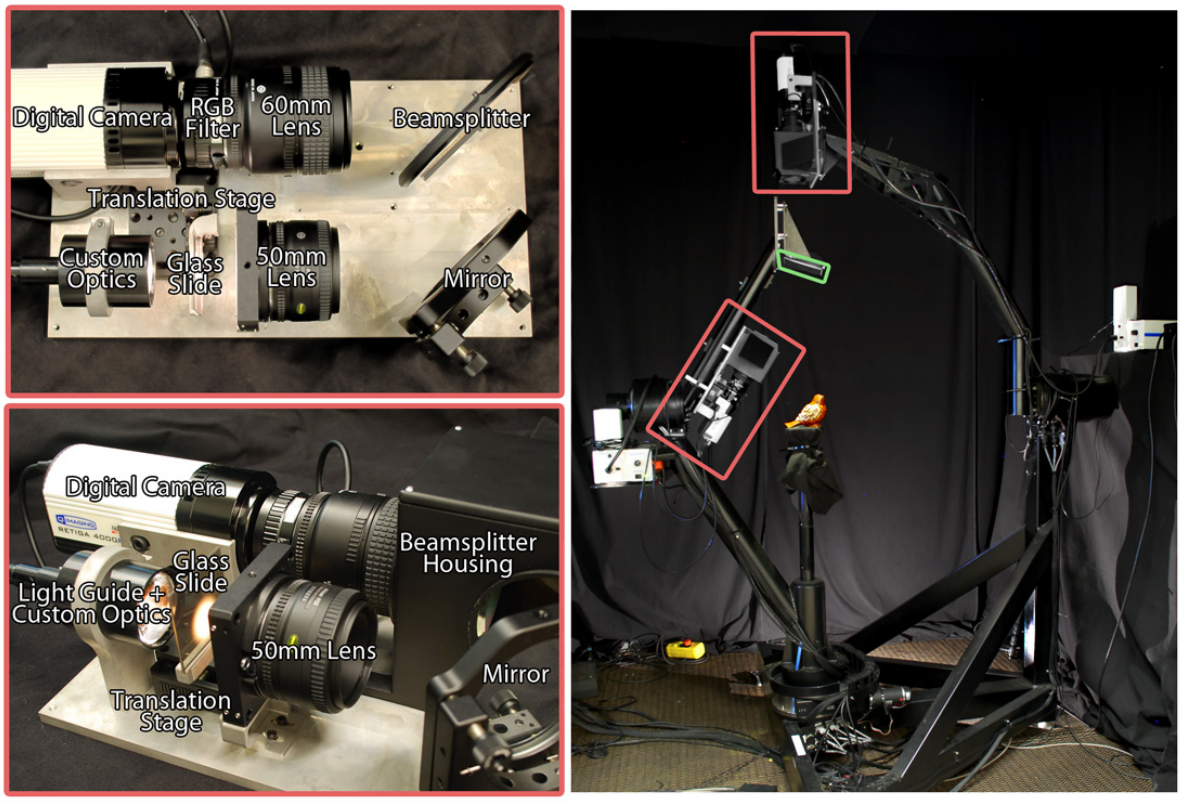


Figure 5.1: Our prototype scanner consists of two identical coaxial camera/lightsource assemblies (left) mounted to the arms of a four-axis spherical gantry (right). The coaxial device mounted on the left arm views the object through a mirror (green) to achieve roughly equal stand-off distances between the two arms. The light is provided by tungsten-halogen sources with computer-controlled mechanical shutters (white boxes in image on right).

et al., 2002]. Reciprocal images allow us to enhance feature-based camera calibration for surfaces with strong view-dependent specular effects [Zickler, 2006]. Perhaps more importantly, surface locations occluded in one camera view correspond to shadowed locations in the reciprocal view. Therefore, these pairs allow one to reason about visibility before any information is known about the geometry, and allows robustly recovering per-pixel geometry through multi-view triangulation. Another important property of reciprocal images is that they provide a constraint on the surface normal by virtue of the reciprocity of BRDFs. However, we choose not to incorporate this constraint in our system due to its sensitivity to the radiometric properties of the device and the difficulty of applying it near high-frequency variation in the reflectance such as edges in the material albedo [Guillemaut et al., 2004].

Third and finally, unlike many designs, our use of coaxial devices and descattering allows measuring surface backscattering, which can be an important attribute of appearance.

Modulated lighting with a glass slide. While it is common to use digital projectors for modulated lighting, we instead create it by coupling a fixed stable light source with a moving glass slide. This avoids the many non-idealities associated with projectors, such as vibration, shallow depth of field, limited resolution, light leakage, screen door effects, and brightness instability (see [Zhang et al., 2006]), and it improves the long-term stability of any radiometric calibration.

An important attribute of our system is that it extracts both shape and reflectance information from lighting that is modulated by a single high-frequency sinusoid, which can be positioned very precisely and repeatably by a translation stage. Like many phase mapping systems, we use shifted sinusoidal illumination patterns to estimate the phase offset at a dense set of surface locations in a reference camera [Srinivasan et al., 1985]. This phase information yields surface depth, but only up to a discrete choice (a 2π

ambiguity). One way to resolve this ambiguity is to project a sequence of lower-frequency patterns, but achieving this "phase unwrapping" reliably often requires capturing many images or, for highly scattering materials, using cross polarization [Chen et al., 2007] that reduces signal strength; or more sophisticated lighting patterns [Chen et al., 2008] that are hard to implement without a projector. Instead, we leverage the fact that we collect images from multiple viewpoints, and we introduce a technique for resolving the phase ambiguity through triangulation. A quantitative analysis of our results shows that this method provides depth with sub-millimeter precision (often tens of microns) for a wide variety of objects. Additionally, we show that after correcting for the amplitude loss discussed in section 4.3, our system can recover BRDF measurements from objects with complex geometry accurate to within 5% of reference measurements.

The remaining sections provide a detailed account of our setup and measurement process before presenting an evaluation of our prototype scanner.

5.2 Experimental Setup

Figure 5.1 shows two photographs of our prototype coaxial assemblies on the left and our complete scanner on the right. Each assembly consists of a QImaging Retiga 4000R 12-bit monochrome camera with 1024x1024 resolution coupled to a Varispec tunable LCD RGB filter and a 60mm Nikkon macro lens. Using an external color filter avoids mosaicing artifacts, but requires taking a separate exposure for each color band. The light source is a tungsten-halogen fiber illuminator connected through a randomized light guide. Each illuminator contains a computer-controlled mechanical shutter that allows interrupting the flow of light onto the guide. The incoming light path is focused onto a glass slide using a custom focusing assembly, and this slide is located at the focal plane

of a 50mm Nikkon lens.[1] The slide is affixed to a translation stage that steps within the focal plane in 3 micron increments which are repeatable to within 3 microns.

The camera and light paths are aligned using a four-inch pellicle beamsplitter. We experimented with several alternatives and found that a large pellicle beamsplitter produced the cleanest images under co-axial lighting (see the top left image in Figure 5.7). On the downside, these attract dust and produce spectral artifacts in the emitted light due to interference at the membrane that must be accounted for during radiometric calibration. We estimate that the camera and light paths diverge by no more than 0.05° over our working volume.

One assembly is mounted to each arm of a computer-controlled four-axis spherical gantry (Figure 5.1 right) which allows sampling the full sphere of directions to within 0.1 degree of precision around a working volume 25cm in diameter. The stand-off distance to each device is roughly 1m. For this working volume and camera resolution, we sample the object surface at 0.2mm intervals.

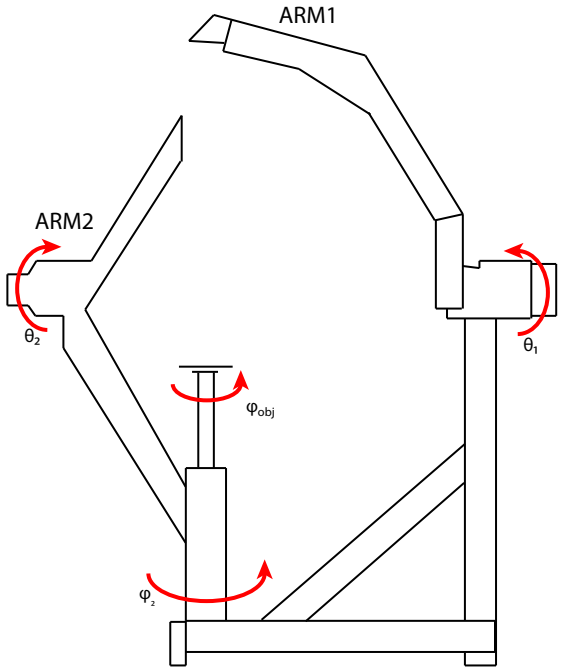
5.2.1 Geometric Calibration

The spherical gantry is capable of precise angular positioning; however, the position of each device relative to the object must be known as well. In addition, to compute a mapping from 3D points to camera pixels, we must know the camera intrinsics including focal length, principal point, radial and tangential distortion coefficients.

In this section we describe a pipeline for generating a rough calibration with about 5 pixel reprojection error, which is later improved for each scan using the automatic feature-based refinement procedure described in section 5.3.3 to subpixel precision. Note that no further steps must be taken to estimate the positions of the light sources since

[1] Although the same lens would ideally be used along both the camera and light path, physical constraints of our setup prevented this.

these are co-located with the cameras. We found this rough geometric calibration to be stable and did not need to readjust the alignment of our coaxial assemblies even after months of use.



ϕ_{obj}	azimuthal angle of object platform
θ_1	elevation angle of ARM1
ϕ_2	azimuthal angle of ARM2
θ_2	elevation angle of ARM2

Figure 5.2: Notation for gantry angles. Each angle is separately controlled by a servo, and moves independently from the others. Note: ϕ_2 only moves ARM2, not the object platform.

Using photographs of a stationary checkerboard pattern, we identify corner points in ever image and perform a non-linear optimization that searches for system parameters that best predict those corner locations. The initial geometric calibration results in a 4×4 transformation matrix $\mathbb{T}(\phi_{\text{obj}}, \theta_1, \phi_2, \theta_2)$ that takes points in world coordinates and converts them to points in a normalized camera coordinate system, as well as the camera intrinsics that allow us to project a point in normalized camera coordinates into

the image.

The best way to conceptualize this matrix is as a combination of many smaller homographies that take us between simply related coordinate systems. The world coordinate system is defined during the one-time calibration process by a checkerboard placed in an arbitrary position on the object platform. No assumption is made about the checkerboard's alignment, but the size of each square and the fact that it is planar are known.

Homographies

The origin of the world coordinates (defined to be the first corner on the checkerboard) is at a distance r_{obj} away from the object platform's center of rotation.

$\mathbf{R}_{\text{world\_platformOrbit}}$

The first homography is between the world coordinate system and the PlatformOrbit coordinate system shown in figure 5.3. The PlatformOrbit coordinate shares the same origin with the world coordinates, but is rotated by $R_{\text{world\_platform}}$ such that the x-axis points directly away from the object platform's axis of rotation, a convention used throughout calibration to define canonical coordinate systems. This rotation has 3 degrees of freedom, and does not depend on any gantry angles, only on the placement of the checkerboard relative to the axis of rotation. Note that we do not assume the z-axis of the checkerboard is parrallel to the axis of rotation.

