

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

Michael Holroyd
University of Virginia

Jason Lawrence
University of Virginia

Todd Zickler
Harvard University

Abstract

We present a novel optical setup and processing pipeline for measuring the 3D geometry and spatially-varying surface reflectance of physical objects. Central to our design is a digital camera and a high frequency spatially-modulated light source aligned to share a common focal point and optical axis. Pairs of such devices allow capturing a sequence of images from which precise measurements of geometry *and* reflectance can be recovered. Our approach is enabled by two technical contributions: a new active multiview stereo algorithm and an analysis of light descattering that has important implications for image-based reflectometry. We show that the geometry measured by our scanner is accurate to within 50 microns at a resolution of roughly 200 microns and that the reflectance agrees with reference data to within 5.5%. Additionally, we present an image relighting application and show renderings that agree very well with reference images at light and view positions far from those that were initially measured.

1 Introduction

Systems for digitizing the appearance of physical objects have applications in fields as diverse as entertainment, manufacturing, and archeology. Although 3D scanners capable of sub-millimeter geometric accuracy are now commonplace, those which allow simultaneously measuring an object’s directionally- and spatially-varying appearance are still a topic of intense research. Many applications, including digital restoration of cultural artifacts, analysis of surface finishes, and photorealistic rendering, would benefit from having tools for efficiently acquiring both of these components.

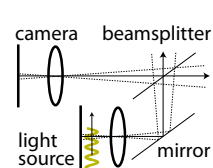
We present a system for recovering aspects of object appearance that, at a pre-chosen scale, are well-represented by a surface and a spatially-varying bi-directional reflectance distribution function (spatially-varying BRDF) [Nicodemus et al. 1977]. For suitably opaque surfaces, this provides an accurate and efficient digital appearance model, and for objects with translucency it can be complemented by additional terms that describe the non-local aspects of appearance, such as sub-surface scattering. In any case, obtaining accurate shape and spatially-varying BRDF (referred to as *reflectance* hereafter) is an essential ingredient of many appearance scanning pipelines. Our goal is to create a system that acquires these measurements with high accuracy and repeatability.

An alternative approach to the one pursued here would be to employ a less structured appearance model, such as a “reflectance field” that is defined on coarsely-approximated geometry [Debevec et al. 2000]. However, estimating accurate geometry allows more efficient object scanning in many cases. It allows one to predict the motion and deformation of sharp specular highlights, for example,

without having to collect dense angular measurements everywhere on the surface. Additionally, recovering explicit models of the geometry and reflectance enables post-processing, such as animating the object or editing and analyzing its material properties, that is difficult or impossible with alternative representations.

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 many of 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 straightforward to capture a sequence of images of an object from different viewpoints under time-varying sinusoidal illumination originating from different locations. Through two technical contributions—a new active multiview stereo algorithm and a theoretical analysis of light descattering—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.

2 Design Rationale and Prior Work

Our system builds on several recent advances in multiview stereo, unstructured space-time stereo, phase-based profilometry, and optical descattering; and its design reflects three main considerations.

Simultaneous recovery of geometry and reflectance. Measuring surface reflectance is inherently an optical task, and it requires images of an object whose geometry is known. In many prior systems, this involves a two-step process whereby the geometry of the object is estimated prior to reflectometry. Examples include the use of laser-stripe triangulation [Sato et al. 1997; Goesele et al.

2004], structured lighting [Debevec et al. 2000; Weyrich et al. 2006], transmissive (CT) scanning [Lensch et al. 2003], shape-from-silhouettes [Müller et al. 2005], and human assisted photogrammetric modeling [Yu et al. 1999]. In all of these two-step approaches, the images that are used for reflectometry must be registered to the 3D model, and although this 2D-3D alignment problem has been studied [Lensch et al. 2001], it remains difficult in practice and invariably leads to corrupted reflectance measurements near misalignments.

To overcome this limitation, we seek a system that incorporates a unified optical design in which range and reflectance are measured in the same raster grid. Although we still must register 3D scans to one another, we avoid the more difficult 2D-3D registration problem, thereby improving final model quality.

Support for a wide range of materials and shapes. We seek a system that provides unbiased measurements of shape and reflectance. In particular, we avoid methods based on photometric stereo, which jointly estimate local surface orientation (surface normals) along with reflectance parameters [Georghiadis 2003; Goldman et al. 2005] since these are inherently limited by the degree to which a surface obeys the chosen analytic reflectance model. For example, recent systems based on robust Lambertian photometric stereo can provide very accurate reconstructions in some cases [Hernández et al. 2008], but as the object reflectance deviates from the Lambertian model, the errors in the resulting geometry (and any subsequent reflectance samples) become hard to characterize and predict. Even non-parametric methods require certain properties in the underlying reflectance such as isotropy [Alldrin et al. 2008] and symmetry [Holroyd et al. 2008; Ma et al. 2007; Ghosh et al. 2009]. Finally, few attempts have been made to compute full 360° models from 2.5D height height fields acquired with these techniques [Esbey et al. 2008].

Instead, we build on advances in phase-based profilometry [Srinivasan et al. 1985] (also referred to as phase mapping or interferometry) and space-time stereo [Zhang et al. 2003; Davis et al. 2005], that have been shown to provide accurate geometry for a wide variety of materials. In particular, we exploit the observation that high frequency sinusoidal illumination can be used to separate local reflections from non-local ones [Nayar et al. 2006] and simultaneously allow the recovery of accurate geometry for objects exhibiting translucency and diffuse interreflections [Chen et al. 2007; Chen et al. 2008].

Note that while our system prioritizes generality, it also prioritizes practicality, and it does both within the confines of an appearance model comprised of shape and reflectance. It is not intended for objects with extremely low albedo, perfect mirror-like reflectance, or refraction. (These are each better handled by alternative systems; see [Ihrke et al. 2008].)

Simple and stable calibration. A major barrier to achieving an accurate scanner *that is practical* is the effort and expertise required to calibrate it before each scan. This typically includes estimating the position, orientation, and internal properties of the cameras and sources, as well as their radiometric sensitivity and output. It often requires an expert to frequently repeat calibration routines that involve geometric and radiometric calibration targets such as checkerboard patterns, mirrored spheres, and diffuse reflectance standards.

In contrast, we seek a system with off-line geometric and radiometric calibration requirements that are manageable for a non-expert, and that rarely (if ever) need to be repeated. We do this by leveraging recent developments in feature-based camera calibration algorithms [Brown and Lowe 2005; Furukawa and Ponce 2008] and by using devices that reduce the number of calibration parameters and the difficulty of estimating them.

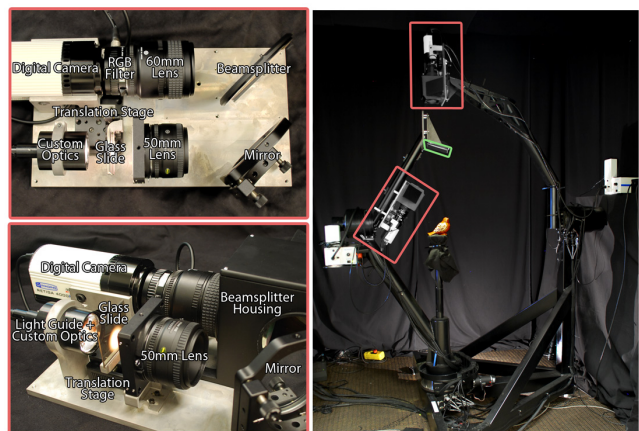


Figure 1: Our prototype scanner consists of two identical coaxial camera/lightsource assemblies (left) mounted to the arms of a four-axis spherical gantry (right). The coaxial device mounted on the left arm views the object through a mirror (green) to achieve roughly equal stand-off distances between the two arms. The light is provided by tungsten-halogen sources with computer-controlled mechanical shutters (white boxes in image on right).

3 System Overview

The basic building block of our system, which is depicted in Figure 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 and Nayar 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 Lowe 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. 2003].

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 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 and Nayar 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. This is enabled by two key technical contributions of this paper:

1. *Multiview phase mapping.* 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.
2. *Local reflectance from sinusoidal illumination.* As shown in previous work [Nayar et al. 2006], shifted sinusoidal illumination allows separating local or “direct” reflections from sub-surface scattering, diffuse interreflections, and other non-local “indirect” effects; and it is tempting to interpret the resulting direct component as samples of a BRDF at the implied scale. We show that this interpretation is correct only after applying a significant correction factor that depends on the local light and view directions, as well as the intrinsic and extrinsic parameters of the measurement devices. A detailed derivation of this correction factor is provided in an associated technical report [Holroyd et al. 2010], and it is validated experimentally.

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

4 Experimental Setup

Figure 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 Nikon 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

Nikon 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 3). 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 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.

4.1 Calibration

Geometric calibration. We used Zhang’s chart-based algorithm [2000] to estimate the intrinsic parameters of each camera: focal length, principal point, and radial and tangential distortion coefficients. The gantry arm angles are then used to estimate the position of each coaxial device to within about 5 pixel reprojection error over the working volume. This rough calibration is improved for each scan using an automatic feature-based refinement procedure described in Section 5.3. Note that no further steps must be taken to estimate the positions of the light sources since 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.

Radiometric calibration. 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: $\sigma(u)$, $u \in \mathbb{R}^2$. 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 $l(x)$, $x \in \mathbb{R}^3$. 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 $l(x)$.

We first estimate the camera sensitivity functions for each device following the procedure illustrated in Figure 2. 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

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

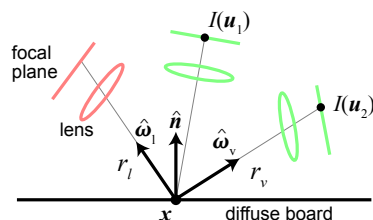


Figure 2: Illustration of our procedure for estimating the camera sensitivity and light emission functions.