$\mathbf{R}_{\text{platformOrbit\_platform}}$, $\mathbf{T}_{\text{platformOrbit\_platform}}$

The platform coordinate system is defined to have its origin centered along the axis of rotation, the z-axis pointed along the center of rotation (unchanged from the platformOrbit coordinate system) and its x-axis pointed such that when $\phi_{\text{obj}} = 0$, it is parallel with the platform coordinate system's x-axis (in other words, when $\phi_{\text{obj}} = 0$, there is no rotation between the two coordinate systems) as shown in figure 5.3. By defining the coordinate system this way, the rotation and translation are completely defined by

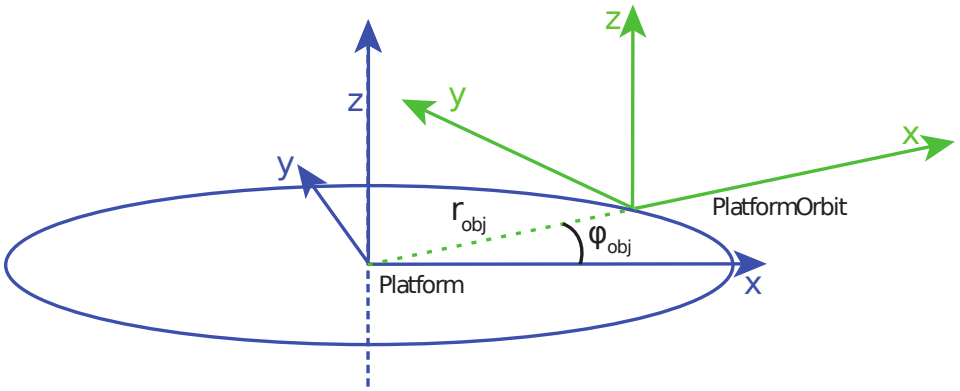


Figure 5.3: Relationship between the platform coordinate system (centered at the object platform's center of rotation) and the PlatformOrbit coordinate system (with the same origin as the world coordinates, but rotated so the y axis points out from the circle).

the radius between the checkerboard origin and the axis of rotation (and of course the known ϕ_{obj}):

```
double l = calib.platformorbit_radius;

Mat4d platformOrbitTOplatform = homography_axes(
          Vec3d(  cos(phi_obj),  sin(phi_obj), 0.),
          Vec3d( -sin(phi_obj),  cos(phi_obj), 0.),
          Vec3d(            0.,            0., 1.),
          Vec3d(l*cos(phi_obj),l*sin(phi_obj), 0.));
```

Figure 5.4: C code for computing the platformOrbit to platform homography.

$\mathbf{R}_{\text{platform\_arm1}}$, $\mathbf{T}_{\text{platform\_arm1}}$

This is the first arm specific coordinate system. The arm1 coordinate system is defined just like the platform coordinate system, but for the axis of rotation made by arm1. Specifically, the origin is at the center of rotation, the z-axis points along the axis of rotation (away from the arm itself) and the x-axis is defined such that when $\theta_1 = 0$, it coincides with the platformOrbit coordinate system we will introduce next (it points toward the camera's position when the arm is at rest). The rotation and transformation have 3 degrees of freedom each, since the arm1 rotation need not have any relation to

the platform's rotation.

$\mathbf{R}_{\text{arm1\_orbit1}}$, $\mathbf{T}_{\text{arm1\_orbit1}}$

This homography is very similar to the platformOrbit to Platform homography. It depends only on θ_1 and the radius of arm1's rotation. Figure 5.4 applies directly, but with θ_1 substituted for ϕ_{obj}, and we need to take the inverse. Arm1 is defined with the x-axis pointed away from the center of rotation, and the z-axis parallel to the arm1 coordinate system's z-axis. This coordinate system shares its origin with the camera's focal point.

$\mathbf{R}_{\text{orbit1\_cam1}}$

The final rotation takes the arm1 orbit and rotates it to be aligned with the camera. It has 3-degrees of freedom. After this final transformation we have arrived at our desired coordinate system.

Full tree

The full tree of homographies is shown in figure 5.5. The first two homographies are shared between both arm1 and arm2, and the rest are separate. Note that there is an additional transformation we have not discussed, which takes platform to platform2. The platform2 coordinate system is simply rotated backwards by ϕ_2 around the z-axis. Essentially, rotating arm2 around the platform is identical to rotating the platform itself in the opposite direction. For example, if we rotate both arm2 and the platform 10 degrees clockwise, there is no net effect on the relationship between the checkerboard and cam2.

It is important to have a good initial guess for all the unknowns so that the non-linear optimization will converge quickly, and with a valid position (for example, avoiding negative radii). Here we step through the initial guess for each homography, which were obtained by rough measurements with rulers and assuming a particular placement of

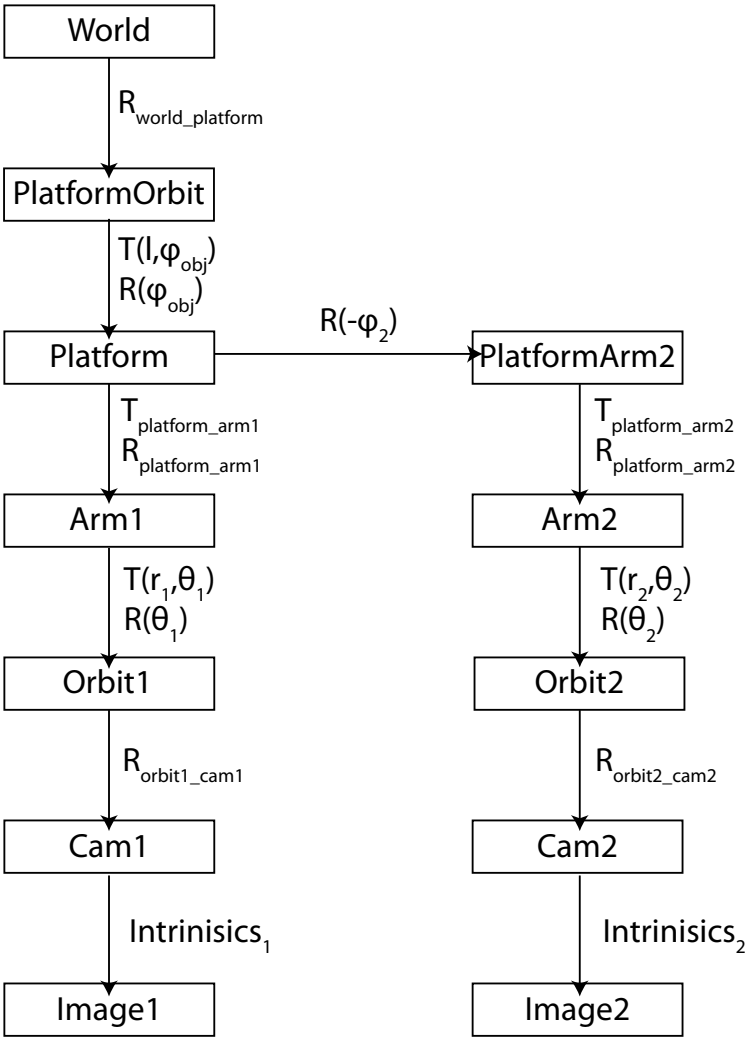


Figure 5.5: Diagram for $\mathbb{T}_1$ and $\mathbb{T}_2$. Each box represents a single coordinate system. The rotations (R) and translations (T) required to move between coordinate systems are listed along the arrows.

the checkerboard.

We start by assuming the checkerboard is placed in the center of the platform, and the x-axis pointed in the direction of arm1, parrallel to its axis of rotation. For our specific checkerboard in this configuration, the platform orbit radius (distance from the checkerboard origin to the platform's axis of rotation) is approximately 65mm. The arm1 and arm2 radii are both approximately 840mm.

The rotation from the world to the platform orbit is approximately a $\frac{3}{4}\pi$ rotation around the z-axis, since the x-axis should point away from the axis of rotation.

When we move from the platform coordinate system to the arm1 coordinate system, we know the z-axis of the arm1 coordinate system should point away from arm1, which is roughly in the same direction as the $-x$ axis of the platform coordinate system. Also, the x axis of the arm1 coordinate system should point straight up, which is the platform coordinate system's z axis. Using these facts we can compute the desired rotation (using `homography_axis(x,y,z)`). We assume there is no translation so the translation is initialized to $(0, 0, 0)$.

Finally, after converting to the arm1 orbit coordinate system, we need to rotate to the final camera position. Again, in the arm1 orbit system the $-x$ axis points in toward the object, and the z axis points away from arm1, whereas in the camera system the z axis points in toward the object and the $-y$ axis points away from arm1. An almost identical process gives us good starting positions for the arm2 side of the homography tree.

We used Zhang's chart-based algorithm [2000] to estimate the intrinsic parameters of each camera independently as a first step. The parameters of each homography are then solved for using a hierarchical simplex optimization. The final calibration is able to predict the position of each coaxial device to within about 5 pixel reprojection error over a 12 inch diameter working volume.

5.2.2 Radiometric calibration

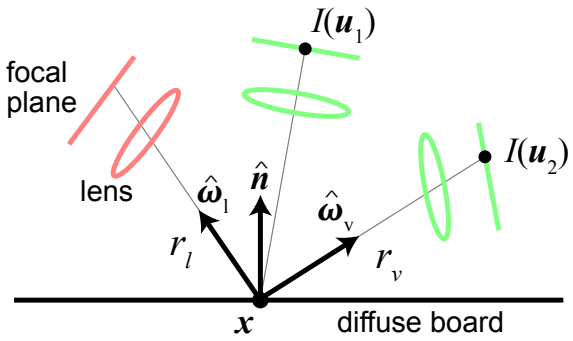


Figure 5.6: Illustration of our procedure for estimating the camera sensitivity and light emission functions.

In order to convert the response at each camera pixel into a measurement of the BRDF, we must know the camera sensitivity and light output, both of which depend on wavelength and position on the respective focal plane. We represent camera sensitivity using three scalar functions, one for each color band, defined over the image plane: $\sigma(\mathbf{u}),\ \mathbf{u} \in \mathbb{R}^2$. We represent light emission using a scalar function that is attached to a coordinate system associated with the coaxial device. This function is defined over a 3D volume that tightly encloses the working volume, and we denote it by $l(\mathbf{x}),\ \mathbf{x} \in \mathbb{R}^3$. We opt to represent this as a 3D function instead of a 2D function on the source focal plane because the beamsplitter produces spectral variations in the emitted light field that are difficult to represent with a simple point-source model. One consequence of this choice is that the inverse-square effect of the light source is absorbed into $l(\mathbf{x})$.

We first estimate the camera sensitivity functions for each device following the procedure illustrated in Figure 5.6. A planar Spectralon target is held fixed relative to one of the coaxial devices (red) which provides the only source of illumination in the scene — the glass slide is removed during this procedure. The other device (green) records several images at different viewpoints. Under the assumption that the board is a perfectly

diffuse reflector, the response at a single pixel $I(\mathbf{u})$ can be written as

$$I(\mathbf{u}) = \sigma(\mathbf{u}) l(\mathbf{x})(\mathbf{n} \cdot \omega_i), \tag{5.1}$$

where the $1/r_l^2$ term is absorbed into the 3D function $l(\mathbf{x})$ as mentioned above. Each point in a discrete set of points on the board induces a linear constraint on the sensitivity function: $\sigma(\mathbf{u}_1)/I(\mathbf{u}_1) - \sigma(\mathbf{u}_2)/I(\mathbf{u}_2) = 0$. Because we expect $\sigma(\mathbf{u})$ to be smooth, we model it as a 10^{th}-order bivariate polynomial whose coefficients are estimated by solving a constrained linear system of the form $Ax = 0,\ \|x\| = 1$.

Once we have estimated the camera sensitivity functions we hold them fixed and estimate the light emission functions $l(\mathbf{x})$. Note that dividing $I(\mathbf{u}_1)$ by $\sigma(\mathbf{u}_1)$ and $(\mathbf{n} \cdot \omega_i)$ gives a value of $l(\mathbf{x})$ at the corresponding 3D position. We record a dense set of measurements at different locations and orientations throughout the working volume and estimate the value of $l(\mathbf{x})$ at the vertices of a volumetric grid by convolving these scattered measurements with a wide Gaussian kernel. We used a $128 \times 128 \times 128$ grid (which corresponds to 1 mm^3 voxels) and a standard deviation of 20.0mm.