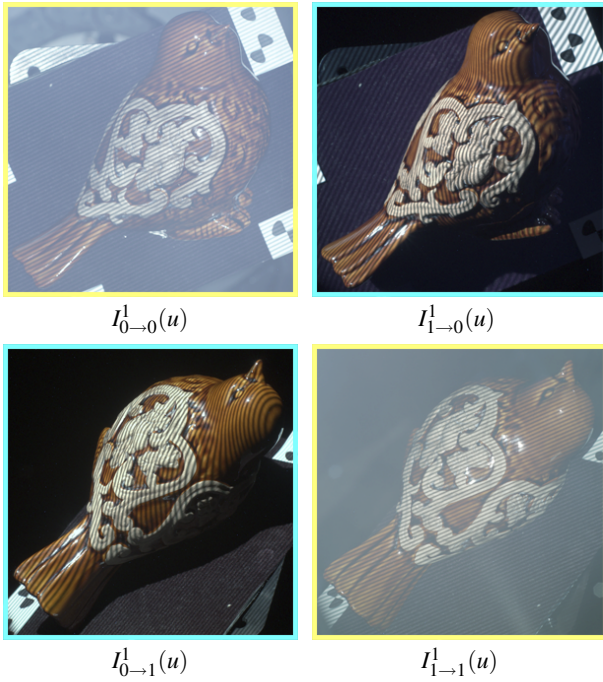


Figure 3: Some of the raw images acquired in one scan. The yellow bordered images are captured under co-axial lighting and the blue bordered images correspond to a reciprocal image pair. Note the high-frequency sinusoidal illumination.

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 $I(u)$ can be written as

$$I(u) = \sigma(u)l(x)(\hat{n} \cdot \hat{\omega}_1), \quad (1)$$

where the $1/r_1^2$ term is absorbed into the 3D function $l(x)$ as mentioned above. Each point in a discrete set of points on the board induces a linear constraint on the sensitivity function: $\sigma(u_1)/I(u_1) - \sigma(u_2)/I(u_2) = 0$. Because we expect $\sigma(u)$ to be smooth, we model it as a 10th-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 $l(x)$. Note that dividing $I(u_1)$ by $\sigma(u_1)$ and $(\hat{n} \cdot \hat{\omega}_1)$ gives a value of $l(x)$ 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 $l(x)$ 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 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-

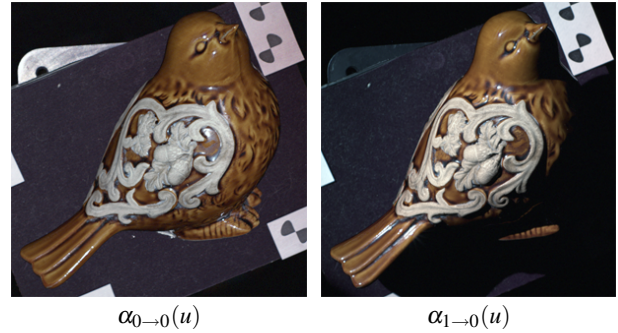


Figure 4: Amplitude maps computed from the stacks along the top row in Figure 3. Note that the strong haze in the coaxial image has been removed. Also, the slight color differences are due to the varying radiometric properties of the light sources and cameras.

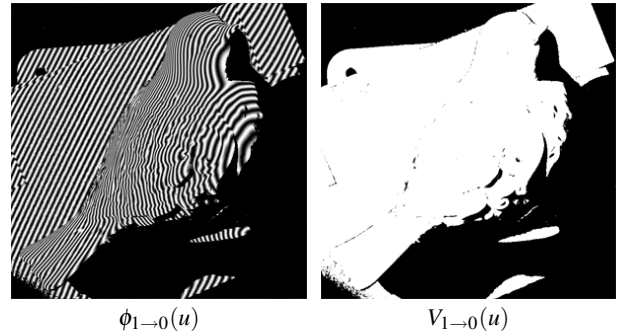


Figure 5: Phase map and visibility map computed from the image stack in the top right corner in Figure 3. The cosine of the phase at each pixel is shown as a grayscale image.

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 5.1).
2. From these images, we perform descattering and compute the apparent sinusoid phase and amplitude at each pixel (Section 5.2).
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).
4. We obtain a dense depth map for each scan using a new multi-view phase mapping algorithm (Section 5.4).
5. We align and merge multiple depth maps into a single watertight model (Section 5.5).
6. Finally, we extract BRDF samples for each vertex in the watertight mesh by applying appropriate radiometric correction factors (Section 5.6).

5.1 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 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 3. We denote images by $I_{i \rightarrow j}^k(u)$, where $i \in [0 \dots N]$ in-

dexes the light source position, $j \in [0 \dots N]$ indexes the camera position, and $k \in [1 \dots M]$ 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 3, 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 beam-splitter, but as described below and shown in Figure 4, this haze is conveniently removed by optical descattering.

Images where $i \neq j$ comprise reciprocal pairs in which i and j are swapped: $(I_{i \rightarrow j}, I_{j \rightarrow i})$. These play an important role in our pose refinement (Section 5.3) and multiview stereo algorithm (Section 5.4).

5.2 Amplitude and Phase Maps

Translations of the glass slide induce time-varying sinusoidal irradiance at the object’s surface. Recent work has demonstrated that for many scenes 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 7.

We model the response at one pixel as: $I^k(u) = \alpha(u) \cos(\gamma t_k + \phi(u)) + \beta(u)$, 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:

$$\begin{bmatrix} \cos(\gamma t_1) & -\sin(\gamma t_1) & 1 \\ \vdots & \vdots & \vdots \\ \cos(\gamma t_M) & -\sin(\gamma t_M) & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} I^1(u) \\ \vdots \\ I^M(u) \end{bmatrix} \quad (2)$$

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

Note that the pseudo-inverse of the coefficient matrix can be pre-computed 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 4 shows the $\alpha(u)$ computed from the image stacks along the top row in Figure 3. Note that the sinusoidal pattern has been removed along with diffuse interreflections (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 5.6). Figure 5 shows the phase $\phi(u)$ computed from the image stack in the top right corner of Figure 3. We also compute a binary visibility map $V(u)$ by applying a threshold to $\alpha(u)$. 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 5.4).

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 5.6 when we describe our procedure for recovering measurements of the BRDF from $\alpha(u)$. The displacements t_k should be chosen so that the system in Equation 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.

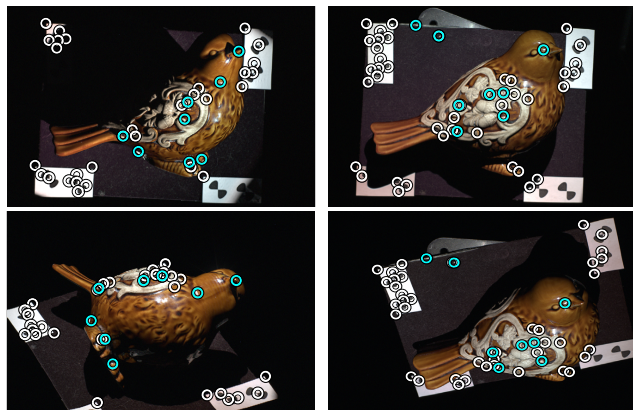


Figure 6: Feature-based pose refinement. Each column shows a reciprocal image pair, with reference view in the top row. By detecting intrinsic texture features (white) and reciprocal specular features (cyan), the geometric calibration is automatically refined from an average reprojection error of 5 pixels to 0.3 pixels.

5.3 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 $\alpha(u)$ 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 6 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 Lowe 2005] or employ a statistical outlier rejection method. Finally, we use the Sparse Bundle Adjustment (SBA) package developed by Lourakis

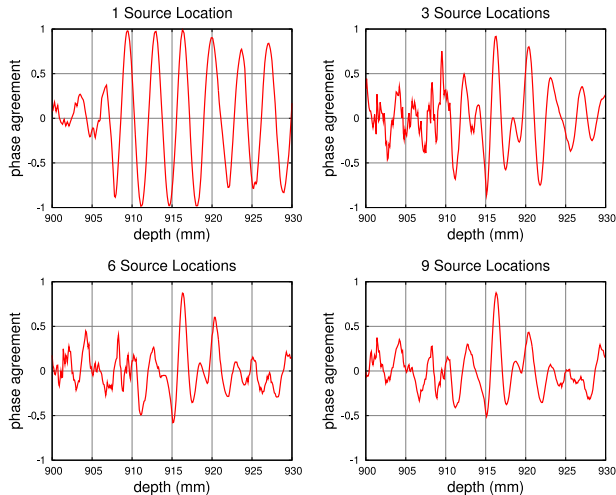


Figure 7: The search for depth correspondences becomes more stable with more auxiliary frames. We recommend $N \geq 7$ in practice.

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.

5.4 Multiview Phase Mapping for Depth Estimation

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 5.2. 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].

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

$$\phi_{0 \rightarrow 0}(u_0) = \phi_{0 \rightarrow i}(u_i) \quad \phi_{i \rightarrow 0}(u_0) = \phi_{i \rightarrow i}(u_i).$$

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 x along each ray

$$\varepsilon(x) = \frac{\sum_{i=1}^N p_i V_{i \rightarrow 0}(u_0)}{\sum_{i=1}^N V_{i \rightarrow 0}(u_0)}, \quad (3)$$

where

$$p_i = \frac{(\phi_{0 \rightarrow 0}(u_0) \cdot \phi_{0 \rightarrow i}(u_i)) + (\phi_{i \rightarrow 0}(u_0) \cdot \phi_{i \rightarrow i}(u_i))}{2}.$$

The expression $(\phi_a(x) \cdot \phi_b(x))$ denotes the inner product of the two vectors associated with these phase values. Note that $\varepsilon(x)$ 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 $\varepsilon(x)$ to become more peaked around the correct depth as false matches are eliminated. This is illustrated in the graphs in Figure 7 which plot $\varepsilon(x)$ as a function of the depth at one of the pixels in the reference view in Figure 3 over a 3cm interval that straddles the surface. We compute the position x that maximizes $\varepsilon(x)$ in two steps. First, we compute values of $\varepsilon(x)$ 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 $\varepsilon(x)$ 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 6).

We record depth values only at pixels that are visible in at least 3 views and for which $\varepsilon(x) > 0.5$. 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).

5.5 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.