We performed random sampling cross validation to evaluate the accuracy of our calibration. We converted the pixel responses into values of the BRDF (presumed to be perfectly diffuse) at a set of 3D positions uniformly sampled over the working volume and over orientations of up to 80° that were not used for training. For both devices, 95% of the values predicted by our model were within 2% of the correct values and 99% were within 5%. Note that our calibration only allows measuring the BRDF up to some unknown global scale. To compute absolute values one would need to measure the power of the light source which we did not do.

5.3 Measurement and Processing Pipeline

We acquire multiple 2.5D *scans* of a target object which are subsequently aligned and merged into a single 3D *model*. The final model is represented as a watertight triangle mesh with a variable-length list of BRDF measurements (local light direction, local view direction, and RGB vector) at each vertex. This section describes the main steps in our pipeline in the order they are performed.

1. We acquire a set of images comprising one scan (Section 5.3.1).

2. From these images, we perform descattering and compute the apparent sinusoid phase and amplitude at each pixel (Section 5.3.2).

3. We apply a feature-based pose refinement algorithm that recovers the positions of the cameras and light sources to within sub-pixel accuracy (Section 5.3.3).

4. We obtain a dense depth map for each scan using a new multiview phase mapping algorithm (Section 5.3.4).

5. We align and merge multiple depth maps into a single watertight model (Section 5.3.5).

6. Finally, we extract BRDF samples for each vertex in the watertight mesh by applying appropriate radiometric correction factors (Section 5.3.6).

5.3.1 Raw Images in One Scan

Here we describe the images that are acquired in a single 2.5D scan. The target object is positioned on a small platform in the center of the working volume, and one of the gantry arms is selected as the *reference frame* and held fixed. The free arm is moved to

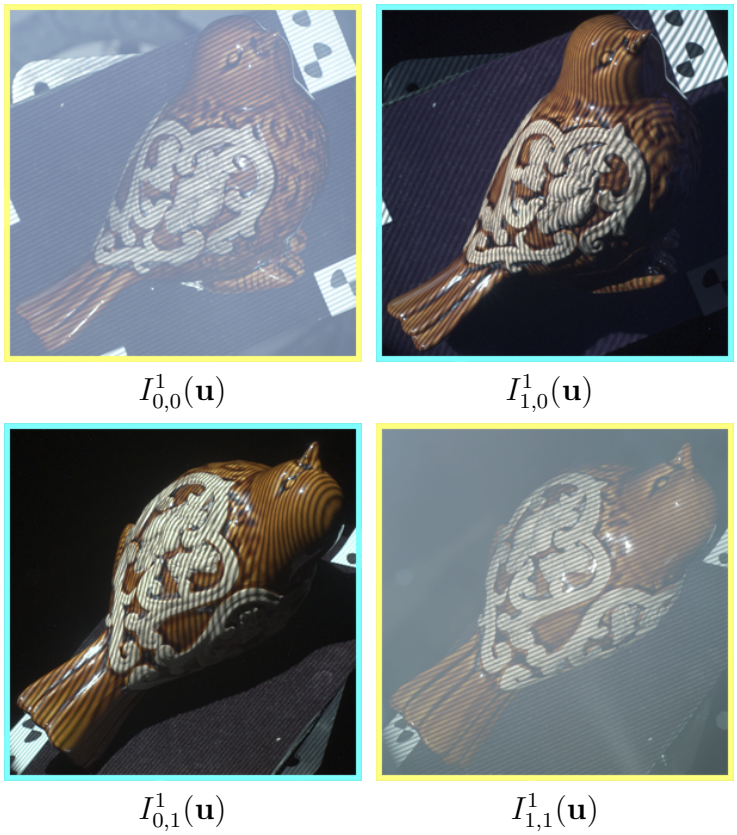


Figure 5.7: Some of the raw images acquired in one scan. The yellow bordered images are captured under co-axial lighting and the blue bordered images correspond to a reciprocal image pair. Note the high-frequency sinusoidal illumination.

N distinct *auxiliary frames*, and for each of these auxiliary frames we collect four *image stacks* corresponding to the four camera/source combinations (two coaxial images and two reciprocal images). Each stack consists of M images recorded at different positions of the active source's translation stage (i.e., different phase shifts of the projected sine wave).

Some of the images from one scan of the `bird` model are shown in Figure 5.7. We denote images by $I^k_{i,j}(\mathbf{u})$, where $i \in [0 \ldots N]$ indexes the light source position, $j \in [0 \ldots N]$ indexes the camera position, and $k \in [1 \ldots M]$ indexes the position within a stack. For index i (or j), the value 0 is used to represent cases where the source (or camera) are at the reference frame.

In images where $i = j$ the object is captured under co-axial illumination, and these images contain measurements of the surface BRDFs at backscattering configurations. As shown in Figure 5.7, the sinusoidal pattern is not distorted by the scene geometry in these cases, and the images are free of cast shadows. These images exhibit a strong haze caused by reflections from dust on the beamsplitter, but as described below and shown in Figure 5.8, this haze is conveniently removed by optical descattering.

$\alpha_{0,0}(\mathbf{u})$ $\alpha_{1,0}(\mathbf{u})$

Figure 5.8: Amplitude maps computed from the stacks along the top row in Figure 5.7. Note that the strong haze in the coaxial image has been removed. Also, the slight color differences are due to the varying radiometric properties of the light sources and cameras.

Images where $i \neq j$ comprise reciprocal pairs in which i and j are swapped: $(I_{i,j}, I_{j,i})$. These play an important role in our pose refinement (Section 5.3.3) and multiview stereo algorithm (Section 5.3.4).

5.3.2 Amplitude and Phase Maps

Translations of the glass slide induce time-varying sinusoidal irradiance at the object's surface. As explained in chapter 4, the response measured at one camera pixel in this scenario will also be a sinusoid, but with modified amplitude, DC offset, and phase. The amplitude of the sinusoid is proportional to the amount of light reflected directly from the surface ("local reflections") and the DC offset includes the effects of sub-

surface scattering and diffuse interreflections ("non-local reflections") [Nayar et al., 2006; Chen et al., 2007; Gupta et al., 2009]. The underlying assumption is that the non-local contributions to the image are constant over the set of phase shifts, and we discuss the ramifications of this important assumption in Section 4.7.

We model the response at one pixel as: $I^k(\mathbf{u}) = \alpha(\mathbf{u})\cos(\gamma\, t_k + \phi(\mathbf{u})) + \beta(\mathbf{u})$, where α is the amplitude of the cosine, t_k is the displacement of the translation stage in the k^{th} image, γ is the observed frequency with respect to t_k, ϕ is the phase offset, and β is the DC offset. The frequency γ is equal to the product of the spatial frequency of the sine wave printed on the glass slide f and $(t_{k+1} - t_k)$, the distance the slide is translated between images.

The values of interest are the amplitude α, which is connected to the surface BRDF, and the phase offset ϕ, which gives information about the scene depth along each ray leaving the camera. Following Chen et al. [2007], we compute these as the least squares solution of the following system:

$$\begin{bmatrix} \cos(\gamma\, t_1) & -\sin(\gamma\, t_1) & 1 \\ & \vdots & \\ \cos(\gamma\, t_M) & -\sin(\gamma\, t_M) & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} I^1(\mathbf{u}) \\ \vdots \\ I^M(\mathbf{u}) \end{bmatrix} \tag{5.2}$$

$$\alpha(\mathbf{u}) = \sqrt{c_1^2 + c_2^2} \qquad \phi(\mathbf{u}) = \tan^{-1}(c_2/c_1) \qquad \beta(\mathbf{u}) = c_3.$$

Note that the pseudo-inverse of the coefficient matrix can be precomputed as it depends only on γ and t_k which are both known and held fixed. Therefore, computing these maps is very fast and is done for each color channel independently. We retain a single phase offset at each pixel by averaging these three values.

Figure 5.8 shows the $\alpha(\mathbf{u})$ computed from the image stacks along the top row in Figure 5.7. Note that the sinusoidal pattern has been removed along with diffuse inter-

reflections (most strikingly in the image captured under co-axial lighting). However, a significant correction factor must be applied before these images can be interpreted as BRDF measurements (Section 5.3.6). Figure 5.9 shows the phase $\phi(\mathbf{u})$ computed from the image stack in the top right corner of Figure 5.7. We also compute a binary visibility map $V(\mathbf{u})$ by applying a threshold to $\alpha(\mathbf{u})$. Values below 1% of the camera's dynamic range are assumed to be in shadow. By identifying which pixels are in shadow we have also identified which pixels are not visible in the co-located camera [Zickler et al., 2002]. We leverage this information in our multiview stereo algorithm (Section 5.3.4).

$\phi_{1,0}(\mathbf{u})$ $V_{1,0}(\mathbf{u})$

Figure 5.9: Phase map and visibility map computed from the image stack in the top right corner in Figure 5.7. The cosine of the phase at each pixel is shown as a grayscale image.

The frequency of the sine wave printed on the glass slide, f, is an important parameter. We chose a value of $f = 10$ cycles/mm because experimentally, we found it to yield reflectance measurements at a scale suitable for a variety of common objects — we discuss this parameter further in Section 5.3.6 when we describe our procedure for recovering measurements of the BRDF from $\alpha(\mathbf{u})$. The displacements t_k should be chosen so that the system in Equation 5.2 is well-posed. Although this requires only three measurements in theory, we found that translating the glass slide by $0.3/f$ (0.3 times the period) and capturing a total of $M = 10$ images gives stable and accurate estimates.

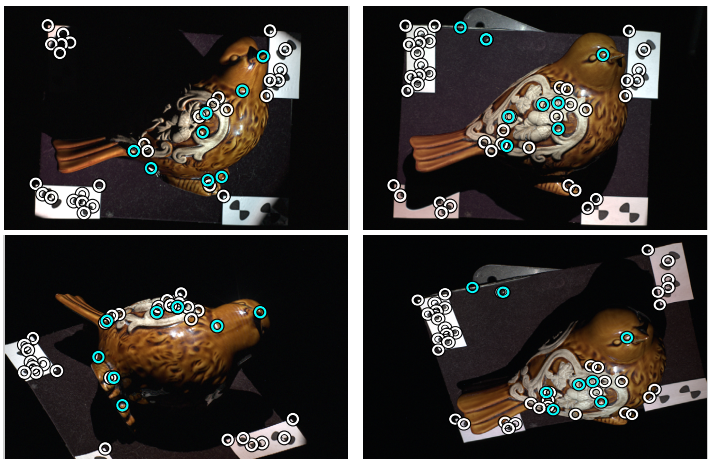


Figure 5.10: Feature-based pose refinement. Each column shows a reciprocal image pair, with reference view in the top row. By detecting intrinsic texture features (white) and reciprocal specularity features (cyan), the geometric calibration is automatically refined from an average reprojection error of 5 pixels to 0.3 pixels.

5.3.3 Pose Refinement

In order to realize the potential resolution of our scanner, the reference and auxiliary locations in each scan must be known to within sub-pixel accuracy. Although standard calibration algorithms provide this level of accuracy, they require placing precise fiducials in the scene such as a planar checkerboard with known dimensions [Zhang, 2000]. We achieve a more automated solution by using a feature-based pose refinement algorithm to improve the rough calibration derived from the gantry configuration. The key challenge is identifying corresponding scene points in images captured at different viewpoints. Fortunately, the fact that we acquire reciprocal image pairs allows reliably locating stable image features even for objects with strong view-dependent appearance that lack "intrinsic" features [Zickler, 2006] — note that specular highlights remain fixed to the object surface in a reciprocal pair. We locate corners in $\alpha(\mathbf{u})$ using a Harris detector [Harris and Stephens, 1988] along with specular highlights by applying a simple intensity threshold. We then match corner features between all pairs of images and match specular features between reciprocal images by simply eliminating false matches

using the epipolar constraints provided by the rough calibration. In some cases, we added additional black and white corners to the scene to increase the number of available features. However, unlike traditional calibration targets, the 3D locations of these corners is not assumed to be known. Figure 5.10 shows a typical set of features extracted from images of the `bird` model. We found this simple procedure to be very reliable and we did not need to use more extensive feature vector matching [Brown and Rusinkiewicz, 2005] or employ a statistical outlier rejection method. Finally, we use the Sparse Bundle Adjustment (SBA) package developed by Lourakis et al. [2004] to jointly optimize the 3D locations of the features and camera positions along with their focal lengths and principal points. We observed final reprojection accuracy of approximately 0.3 pixels for all of the models we scanned.

5.3.4 Multiview Phase Mapping

The next step in our pipeline is to compute a dense depth map in the reference frame of each scan using the phase maps and visibility maps described in Section 5.3.2. Traditional phase-based profilometry methods use the phase shift observed at each pixel in conjunction with the pose of the light source to recover the depth up to a 2π ambiguity [Srinivasan et al., 1985]. This ambiguity is resolved through a process called phase unwrapping, either by varying the frequency of the modulated light or with some heuristic such as favoring locally smooth surfaces [Strand and Taxt, 1999].

We take a different approach that does not rely on knowing the pose of the light source, and we resolve the 2π ambiguity by considering information across multiple viewpoints: we search along the ray through each pixel in the reference camera to locate a 3D position that gives consistent phase information over the set of views. This avoids the need to vary the frequency of the modulated light [Chen et al., 2007] or use more sophisticated lighting patterns [Chen et al., 2008], thereby simplifying the overall design.

Also, our multiview matching process benefits from the fact that visibility information is available before anything is known about the surface by virtue of having captured reciprocal image pairs [Zickler et al., 2002]. This eliminates the need for reasoning about scene visibility during matching, which is a significant source of complexity in standard multiview stereo [Seitz et al., 2006].

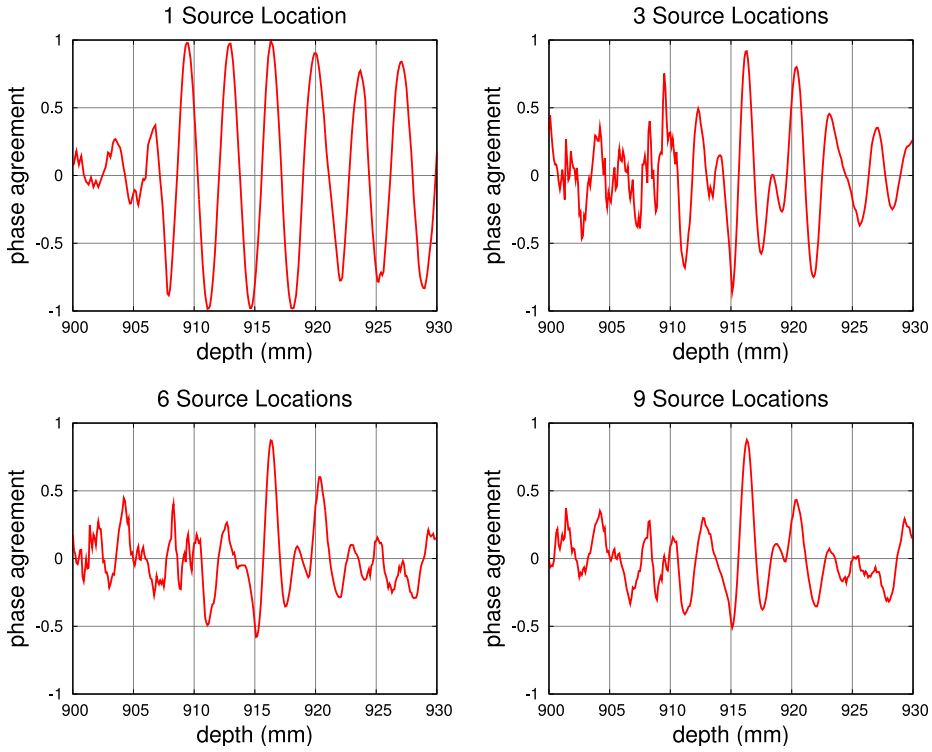


Figure 5.11: The search for depth correspondences becomes more stable with more auxiliary frames. We recommend $N \geq 7$ in practice.

Our approach is related to space-time stereo [Zhang et al., 2003; Davis et al., 2005] in that we deviate from conventional phase-based profilometry by not assuming knowledge of the light source pose. Like space-time stereo, we instead treat phase information as "texture" projected into the scene to help establish stereo correspondence. However, unlike conventional space-time stereo, we use sinusoidal illumination, and this allows reliably measuring the geometry of objects in the presence of scene interreflections and subsurface scattering [Nayar et al., 2006; Chen et al., 2007].

Let $(\mathbf{u}_i) = \Pi_i(\mathbf{x})$ denote the projection of a 3D point $\mathbf{x}$ into the image plane of the i^{th} camera. At each source location i with an unoccluded view of $\mathbf{x}$, two conditions must be satisfied for $\mathbf{x}$ to lie on the object surface:

$$\phi_{0,0}(\mathbf{u}_0) = \phi_{0,i}(\mathbf{u}_i) \qquad \phi_{i,0}(\mathbf{u}_0) = \phi_{i,i}(\mathbf{u}_i).$$

This corresponds to observing the same phase offset in the temporal response measured at different views when the object is illuminated by the same light source. We treat phase offsets as unit-length vectors in the plane and measure the agreement between two phases as the dot product of their respective vectors. This allows writing an objective function over points $\mathbf{x}$ along each ray

$$\epsilon(\mathbf{x}) = \frac{\sum_{i=1}^{N} p_i V_{i,0}(\mathbf{u}_0)}{\sum_{i=1}^{N} V_{i,0}(\mathbf{u}_0)}, \tag{5.3}$$

where

$$p_i = \frac{(\phi_{0,0}(\mathbf{u}_0) \cdot \phi_{0,i}(\mathbf{u}_i)) + (\phi_{i,0}(\mathbf{u}_0) \cdot \phi_{i,i}(\mathbf{u}_i))}{2}.$$

The expression $(\phi_a(\mathbf{x}) \cdot \phi_b(\mathbf{x}))$ denotes the inner product of the two vectors associated with these phase values. Note that $\epsilon(\mathbf{x})$ is bounded between -1 and 1 as long as the denominator is greater than 0. As expected, increasing the number of auxiliary frames N causes the function $\epsilon(\mathbf{x})$ to become more peaked around the correct depth as false matches are eliminated. This is illustrated in the graphs in Figure 5.11 which plot $\epsilon(\mathbf{x})$ as a function of the depth at one of the pixels in the reference view in Figure 5.7 over a 3cm interval that straddles the surface. We compute the position $\mathbf{x}$ that maximizes $\epsilon(\mathbf{x})$ in two steps. First, we compute values of $\epsilon(\mathbf{x})$ at 0.1mm increments across the working volume and locate the sample with the largest value. Second, we fit a quadratic function to the 3 samples of $\epsilon(\mathbf{x})$ surrounding this maximum and output the depth that

maximizes this quadratic. We have found that $N \geq 7$ is sufficient to recover accurate geometry (Section 3.3).

We record depth values only at pixels that are visible in at least 3 views and for which $\epsilon(\mathbf{x}) > 0.5$. In order to reject outliers, we embed the depth map in a graph whereby pixels are represented as nodes and edges connect neighboring pixels whose depth values are within 1.5mm (0.5% of the working volume). We then retain the largest connected component. This removes unwanted elements of the background that are often captured in the phase maps (Figure 5.9).

5.3.5 Alignment and Merging

We use well established methods for aligning multiple depth maps to one another and merging them into a single watertight model. Specifically, we use the Iterative Closest Points (ICP) algorithm [Zhang, 1994] followed by the Poisson surface reconstruction algorithm of Kazhdan et al. [2006]. At this point we have the original depth maps and their corresponding reference and auxiliary images registered to the merged model. The final step is extracting BRDF measurements.

5.3.6 Recovering BRDF Measurements

A significant advantage of using sinusoidal illumination is that the amplitude maps $\alpha_{i,j}(\mathbf{u})$ carry information about the local reflectance at the measurement scale of the camera and projected sine wave, even when certain non-local effects, such as diffuse interreflections, are present in the scene [Nayar et al., 2006].

Amplitude loss. In chapter 4 we derived a closed-form expression for the amplitude of the time-varying sinusoid measured at a single camera pixel imaging an opaque surface under projected sinusoidal illumination (equation 4.8).

This result confirms that the values in $\alpha(\mathbf{u})$ are proportional to the product of the surface irradiance under point lighting $l(\mathbf{x})(\mathbf{n} \cdot \omega_i)$ (recall that $1/r_l^2$ is captured by $l(\mathbf{x})$), the amplitude of the projected sinusoid A, the BRDF $\rho(\omega_i, \omega_o)$, and the camera sensitivity function $\sigma(\mathbf{u})$. However, it also predicts a less obvious effect we call "amplitude loss" (Equation 4.8) whereby the measured response is inversely proportional to the product of the pixel width w and the frequency f of the source radiance patterns in addition to the relative distances between $\mathbf{x}$ and the camera and source. In words, if either f or w increase (holding everything else fixed) the measured amplitude will decrease at a rate predicted by the product of the sinc functions in Equation 4.8 and eventually reach zero — this corresponds to the point at which the sine pattern is no longer visible in the image. Similarly, as the camera approaches a grazing view of the surface with the source held fixed overhead or the distance r_v increases then the measured amplitude will similarly decrease. The graphs in Figure 5.12 confirm this effect and validate our analytic model.

Based on this analysis, we harvest BRDF samples as follows. Each vertex in the merged model is projected into the amplitude maps $\alpha_{i,j}(\mathbf{u})$ in which it is visible. No additional registration between images and geometry is required because the depth maps and amplitude maps are both computed in the same raster grid, but we do avoid samples near depth discontinuities (any pixel within 3 pixels of a change in depth $\geq$10 mm), since slight misalignments between the individual scans and the merged model may exist. To convert these samples to BRDF values, we divide them by $\sigma(\mathbf{u})$, $l(\mathbf{x})$, $(\mathbf{n} \cdot \omega_i)$, and the product of the sinc functions in Equation 4.8. This produces BRDF measurements up to a single global scale factor that we do not attempt to estimate.