5.6 Recovering BRDF Measurements

A significant advantage of using sinusoidal illumination is that the amplitude maps $\alpha_{i \rightarrow j}(u)$ 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]. It makes sense, then, that these amplitude maps can be used to obtain samples of the BRDF. The remainder of this section describes a significant correction factor that must be applied when relating amplitudes to BRDF values.

Amplitude loss. We have 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 [Holroyd et al. 2010]. Figure 8 illustrates the geometric setup and key notation used in our analysis. We briefly summarize the main result here.

Let A_{surf} denote the portion of the scene that is imaged by a single camera pixel A_{pixel} . Under a pinhole camera model, the size and shape of A_{surf} is determined by the orientation $(\hat{n} \cdot \hat{\omega}_v)$ and distance r_v , along with the dimensions of A_{pixel} and the focal length f_c . Additionally, consider the image that A_{surf} forms on the focal plane of

the light source, A_{light} , which will depend as well on the orientation ($\hat{n} \cdot \hat{\omega}_l$), distance r_l , and focal length f_l . Due to the small extent of a single pixel, the region A_{light} is well-approximated by a parallelogram, as depicted in Figure 8, which is characterized by the lengths of its sides, a and b , its orientation ξ relative to the horizontal axis of the source focal plane, and its skew, which we parameterize here by the signed distance c . We show in the technical report that the area of this parallelogram is equal to

$$ab = w^2 \frac{(\hat{n} \cdot \hat{\omega}_l) r_c^2 f_l^2 \cos^3 \psi_c}{(\hat{n} \cdot \hat{\omega}_v) r_l^2 f_c^2 \cos^3 \psi_l}, \quad (4)$$

where w is the width of one camera pixel and ψ_l and ψ_c are the angles made by the rays leaving the light and camera with their respective optical axes (Figure 8). We further show that the amplitude of the cosine measured at A_{pixel} satisfies

$$\alpha(u) \propto l(x) \sigma(u) \rho(\hat{\omega}_l, \hat{\omega}_v) (\hat{n} \cdot \hat{\omega}_l) A', \quad (5)$$

with

$$A' = A \operatorname{sinc}(af \cos \xi) \operatorname{sinc}(bf \sin \xi + cf \cos \xi). \quad (6)$$

This result confirms that the values in $\alpha(u)$ are proportional to the product of the surface irradiance under point lighting $l(x)(\hat{n} \cdot \hat{\omega}_l)$ (recall that $1/r_l^2$ is captured by $l(x)$), the amplitude of the projected sinusoid A , the BRDF $\rho(\hat{\omega}_l, \hat{\omega}_v)$, and the camera sensitivity function $\sigma(u)$. However, it also predicts a less obvious effect we call ‘‘amplitude loss’’ (Equation 6) 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 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 6 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 8 confirm this effect and validate our analytic model. They show the amplitude measured at a camera pixel as it moves toward the horizon ($\hat{n} \cdot \hat{\omega}_v \rightarrow 0$) while its up vector remains in the epipolar plane with a stationary overhead light ($\hat{n} \cdot \hat{\omega}_l = 1$). For this ‘‘in-plane’’ configuration, $c = 0$ and so the parallelogram reduces to a rectangle. The two graphs correspond to the sine wave being perpendicular and parallel to the plane of motion, respectively. These graphs also include predictions by a numerical simulation which agree with our analytic model exactly. Note that this amplitude loss can be significant. In the case where $\xi = 0$, a roughly 20% decrease is observed at an elevation angle of 60 degrees which falls off to roughly 90% at 80 degrees.

Based on this analysis, we harvest BRDF samples as follows. Each vertex in the merged model is projected into the amplitude maps $\alpha_{i \rightarrow j}(u)$ 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 ≥ 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 $\sigma(u)$, $l(x)$, $(\hat{n} \cdot \hat{\omega}_l)$, and the product of the sinc functions in Equation 6. This produces BRDF measurements up to a single global scale factor that we do not attempt to estimate.

6 Results

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

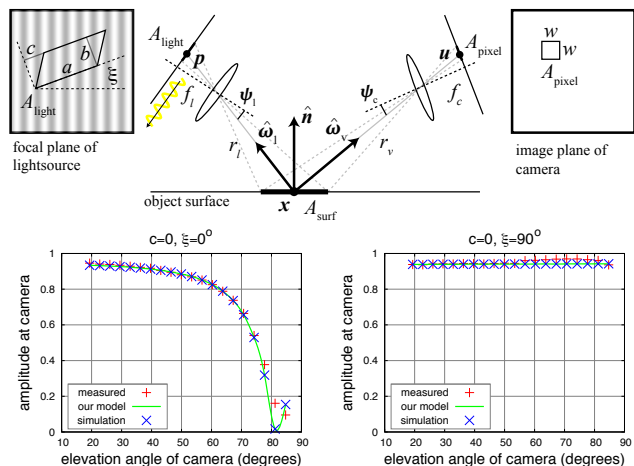


Figure 8: Top: Geometric setup and notation used in our analysis of ‘‘amplitude loss’’. **Bottom:** Measurements of the amplitude measured at one pixel as a camera moves towards the horizon with a fixed overhead light in a direction perpendicular and parallel to the projected sine wave, respectively. The measurements correspond to images of a Spectralon board corrected to account for deviations from a perfectly Lambertian reflector.

6.1 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 12 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 9 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 ~ 40 microns and the maximum error in any single scan is 0.25mm. The median error for the bird model is ~ 50 microns and the maximum in any scan is 0.25mm. Figure 9 also demonstrates the typical amount of coverage in one scan. Our 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 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.

6.2 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 10 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

²<http://www.xyzrgb.com>

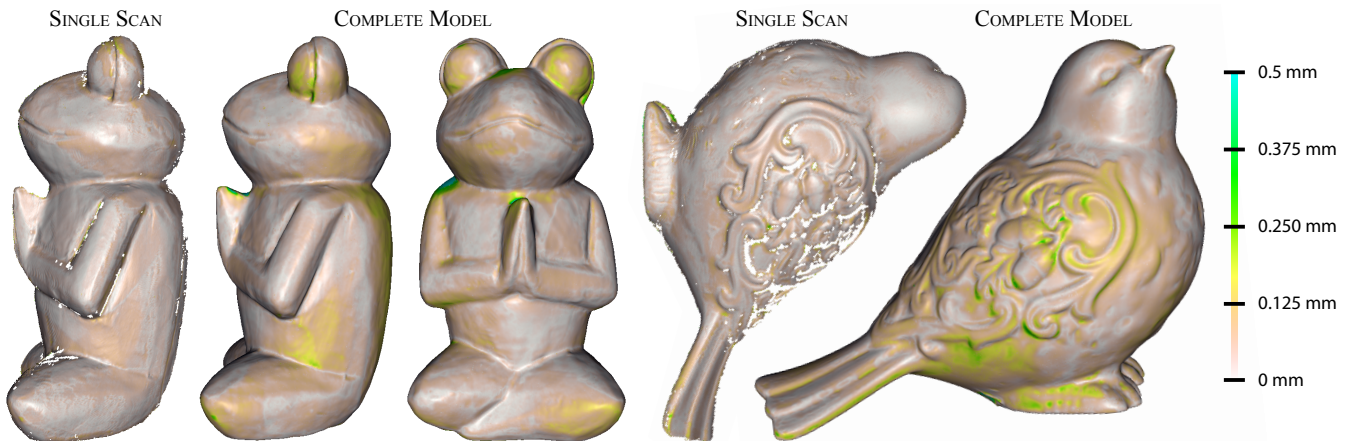


Figure 9: Error analysis of geometry acquired with our scanner for the frog and bird models. Comparisons are made to a reference surface obtained by coating these objects in a diffuse powder and scanning them with a high-end laser rangefinder.

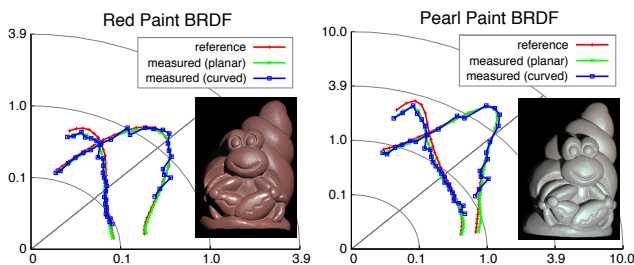


Figure 10: Polar plots of reflectance data acquired with our scanner for two types of spray paint compared to reference measurements. The curved samples are shown in the insets. The two lobes in each graph show in-plane measurements of changing view elevation angle, for a fixed light elevation angle equal to 20 degrees and 40 degrees, respectively. Intensity is plotted relative to a standard diffuse target.

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 7.

6.3 Geometry and Reflectance

We also analyzed the accuracy of our scanner by fitting an analytic spatially-varying BRDF to the models shown in Figure 12. 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

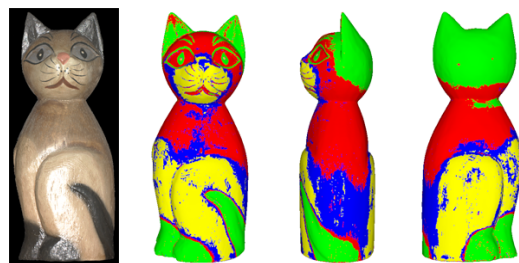


Figure 11: Material clusters computed for the `cat` model. A photograph is shown on the left next to false-color images showing the coverage of the four clusters over the object.

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. 2003; 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 11 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 12 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.



Figure 12: 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.

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 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.

6.4 Capture and Processing Times

Table 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 images at each auxiliary frame (Section 5.1), 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.

7 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 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 Lowe 2005]. Thus, future work might allow scanning outside a lab setting with rough placement of multiple co-axial assemblies.

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 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 ef-

fects in order to incorporate photometric constraints that would help further improve the quality of the resulting models.

This work also indicates new directions of research in view planning, alignment and merging. The location of auxiliary frames could be chosen adaptively, guided by an on-line reconstruction of the shape and reflectance of the target surface. 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. We will make all of the data that we have captured publicly available to support research in these areas.