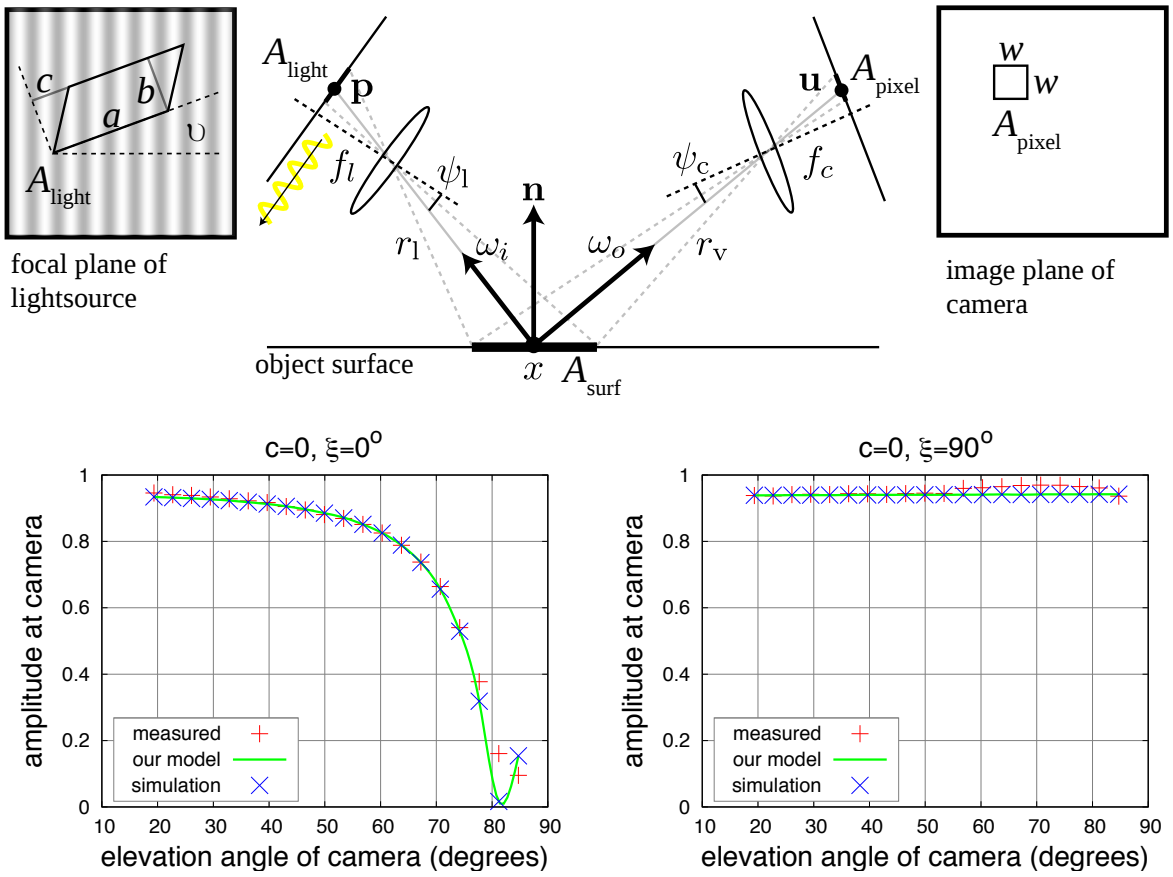


Figure 5.12: **Top:** Geometric setup and notation used in our analysis of "amplitude loss". **Bottom:** Measurements of the amplitude measured at one pixel as a camera moves towards the horizon with a fixed overhead light in a direction perpendicular and parallel to the projected sine wave, respectively. The measurements correspond to images of a Spectralon board corrected to account for deviations from a perfectly Lambertian reflector.

5.4 Results

In this section we analyze the accuracy of models acquired with our scanner for several objects that exhibit a range of properties.

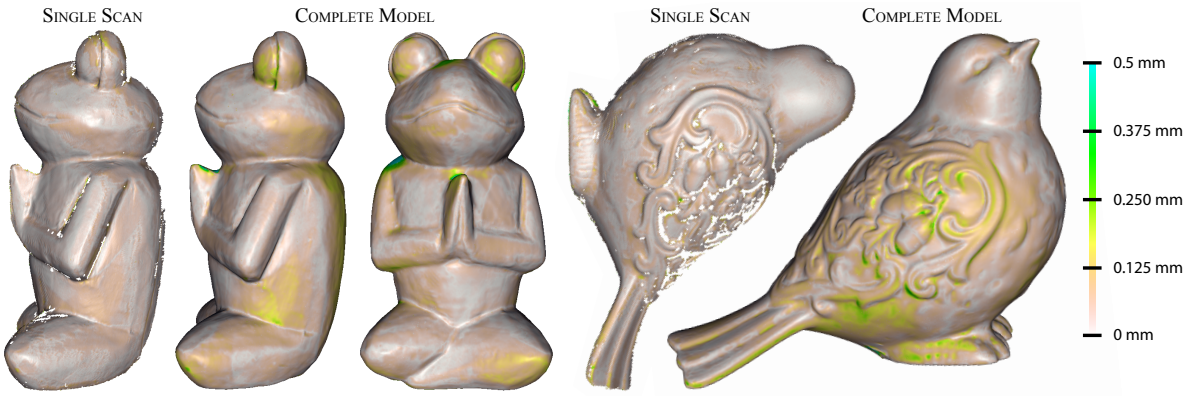


Figure 5.13: Error analysis of geometry acquired with our scanner for the `frog` and `bird` models. Comparisons are made to a reference surface obtained by coating these objects in a diffuse powder and scanning them with a high-end laser rangefinder.

5.4.1 Geometry

We measured the geometric accuracy of our scanner by comparing it to a high-end laser rangefinder operated by XYZRGB, Inc.[2] This laser scanner produces scans at 300 micron resolution with an accuracy of 20 microns. We scanned the first two objects shown in Figure 5.16 after first coating them in a removable diffuse powder. The `bird` is ceramic and exhibits spatially-varying reflectance including both very shiny glazed regions and diffuse unglazed regions. The `frog` is made of painted wood and also exhibits spatial variation in both the material albedo and in the shape and strength of the specular reflection.

Figure 5.13 shows individual scans and the final models acquired with our scanner. We used ICP to align these models to the laser-scanned reference and report the distance between each vertex and the nearest point on the reference surface. The median error for the `frog` model is $\sim$40 microns and the maximum error in any single scan is 0.25mm. The median error for the `bird` model is $\sim$50 microns and the maximum in any scan is 0.25mm. Figure 5.13 also demonstrates the typical amount of coverage in one scan. Our

[2] `http://www.xyzrgb.com`

data agrees very well with the reference, although there are areas that contain slight ringing artifacts (e.g., near the neck of the bird and on the arm of the frog) due to deviations between our measurements and the model expressed in Equation 5.2 as well as reprojection errors that remain after pose refinement. However, note that these errors never exceed more than 0.05mm and are typically on the order of 0.02mm. The right shoulder of the frog and the top of its hands were not visible in any of the scans and were filled in during the merging process so they deviate more substantially from the reference.

5.4.2 Reflectance

We measured the BRDFs of two types of spray paint after applying them to a planar board, removing the glass sinusoidal slides from our co-axial assemblies, and using the gantry as a traditional reflectometer. We compared this data to measurements obtained using our scanner for the same planar sample and a curved sample. Figure 5.14 shows comparisons along with the curved samples for Valspar #66304 Aubergine Silk Interior Enamel ("red paint)" and Valspar #66307 Lime Pearl Interior Enamel ("pearl paint"). The two lobes correspond to an elevation angle in the light source of 20 degrees and 40 degrees, respectively, and densely sampled view directions within the plane formed by the source and surface normal. For both the planar and curved samples, we observed close agreement with the reference data (within 5.5%) over this range of light and view directions. We also measured the degree of reciprocity in our data using the set of reciprocal images we acquire in each scan. We observed a median error of 13% for the red paint and 15% for the pearl paint — within individual scans these errors were roughly 10%. Although these errors are competitive with prior image-based reflectometers [Marschner, 1998; Dana et al., 1999], they were not low enough to enable refining the surface normals as is done in Helmholtz stereopsis [Zickler et al., 2002]. We discuss

this point further in Section 5.5.

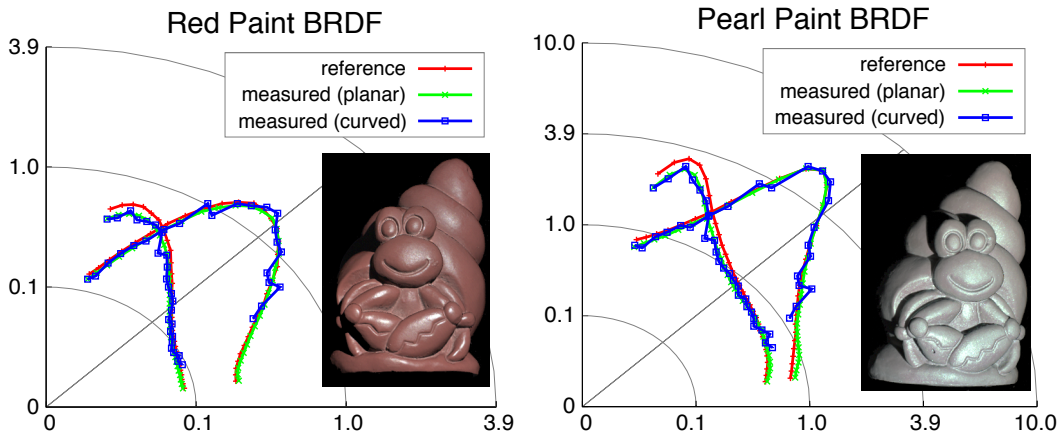


Figure 5.14: Polar plots of reflectance data acquired with our scanner for two types of spray paint compared to reference measurements. The curved samples are shown in the insets. The two lobes in each graph show in-plane measurements of changing view elevation angle, for a fixed light elevation angle equal to 20 degrees and 40 degrees, respectively. Intensity is plotted relative to a standard diffuse target.

5.4.3 Geometry and Reflectance

We also analyzed the accuracy of our scanner by fitting an analytic spatially-varying BRDF to the models shown in Figure 5.16. This allows rendering images at view and light positions that were not initially sampled. The `cat` object is made of wood and has both unfinished and polished areas as well as painted regions. For each object, we acquired 6 scans with 7 auxiliary frames in each scan chosen uniformly within a cone of 60 degrees around the reference camera. This yields an average of 10 BRDF measurements per vertex. With this number of measurements, important features of the reflectance such as specular highlights will only be observed at a relatively small number of vertices, especially for very shiny objects. Therefore, independently fitting a BRDF model to the data at each vertex would give poor results. Building on prior work, we instead model the spatially-varying reflectance as a low-dimensional subspace

spanned by a small set of "basis" BRDFs [Lensch et al., 2003b; Goldman et al., 2005; Lawrence et al., 2006]. This allows sharing reflectance data between different vertex locations (with potentially different orientations) that cover the same material.

We first use the k-means algorithm to separate the surface into clusters based on the diffuse albedo at each vertex (Figure 5.15 shows the clusters for the `cat` model). We fit the parameters of a Cook-Torrance BRDF [1981] to the reflectance data within each of these clusters (tens of thousands of scattered BRDF measurements). Finally, at each vertex we estimate a diffuse color and set of convex linear blending weights (partition of unity) over the specular terms in these k Cook-Torrance BRDFs that give the best agreement with the measurements. Figure 5.16 compares renderings to reference images at light and camera positions that are approximately 20 degrees away from the closest measured ones. The reference image is the amplitude map for one auxiliary frame that was held out during training. We observed very close agreement for all of these objects over the entire range of view and light positions. The supplemental video includes animations of these models rotating under point lighting. Note that achieving accurate smooth motion of specular highlights over the object surface (e.g., in the `bird` model) would require an infeasible amount of data with image-based methods that rely on approximate geometry. With those methods, these features would be improperly interpolated wherever the proxy geometry deviated from the true object surface resulting in "ghosting" artifacts. The fact that we acquire very accurate geometry and reflectance allows producing convincing smooth interpolations of these high-frequency features from only 36 light positions.

5.4.4 Capture and Processing Times

Table 5.1 lists the number of individual scans we acquired for each object along with capture and processing times. It requires roughly 7 minutes to collect the set of HDR

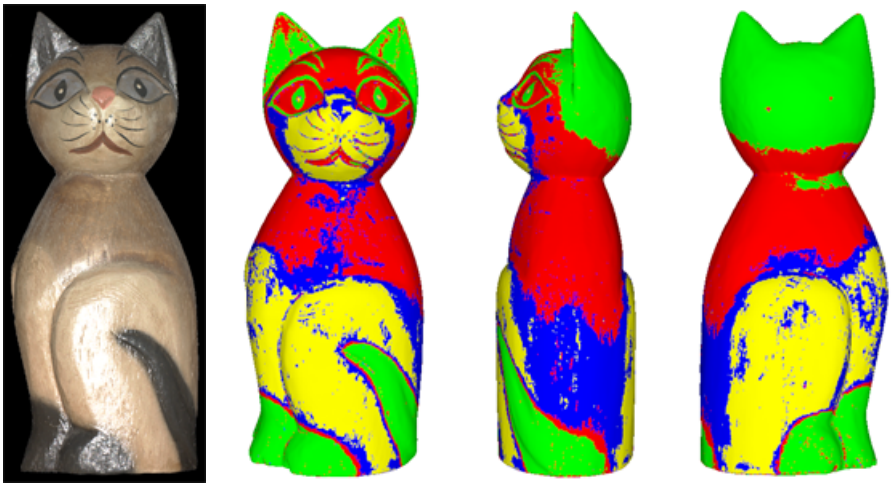


Figure 5.15: Material clusters computed for the `cat` model. A photograph is shown on the left next to false-color images showing the coverage of the four clusters over the object.

images at each auxiliary frame (Section 5.3.1), 30% of which is spent operating the translation stage and mechanical shutters and positioning the gantry. The remaining time could be significantly reduced by using more powerful light sources. Computing the amplitude and phase maps along with performing pose refinement for one scan with 7 auxiliary frames takes approximately 5 minutes in addition to $10 - 30$ minutes to estimate depth maps. Alignment and merging multiple scans and reconstructing BRDF samples requires an additional 50 minutes on average.

5.5 Discussion and Future Work

We have demonstrated an acquisition setup and processing pipeline for obtaining accurate high-resolution measurements of the 3D shape and reflectance of opaque objects. This was enabled by an optical design centered around a co-located camera and high frequency spatially-modulated light source, as well as a new active multiview stereo algorithm and a theoretical analysis of light descattering with sinusoidal illumination. We presented results that show geometry captured with our system is accurate to within

Figure 5.16: Rendered images computed using models captured by our scanner compared to reference images. The chosen light and view positions are approximately 20 degrees away from the closest measurement location. Note that we do not attempt to render shadows.

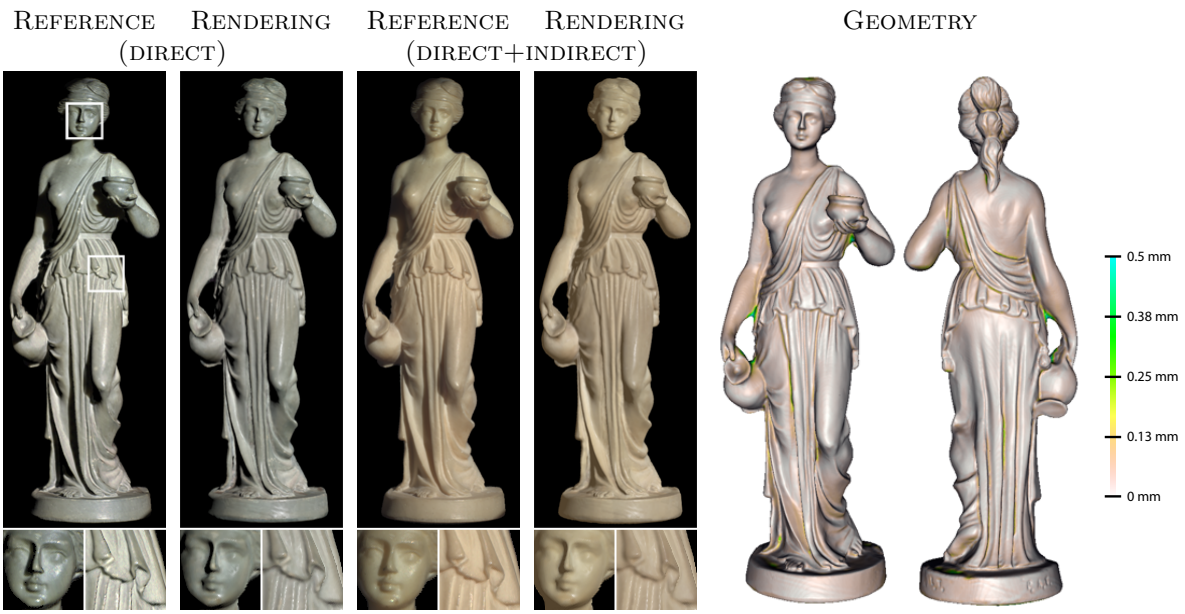


Figure 5.17: Preliminary result of using our scanner to measure the appearance of a translucent object. We do not attempt to render shadows.

50 microns of the true surface on average and agrees with reference reflectance data to within 5.5%.

Our proposed system represents one point in the larger design space of scanners that might use coaxial devices, and we anticipate future systems with goals and requirements that are different from those considered here such as scanning translucent objects. Additionally, we believe this work could lead to new tools for appearance scanning "in the field". Co-axial camera and light source assemblies reduce the calibration burden significantly, because the position and orientation of the light source does not need to be continually re-calibrated. While we rely on a rough chart-based calibration of our gantry, it is likely that this step could also be replaced by automatic feature-based techniques [Brown and Rusinkiewicz, 2005]. Thus, future work might allow scanning outside a lab setting with rough placement of multiple co-axial assemblies.

In particular, we believe that similar systems can be used for scanning translucent objects, and Figure 5.17 shows a preliminary step in this direction. In this example,

Model	Scans	Capture Time	Processing Time	Total
Bird	7	5.8 h	3.7 h	7.0 h
Frog	6	5.0 h	3.6 h	6.2 h
Cat	6	5.0 h	3.2 h	6.2 h

Table 5.1: Statistics for the results reported in the paper. A portion of the capture and processing are done in parallel and so the total is less than the sum of these parts.

the geometry and local reflectance were measured and modeled using the same steps described in Sections 5.3 and 5.4.3, respectively. In addition, we augmented this model with a single ambient term computed by averaging the response in shadowed regions in the "indirect" images — the $\beta(\mathbf{u})$ maps computed in Equation 5.2 — along with an additional diffuse component at each vertex. For evaluation, we compare the recovered geometry to a laser scan, and we qualitatively compare a rendering to a "direct+indirect" reference formed by taking the sum of $\alpha(\mathbf{u})$ and $\beta(\mathbf{u})$ and interpreting their combination as the reflection at the object surface and internal scattering below it.

In this example, the geometry is very accurate as seen in a comparison to a laser scan, but while the local reflectance generally agrees with the reference, there are larger differences than were observed for the opaque objects in Figure 5.16. We attribute this to the fact that the sinusoidal illumination pattern is "stretched" by surface foreshortening so that the effective frequency on an object's surface changes with the local lighting direction. This means that different fractions of the scattering are being separated into the "local" and "non-local" components, and the spatial scale at which the BRDF is being measured is not constant. We did not see evidence of this effect in the earlier results because it is negligible for surfaces that are effectively opaque at our chosen measurement scale. We expect that reflectometry for translucent objects will benefit from a careful analysis of light descattering for materials with sub-surface scattering along the lines of our analysis for opaque surfaces.

Another aspect of these illumination systems that deserves more attention is defocus

in the lightsource. Although we were able to set the aperture on the light path to avoid artifacts due to a limited depth of field, a proper analysis of defocus would allow using larger apertures and decrease acquisition time. Some work was recently published on this topic [Gupta et al., 2009].

We inherit limitations of the light descattering method we build upon and point out that only diffuse interreflections are completely removed by our system. Other interreflections could corrupt the data, although we have found that they are often detectable either as a high residual error in the solution to Equation 5.2 or as outliers during the search for correspondences in our multiview stereo algorithm.

Another weakness in our final design is the complexity of the radiometric properties of the lightsource. Although the errors we report are comparable to prior image-based reflectometry systems, the lack of reciprocity in our measurements prevented us from refining the surface normals using a method like Helmholtz stereopsis [Zickler et al., 2002]. Future research is warranted into alternative beamsplitters that minimize wavelength interference effects in order to incorporate photometric constraints that would help further improve the quality of the resulting models.

Chapter 6

Conclusion

Measuring the shape and appearance of objects is a challenging task, but has a wide range of applications including cultural heritage preservation, architecture, law enforcement, industry, and entertainment. This thesis has demonstrated how acquisition can be simplified and improved by the **synchronous measurement of shape and appearance**, in contrast to prior work that measured each independently.

In chapter 3, we demonstrated the first photometric approach for measuring the shape and appearance of objects with anisotropic materials. For such materials, both the surface normal as well as tangent direction must be recovered, unlike in the case of isotropic materials where the tangent direction can be ignored. We observed that the microfacet distribution of most analytic models as well as measured data exhibit planar symmetries across the plane formed by the normal and tangent vectors. We then designed an optimization strategy to locate these vectors using a dense set of images taken under point light sources, resulting in a normal and tangent vector recovered at each pixel along with a dense 2D slice of the BRDF.

Chapter 4 introduced our analysis of sinusoidal illumination, which shows how to simultaneously perform optical descattering, recover 3D geometry, and acquire mea-

surements of the surface BRDF. In addition, we showed that acquiring the geometry of translucent objects using sinusoidal illumination is possible for a wide range of materials and sinusoid frequencies. Both of these theoretical results provide a much needed guide for the design of practical scanning systems, and provide the analysis needed to simultaneously recover shape and appearance.

A full pipeline for acquiring the shape and appearance of opaque objects was presented in chapter 5. We combined sinusoidal illumination with a novel optical design to create *coaxial devices*, which greatly simplify the joint measurement task compared to prior work. Using the key result from our analysis of sinusoidal illumination and a novel multiview stereo algorithm, we demonstrated the capture of complete 360° surface meshes with spatially varying BRDF from physical objects. Our evaluation of the system compares the final geometry with state-of-the-art laser scans; the reflectance with a reference gonioreflectometer; and the final model with reference photographs.

6.1 Quantitative Evaluation

We hope that our rigorous quantitative evaluation of the coaxial scanning system sets a precedent for future work on shape and appearance acquisition. Such systems can be very difficult to implement and often require special equipment, therefore it is impractical for future researchers to reimplement existing systems to make comparisons. As we have shown, current measurement systems are capable of producing images comparable with real photographs; however, we believe that reporting direct comparisons with ground truth geometry and reflectance is important. These comparisons enable future work to compare quantitatively against prior methods, without relying on vague arguments about image quality.

6.2 Future Work

The work presented in this thesis suggests several directions for new research. In particular, the datasets generated by this work have all been made available to the research community, and will hopefully enable new experiments that were not previously possible.

Tangent Fields

As discussed in section 3.4, the time-consuming capture of images under dense point light sources could potentially be alleviated using wider area lightsources. An interesting observation is that symmetric functions, such as the microfacet distribution, remain symmetric after convolution with any symmetric filter. Thus, the approach described in chapter 3 could still be applied to data captured with area lightsources. Using a large area light would significantly speed up acquisition by reducing camera exposure times, and also prefilter the BRDF making reconstruction of the function well-posed with fewer samples even for shiny materials.

Analysis of Sinusoidal Illumination

Our analysis of sinusoidal illumination ignores camera and lightsource depth of field. In practice, it is often impractical to keep an entire scene in-focus, and so a further analysis of this effect is needed if we wish to accurate estimate the BRDF for such scenes. If defocus can be reformulated or approximated as convolutions with the original signal, it may be possible to reach a simple analytic formula even in this more complex setting.

Acquiring the full appearance of translucent objects using sinusoidal illumination remains an open problem. In figure 5.17 we show early results using out system for this purpose, however our approximation of the subsurface scattering is very naive (just a diffuse color at each vertex). It may be possible, for example by capturing data under

different frequency sinusoids, to estimate internal scattering parameters of an object and enable accurate subsurface scattering renderings (figure 2.3 shows an example of such a rendering, but with manually set parameters).