8 Acknowledgements

We wish to thank the staff at OrbOptronix and Cyberware who helped with the design and assembly of our coaxial scanner. We also acknowledge the helpful and constructive comments from the reviewers. This research was supported by NSF CAREER Awards CCF-0747220 and IIS-0546408, NSF grant CCF-0811493, and a Sloan Research Fellowship.

References

- ALLDRIN, N., ZICKLER, T., AND KRIEGMAN, D. 2008. Photometric stereo with non-parametric and spatially-varying reflectance. In *Computer Vision and Pattern Recognition*.
- BROWN, M., AND LOWE, D. 2005. Unsupervised 3D object recognition and reconstruction in unordered datasets. In *Proc. Int. Conf. on 3D Digital Imaging and Modelling (3DIM)*.
- CHEN, T., LENSCH, H. P. A., FUCHS, C., AND SEIDEL, H.-P. 2007. Polarization and phase-shifting for 3d scanning of translucent objects. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.
- CHEN, T., SEIDEL, H.-P., AND LENSCH, H. P. A. 2008. Modulated phase-shifting for 3D scanning. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.
- COOK, R. L., AND TORRANCE, K. E. 1981. A reflectance model for computer graphics. *Computer Graphics (SIGGRAPH 1981)*, 7–24.
- DANA, K. J., VAN GINNEKEN, B., NAYAR, S. K., AND KOENDERINK, J. J. 1999. Reflectance and texture of real-world surfaces. *ACM Transactions on Graphics* 18, 1, 1–34.
- DAVIS, J., NEHAB, D., RAMAMOORTHI, R., AND RUSINKIEWICZ, S. 2005. Spacetime stereo: A unifying framework for depth from triangulation. *IEEE Trans. on Pattern Analysis and Machine Intelligence*.
- DEBEVEC, P. E., HAWKINS, T., TCHOU, C., DUIKER, H.-P., SAROKIN, W., AND SAGAR, M. 2000. Acquiring the reflectance field of a human face. In *SIGGRAPH*, 145–156.
- ESTEBAN, C. H., VOGIATZIS, G., AND CIPOLLA, R. 2008. Multiview photometric stereo. *IEEE Trans. Pattern Analysis and Machine Intelligence* 30, 3 (Mar.), 548–554.
- FURUKAWA, Y., AND PONCE, J. 2008. Accurate camera calibration from multi-view stereo and bundle adjustment. In *IEEE CVPR*.
- GARG, G., TALVALA, E.-V., LEVOY, M., AND LENSCH, H. P. A. 2006. Symmetric photography: Exploiting data-sparseness in reflectance fields. In *Proc. of the Eurographics Symposium on Rendering (EGSR)*.
- GEORGHIADES, A. 2003. Recovering 3-d shape and reflectance from a small number of photographs. In *Proceedings of the 14th Eurographics workshop on Rendering*, 230–240.
- GHOSH, A., ACHUTHA, S., HEIDRICH, W., AND O’TOOLE, M. 2007. Brdf acquisition with basis illumination. In *Proc. IEEE ICCV*.
- GHOSH, A., CHEN, T., PEERS, P., WILSON, C. A., AND DEBEVEC, P. E. 2009. Estimating specular roughness and anisotropy from second order spherical gradient illumination. *Comput. Graph. Forum* 28, 4, 1161–1170.
- GOESELE, M., LENSCH, H., LANG, J., FUCHS, C., AND SEIDEL, H. 2004. DISCO: acquisition of translucent objects. *ACM Transactions on Graphics (Proc. ACM SIGGRAPH)*, 835–844.