Scanning Systems

Our scanning pipeline indicates new directions of research in view planning. The location of auxiliary frames could be chosen adaptively, guided by an online reconstruction of the shape and reflectance of the target surface. Although prior work has investigated view planning for SVBRDF acquisition [Lensch et al., 2003b], they assume known geometry. How to prioritize the goals of accurate geometry reconstruction and SVBRDF sampling is not obvious.

The fact that our setup produces perfectly registered depth and BRDF maps opens up new possibilities for alignment and merging techniques that consider both of these components as opposed to focusing only on geometry. Some related work has been done on alignment and merging of 3D meshes with RGB texture over their surface [Johnson and Kang, 1999], but measured BRDF data presents additional challenges due to its 4-dimensional nature and sparse sampling. The datasets we have created should enable new research in this area that was previous impractical due to the difficulty in creating a measurement pipeline from scratch.

Finally, the simplified calibration requirements of our coaxial devices suggests new system designs that would make scanning possible outside the laboratory setting. As mobile devices become equipped with increasingly powerful cameras, "in the field" measurements that could be taken by a novice user become increasingly practical.

Bibliography

[Alldrin and Kriegman, 2007] Neil Alldrin and D Kriegman. Toward Reconstructing Surfaces With Arbitrary Isotropic Reflectance: {A} Stratified Photometric Stereo Approach. In *ICCV*, pages 1–8, 2007.

[Alldrin et al., 2008] Neil Alldrin, Todd Zickler, and D Kriegman. Photometric Stereo with Non-Parametric and Spatially-Varying Reflectance. In *Computer Vision and Pattern Recognition*, 2008.

[Ashikhmin et al., 2000] Michael Ashikhmin, Simon Premoze, and Peter Shirley. A Microfacet-based BRDF Generator. In *Computer Graphics (SIGGRAPH 2000 Proceedings)*, pages 65–74. ACM SIGGRAPH, 2000.

[Barsky and Petrou, 2001] S Barsky and M Petrou. Colour photometric stereo: Simultaneous reconstruction of local gradient and colour of rough textured surfaces. In *ICCV*, volume 2, pages 600–605, 2001.

[Bernardini et al., 1999] F. Bernardini, J. Mittleman, H. Rushmeier, C. Silva, and G. Taubin. The ball-pivoting algorithm for surface reconstruction. *Visualization and Computer Graphics, IEEE Transactions on*, 5(4):349–359, 1999.

[Blinn, 1977] J.F. Blinn. Models of light reflection for computer synthesized pictures. *ACM SIGGRAPH Computer Graphics*, 11(2):198, 1977.

[Boehler et al., 2003] Wolfgang Boehler, M Vicent, and Andreas Marbs. Investigating Laser Scanner Accuracy. presented at the XIX Symposium, October 2003.

[Brown and Rusinkiewicz, 2005] Benedict J. Brown and Szymon Rusinkiewicz. Non-rigid global alignment using thin-plate splines. *ACM SIGGRAPH 2005 Sketches on - SIGGRAPH '05*, page 31, 2005.

[Chen et al., 2006] Tongbo Chen, Michael Goesele, and Hans-Peter Seidel. Mesostructure from Specularity. *CVPR*, 2006.

[Chen et al., 2007] Tongbo Chen, Hendrik Lensch, Christian Fuchs, and Hans-Peter Seidel. Polarization and Phase-Shifting for {3D} Scanning of Translucent Objects. In *CVPR*, 2007.

[Chen et al., 2008] Tongbo Chen, Hans-Peter Seidel, and Hendrik Lensch. Modulated phase-shifting for {3D} scanning. In *CVPR*, 2008.

[Chung and Jia, 2008] H Chung and J Jia. Efficient Photometric Stereo on Glossy Surfaces with Wide Specular Lobes. In *CVPR*, 2008.

[Coleman and Jain, 1982] E Coleman and R Jain. Obtaining 3-dimensional shape of textured and specular surfaces using four-source photometry. *Computer Graphics and Image Processing*, 18(4):309–328, April 1982.

[Cook and Torrance, 1981] Robert L. Cook and Kenneth E. Torrance. A reflectance model for computer graphics. *ACM SIGGRAPH Computer Graphics*, 15(3):307–316, August 1981.

[Dana et al., 1999] Kristin Dana, Bram van Ginneken, Shree Nayar, and Jan Koenderink. Reflectance and Texture of Real-World Surfaces. *ACM Transacations on Graphics*, 18(1):1–34, January 1999.

[Dana, 2001] Kristin Dana. {BRDF/BTF} Measurement Device. In *International Conference on Computer Vision*, pages 460–466, 2001.

[Davis et al., 2005] James Davis, Diego Nehab, Ravi Ramamoorthi, and Szymon Rusinkiewicz. Spacetime Stereo: {A} Unifying Framework for Depth from Triangulation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(2):296–302, February 2005.

[Debevec et al., 2000] Paul Debevec, Tim Hawkins, C. Tchou, H.P. Duiker, W. Sarokin, and Mark Sagar. Acquiring the reflectance field of a human face. In *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, pages 145–156. ACM Press/Addison-Wesley Publishing Co., 2000.

[Fisher et al., 2007] Matthew Fisher, P. Schroder, Mathieu Desbrun, and Hugues Hoppe. Design of tangent vector fields. *ACM Transactions on Graphics*, 26(3):56, July 2007.

[Francken et al., 2008] Yannick Francken, Tom Cuypers, Tom Mertens, Jo Gielis, and Philippe Bekaert. High Quality Mesostructure Acquisition Using Specularities. In *CVPR*, 2008.

[Furukawa and Ponce, 2008] Yasutaka Furukawa and Jean Ponce. Accurate Camera Calibration from Multi-View Stereo and Bundle Adjustment. In *IEEE CVPR*, 2008.

[Garg et al., 2006] Gaurav Garg, Eino-Ville Talvala, Marc Levoy, and Hendrik Lensch. Symmetric Photography: Exploiting Data-sparseness in Reflectance Fields. In *Eurographics Symposium on Rendering (EGSR)*, 2006.

[Georghiades, 2003] Athinodoros Georghiades. Incorporating the Torrance and Sparrow model of reflectance in uncalibrated photometric stereo. In *ICCV*, pages 816–823, 2003.

[Ghiglia and Pritt, 1998] Dennis Ghiglia and Mark Pritt. *Two-Dimensional Phase Unwrapping: Theory, Algorithms, and Software*. Wiley-Interscience, 1998.

[Ghosh et al., 2007] Abhijeet Ghosh, Shruthi Achutha, Wolfgang Heidrich, and Matthew O'Toole. {BRDF} Acquisition with Basis Illumination. In *International Conference on Computer Vision*, pages 1–8, 2007.

[Goesele et al., 2004] Michael Goesele, Hendrik Lensch, Jochen Lang, Christian Fuchs, and Hans-Peter Seidel. {DISCO}: Acquisition of Translucent Objects. *ACM Transactions on Graphics*, 23(3):835–844, August 2004.

[Goldman et al., 2005] Dan Goldman, Brian Curless, Aaron Hertzman, and Steven Seitz. Shape and Spatially-Varying {BRDF}s from Photometric Stereo. In *International Conference on Computer Vision*, pages 341–348, 2005.

[Gortler et al., 1996] Steven J. Gortler, Radek Grzeszczuk, Richard Szeliski, and Michael F. Cohen. The lumigraph. *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques - SIGGRAPH '96*, pages 43–54, 1996.

[Guillemaut et al., 2004] J.-Y. Guillemaut, O. Drbohlav, R. Sara, and J. Illingworth. Helmholtz Stereopsis on rough and strongly textured surfaces. In *3D Data Processing, Visualization and Transmission, 2004. 3DPVT 2004. Proceedings. 2nd International Symposium on*, number 2, pages 10–17. IEEE, 2004.

[Gupta et al., 2009] Mohit Gupta, Yuandong Tian, Srinivasa Narasimhan, and Li Zhang. (De)Focusing on Global Light Transport for Active Scene Recovery. *CVPR*, 2009.

[Han and Perlin, 2003] Jefferson Han and Ken Perlin. Measuring Bidirectional Texture Reflectance with a Kaleidoscope. In *Proceedings of ACM SIGGRAPH 2003*, volume 22, pages 741–748, 2003.

[Harris and Stephens, 1988] Chris Harris and Mike Stephens. A combined corner and edge detector, 1988.

[He et al., 1991] Xiao D. He, Kenneth E. Torrance, François X. Sillion, and Donald P. Greenberg. A comprehensive physical model for light reflection. *ACM SIGGRAPH Computer Graphics*, 25(4):175–186, July 1991.

[Heidrich and Seidel, 1999] Wolfgang Heidrich and H.P. Seidel. Realistic, hardware-accelerated shading and lighting. In *Proceedings of the 26th annual conference on Computer graphics and interactive techniques*, pages 171–178, New York, New York, USA, 1999. ACM Press/Addison-Wesley Publishing Co.

[Henyey and Greenstein, 1940] L.G. Henyey and J.L. Greenstein. Diffuse radiation in the galaxy. *Annales d'Astrophysique*, 3:117, 1940.

[Hertzmann and Seitz, 2003] A. Hertzmann and S.M. Seitz. Shape and materials by example: a photometric stereo approach. *Conference on Computer Vision and Pattern Recognition*, 1:533–540, 2003.

[Holroyd and Lawrence, 2011] Michael Holroyd and Jason Lawrence. An Analysis of Using High-Frequency Sinusoidal Illumination to Measure the 3D Shape of Translucent Objects. *IEEE Computer Vision and Pattern Recognitiion*, 2011.

[Holroyd et al., 2008] Michael Holroyd, Jason Lawrence, Greg Humphreys, and Todd Zickler. A photometric approach for estimating normals and tangents. *International Conference on Computer Graphics and Interactive Techniques*, 27(5), 2008.

[Holroyd et al., 2010a] Michael Holroyd, J. Lawrence, and Todd Zickler. A coaxial optical scanner for synchronous acquisition of 3D geometry and surface reflectance. *ACM SIGGRAPH 2010*, pages 1–12, 2010.

[Holroyd et al., 2010b] Michael Holroyd, J. Lawrence, and Todd Zickler. A radiometric analysis of projected sinusoidal illumination for opaque surfaces, 2010.

[Horn, 1975] B Horn. Obtaining Shape from Shading Information. In *Shape from Shading*, pages 123–173, 1975.

[Jakob, 2010] Wenzel Jakob. http://mitsuba-renderer.org, 2010.

[Jensen et al., 2001] HW Jensen, SR Marschner, and Marc Levoy. A practical model for subsurface light transport. *ACM SIGGRAPH*, 2001.

[Johnson and Kang, 1999] Andrew Johnson and Sing Kang. Registration and Integration of Textured {3-D} Data. *Image and Vision Computing*, 17(2):135–147, February 1999.

[Kazhdan et al., 2006] Michael Kazhdan, Matthew Bolitho, and Hugues Hoppe. Poisson Surface Reconstruction. In *Proceedings of Eurographics Symposium on Geometry Processing 2006*, volume 256, pages 61–70. Eurographics Association, 2006.

[Lawrence et al., 2006] Jason Lawrence, Aner Ben-Artzi, Christopher DeCoro, Wojciech Matusik, Hanspeter Pfister, Ravi Ramamoorthi, and Szymon Rusinkiewicz. Inverse shade trees for non-parametric material representation and editing. *ACM Transactions on Graphics*, 25(3):735, July 2006.

[Lensch et al., 2003a] Hendrik Lensch, Jan Kautz, Michael Goesele, Wolfgang Heidrich, and Hans-Peter Seidel. Image-based Reconstruction of Spatial Appearance and Geometric Detail. *ACM Transactions on Graphics*, 22(2):234–257, April 2003.

[Lensch et al., 2003b] Hendrik P.a. Lensch, Jochen Lang, Asla M. Sa, and Hans-Peter Seidel. Planned Sampling of Spatially Varying BRDFs. *Computer Graphics Forum*, 22(3):473–482, September 2003.

[Levoy and Curless, 2002] Marc Levoy and Brian Curless. Stanford Spherical Gantry, 2002.

[Levoy and Hanrahan, 1996] Mark Levoy and Pat Hanrahan. Light Field Rendering. In *Proceedings of ACM SIGGRAPH 1996*, pages 31–42. Addison Wesley, August 1996.

[Lourakis and Argyros, 2004] M I A Lourakis and A A Argyros. The Design and Implementation of a Generic Sparse Bundle Adjustment Software Package Based on the Levenberg-Marquardt Algorithm. Technical Report 340, Institute of Computer Science - FORTH, Heraklion, Crete, Greece, 2004.

[Lu and Little, 1999] Jiping Lu and J.J. Little. Reflectance and shape from images using a collinear light source. *International Journal of Computer Vision*, 32(3):213–240, 1999.

[Ma et al., 2007] Wan-Chun Ma, Tim Hawkins, Pieter Peers, Charles-Felix Chabert, Malte Weiss, and Paul Debevec. Rapid Acquisition of Specular and Diffuse Normal Maps from Polarized Spherical Gradient Illumination. In *Rendering Techniques*, pages 183–194, 2007.

[Mallick et al., 2005] Satya Mallick, Todd Zickler, David Kriegman, and Peter Belhumeur. Beyond Lambert: Reconstructing Specular Surfaces Using Color. In *CVPR*, pages II: 619—-626, 2005.

[Marr and Poggio, 1976] D Marr and T Poggio. Cooperative Computation of Stereo Display. Technical Report AIM-364, MIT Artificial Intelligence Laboratory, June 1976.

[Marschner et al., 1999] Stephen Marschner, Stephen Westin, Eric Lafortune, Kenneth Torrance, and Donald Greenberg. Image-Based {BRDF} Measurement Including Human Skin. In *Eurographics Workshop on Rendering*, pages 131–144, 1999.

[Marschner et al., 2000] Stephen Marschner, Stephen Westin, Eric Lafortune, and Kenneth Torrance. Image-Based Bidirectional Reflectance Distribution Function Measurement. *Applied Optics*, 39(16):2592–2600, June 2000.

[Marschner et al., 2005] Stephen R Marschner, Stephen H Westin, Adam Arbree, and Jonathan T Moon. Measuring and Modeling the Appearance of Finished Wood. *Computer Graphics (SIGGRAPH 2005 Proceedings)*, 24(3):727–734, August 2005.

[Marschner, 1998] S.R. Marschner. *Inverse rendering for computer graphics*. Ph.d., Cornell University, 1998.

[Masselus et al., 2002] Vincent Masselus, Philip Dutré, and Frederik Anrys. The free-form light stage. In *ACM SIGGRAPH 2002 conference abstracts and applications*, page 262, New York, New York, USA, 2002. ACM.

[Matusik et al., 2003] Wojciech Matusik, Hanspeter Pfister, Matthew Brand, and Leonard McMillan. Efficient Isotropic {BRDF} Measurement. In *Proceedings of the 14th Euographics Workshop on Rendering*, pages 241–248. Eurographics Association, 2003.

[McAllister, 2002] David McAllister. *A Generalized Surface Appearance Representation for Computer Graphics*. PhD thesis, Department of Computer Science, University of North Carolina - Chapel Hill, June 2002.

[Mukaigawa et al., 2010] Yasuhiro Mukaigawa, Y Yagi, and R Raskar. Analysis of light transport in scattering media. *CVPR*, 2010.

[Müller et al., 2005] Gero Müller, Gerhard Bendels, and Reinhard Klein. Rapid Synchronous Acquisition of Geometry and Appearance of Cultural Heritage Artefacts. In *The 6th International Symposium on Virtual Reality, Archaeology and Cultural Heritage*. Eurographics Association, 2005.

[Nayar et al., 2006] Shree Nayar, Gurunandan Krishnan, Michael Grossberg, and Ramesh Raskar. Fast Separation of Direct and Global Components of a Scene Using High Frequency Illumination. *ACM Transactions on Graphics*, 25(3):935–944, July 2006.

[Nelder and Mead, 1965] J. A. Nelder and R. Mead. A simplex method for function minimization. *The computer journal*, 7:308–313, July 1965.

[Ngan et al., 2005] Addy Ngan, Frédo Durand, and Wojciech Matusik. Experimental Analysis of {BRDF} Models. In *Eurographics Symposium on Rendering*, pages 117–126. Eurographics Association, 2005.

[Nicodemus et al., 1977] F.E. Nicodemus, J.C. Richmond, J.J. Hsia, IW Ginsberg, and T. Limperis. Geometrical considerations and nomenclature for reflectance. *NBS monograph*, 160(October):201–231, 1977.

[Peers et al., 2006] Pieter Peers, Karl vom Berge, Wojciech Matusik, Ravi Ramamoorthi, Jason Lawrence, Szymon Rusinkiewicz, and Philip Dutré. A compact factored representation of heterogeneous subsurface scattering. *ACM Transactions on Graphics*, 25(3):746, July 2006.

[Romeiro and Zickler, 2010] Fabiano Romeiro and Todd Zickler. Blind Reflectometry. *ECCV*, 2010.

[Romeiro et al., 2008] Fabiano Romeiro, Y Vasilyev, and Todd Zickler. Passive Reflectometry. *Proc. ECCV*, 2008.

[Rusinkiewicz, 1998] Szymon M Rusinkiewicz. A New Change of Variables for Efficient BRDF Representation. *EuroGraphics Workshop on Rendering*, page 11, 1998.

[Salvi et al., 2010] Joaquim Salvi, Sergio Fernandez, Tomislav Pribanic, and Xavier Llado. A state of the art in structured light patterns for surface profilometry. *Pattern Recognition*, 43(8):2666–2680, August 2010.

[Sato et al., 1997] Yoichi Sato, Mark Wheeler, and Katsushi Ikeuchi. Object Shape and Reflectance Modeling from Observation. In *Proceedings of ACM SIGGRAPH 1997*, pages 379–388. Addison Wesley, 1997.

[Schlick, 1994] Christophe Schlick. A Survey of Shading and Reflectance Models. *Computer Graphics Forum*, 13(2):121–131, June 1994.

[Seitz et al., 2006] Steven Seitz, Brian Curless, James Diebel, Daniel Scharstein, and Richard Szeliski. A Comparison and Evaluation of Multi-View Stereo Reconstruction Algorithms. In *CVPR*, pages I: 519—-528, 2006.

[Shewchuk, 1996] J. Shewchuk. Triangle: Engineering a 2D quality mesh generator and Delaunay triangulator. *Applied Computational Geometry Towards Geometric Engineering*, pages 203–222, 1996.

[Srinivasan et al., 1985] V Srinivasan, HC Liu, and Maurice Halioua. Automated phase-measuring profilometry: a phase mapping approach. *Applied Optics*, 24(2):185–188, 1985.

[Strand and Taxt, 1999] J Strand and T Taxt. Performance evaluation of two-dimensional phase unwrapping algorithms. *Applied optics*, 38(20):4333–44, July 1999.

[Tagare and DeFigueiredo, 1991] H D Tagare and R J P DeFigueiredo. A Theory of Photometric Stereo for a Class of Diffuse non-Lambertian Surfaces. *IEEE Pattern Analysis and Machine Intelligence*, 13(2):133–152, February 1991.

[Toler-Franklin et al., 2007] C. Toler-Franklin, A. Finkelstein, and S. Rusinkiewicz. Illustration of complex real-world objects using images with normals. In *Proceedings of the 5th*

international symposium on Non-photorealistic animation and rendering, pages 111–119. ACM, 2007.

[Tomasi and Kanade, 1992] C Tomasi and T Kanade. Shape and Motion from Image Streams under Orthography: {A} Factorization Method. *International Journal of Computer Vision*, 9(2):137–154, November 1992.

[Torrance and Sparrow, 1967] Kenneth Torrance and Edward Sparrow. Theory for off-specular reflection from roughened surfaces. *Journal of Optical Society of America*, 57(9):1105–1114, 1967.

[Touloukian, 1970] Y. S Touloukian. *Thermophysical properties of matter*. 1970.

[Tyo et al., 1996] JS Tyo, MP Rowe, EN Pugh Jr, and N Engheta. Target detection in optically scattering media by polarization-difference imaging. *Applied Optics*, 35(11):1855–1870, 1996.

[Valkenburg and McIvor, 1998] R Valkenburg and A McIvor. Accurate 3d Measurement Using a Structured Light System. *Image and Vision Computing*, 16(2):99–110, February 1998.

[Wang and Dana, 2006] Jiaping Wang and Kristin Dana. Relief Texture from Specularities. *PAMI*, 28(3):446–457, March 2006.

[Ward, 1992] Gregory J Ward. Measuring and Modeling Anisotropic Reflection. In *Computer Graphics (SIGGRAPH '92 Proceedings)*, volume 26, pages 265–272, July 1992.

[Weyrich et al., 2006] Tim Weyrich, Wojciech Matusik, Hanspeter Pfister, Bernd Bickel, Craig Donner, Chien Tu, Janet McAndless, Jinho Lee, Addy Ngan, Henrik Jensen, and Markus Gross. Analysis of Human Faces Using a Measurement-based Skin Reflectance Model. *ACM Transactions on Graphics*, 25(3):1013–1024, 2006.

[Weyrich et al., 2007] Tim Weyrich, Jason Lawrence, Hendrik P. A. Lensch, Szymon Rusinkiewicz, and Todd Zickler. Principles of Appearance Acquisition and Representation. *Foundations and Trends in Computer Graphics and Vision*, 4(2):75–191, 2007.

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

[Wu et al., 2010] Lun Wu, Arvind Ganesh, Boxin Shi, Yasuyuki Matsushita, Yongtian Wang, and Y. Ma. Robust Photometric Stereo via Low-Rank Matrix Completion and Recovery. *Asian Conference on Computer Vision*, 2010.

[Zhang et al., 2003] Li Zhang, Brian Curless, and Steven Seitz. Spacetime stereo: shape recovery for dynamic scenes. In *Proceedings of IEEE Computer Vision and Pattern Recognition 2003*, pages 367–374, 2003.

[Zhang et al., 2006] Jeffrey Zhang, D.G. Aliaga, M. Boutin, and Robert Insley. Angle independent bundle adjustment refinement. In *3D Data Processing, Visualization, and Transmission, Third International Symposium on*, pages 1108–1116, 2006.

[Zhang, 1994] Zhengyou Zhang. Iterative point matching for registration of free-form curves and surfaces. *International Journal of Computer Vision*, 13(2):119–152, October 1994.

[Zhang, 2000] Z. Zhang. A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(11):1330–1334, 2000.

[Zickler et al., 2002] Todd Zickler, Peter Belhumeur, and David Kriegman. Helmholtz Stereopsis: Exploiting Reciprocity for Surface Reconstruction. *International Journal of Computer Vision*, 49(2-3):215–227, 2002.

[Zickler et al., 2005] Todd Zickler, Sebastian Enrique, Ravi Ramamoorthi, and Peter Belhumeur. Reflectance Sharing: Image-based Rendering from a Sparse Set of Images. In *Rendering Techniques*, pages 253–264, 2005.

[Zickler, 2006] Todd Zickler. Reciprocal Image Features for Uncalibrated Helmholtz Stereopsis. In *IEEE Computer Vision and Pattern Recognitiion*, pages II: 1801—-1808, 2006.