- GOLDMAN, D. B., CURLESS, B., HERTZMANN, A., AND SEITZ, S. M. 2005. Shape and spatially-varying BRDFs from photometric stereo. In *Proc. of the International Conference on Computer Vision (ICCV)*.
- GUILLEMAUT, J.-Y., DRBOHLAV, O., SARA, R., AND ILLINGWORTH, J. 2004. Helmholtz stereopsis on rough and strongly textured surfaces. In *3D Data Processing, Visualization and Transmission (3DPVT)*, 10–17.
- GUPTA, M., TIAN, Y., NARASIMHAN, S. G., AND ZHANG, L. 2009. (De)Focusing on Global Light Transport for Active Scene Recovery. In *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*.
- HAN, J. Y., AND PERLIN, K. 2003. Measuring bidirectional texture reflectance with a kaleidoscope. (*Proc. SIGGRAPH*) 22, 3, 741–748.
- HARRIS, C., AND STEPHENS, M. 1988. A combined corner and edge detector. In *Alvey Vision Conference*, 147–151.
- HERNÁNDEZ, C., VOGIATZIS, G., AND CIPOLLA, R. 2008. Multiview photometric stereo. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 548–554.
- HOLROYD, M., LAWRENCE, J., HUMPHREYS, G., AND ZICKLER, T. 2008. A photometric approach for estimating normals and tangents. *ACM Trans. on Graphics (Proc. of SIGGRAPH Asia)* 27, 5.
- HOLROYD, M., LAWRENCE, J., AND ZICKLER, T. 2010. A radiometric analysis of projected sinusoidal illumination for opaque surfaces. Tech. Rep. CS-2010-7, University of Virginia, May.
- IHRKE, I., KUTULAKOS, K. N., LENSCH, H. P. A., MAGNOR, M., AND HEIDRICH, W. 2008. State of the art in transparent and specular object reconstruction. In *STAR Proceedings of Eurographics*.
- KAZHDAN, M., BOLITHO, M., AND HOPPE, H. 2006. Poisson surface reconstruction. In *Proc. of the Eurographics Symposium on Geometry Processing*.
- LAWRENCE, J., BEN-ARTZI, A., DECORO, C., MATUSIK, W., PFISTER, H., RAMAMOORTHY, R., AND RUSINKIEWICZ, S. 2006. Inverse shade trees for non-parametric material representation and editing. *ACM Trans. Graph* 25, 3, 735–745.
- LENSCH, H. P. A., HEIDRICH, W., AND SEIDEL, H.-P. 2001. A silhouette-based algorithm for texture registration and stitching. *Graphical Models* 64, 4, 245–262.
- LENSCH, H. P. A., KAUTZ, J., GOESELE, M., AND HEIDRICH, W. 2003. Image-based reconstruction of spatial appearance and geometric detail. *ACM Transactions on Graphics* 22, 2.
- LOURAKIS, M., AND ARGYROS, A. 2004. The design and implementation of a generic sparse bundle adjustment software package based on the levenberg-marquardt algorithm. Tech. rep.
- LU, J., AND LITTLE, J. 1999. Reflectance and shape from images using a collinear light source. *Int. Journal of Computer Vision* 32, 3, 213–240.
- MA, W.-C., HAWKINS, T., PEERS, P., CHABERT, C.-F., WEISS, M., AND DEBEVEC, P. 2007. Rapid acquisition of specular and diffuse normals from polarized spherical gradient illumination. In *Eurographics Symposium on Rendering*.
- MARSCHNER, S. R. 1998. *Inverse Rendering for Computer Graphics*. PhD thesis, Cornell University.
- MÜLLER, G., BENDELS, G., AND KLEIN, R. 2005. Rapid synchronous acquisition of geometry and appearance of cultural heritage artefacts. In *Proc. of VAST*.
- NAYAR, S. K., KRISHNAN, G., GROSSBERG, M. D., AND RASKAR, R. 2006. Fast separation of direct and global components of a scene using high frequency illumination. *ACM Transactions on Graphics (Proc. SIGGRAPH)* 25, 3, 935–944.
- NICODEMUS, F., RICHMOND, J., HSIA, J., GINSBERG, I., AND LIMPÉRIS, T. 1977. Geometrical considerations and nomenclature for reflectance. *National Bureau of Standards Monograph* 160.
- SATO, Y., WHEELER, M. D., AND IKEUCHI, K. 1997. Object shape and reflectance modeling from observation. In *Proc. of SIGGRAPH*.
- SEITZ, S., CURLESS, B., DIEBEL, J., SCHARSTEIN, D., AND SZELISKI, R. 2006. A comparison and evaluation of multi-view stereo reconstruction algorithms. In *Proc. of IEEE ICCV*, 519–528.
- SRINIVASAN, V., LIU, H. C., AND HALIOUA, M. 1985. Automated phase-measuring profilometry: A phase mapping approach. *Applied Optics* 24, 185–188.
- STRAND, J., AND TAXT, T. 1999. Performance evaluation of 2d phase unwrapping algorithms. *Applied Optics* 38, 20, 4333–4344.
- WEYRICH, T., MATUSIK, W., PFISTER, H., BICKEL, B., DONNER, C., TU, C., MCANDLESS, J., LEE, J., NGAN, A., JENSEN, H. W., AND GROSS, M. H. 2006. Analysis of human faces using a measurement-based skin reflectance model. *ACM Trans. Graph* 25, 3, 1013–1024.
- YU, Y., DEBEVEC, P., MALIK, J., AND HAWKINS, T. 1999. Inverse global illumination: Recovering reflectance models of real scenes from photographs. In *Proc. of SIGGRAPH*.
- ZHANG, L., AND NAYAR, S. K. 2006. Projection defocus analysis for scene capture and image display. *ACM Trans. on Graphics (Proc. SIGGRAPH)* 25, 3.
- ZHANG, L., CURLESS, B., AND SEITZ, S. M. 2003. Spacetime stereo: Shape recovery for dynamic scenes. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.
- ZHANG, Z. 1994. Iterative point matching for registration of free-form curves and surfaces. *International Journal of Computer Vision* 13, 2, 119–152.
- ZHANG, Z. 2000. A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22, 11.
- ZICKLER, T., BELHUMEUR, P. N., AND KRIEGMAN, D. J. 2002. Helmholtz stereopsis: Exploiting reciprocity for surface reconstruction. *International Journal of Computer Vision* 49, 2-3, 215–227.
- ZICKLER, T. E. 2006. Reciprocal image features for uncalibrated helmholtz stereopsis. In *IEEE Computer Vision and Pattern Recognition or CVPR*, II: 1801–1808.