Shape and Spatially-Varying BRDFs From Photometric Stereo

Dan B Goldman[1] Brian Curless[1] Aaron Hertzmann[2] Steve Seitz[1]

[1]University of Washington [2]University of Toronto

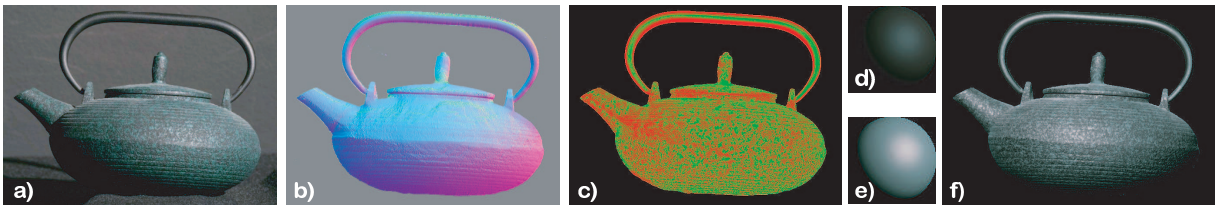

Figure 1: From (a) 10 photographs of an object taken under varying illumination, we can reconstruct (b) its normals and materials, represented as (c) a material weight map controlling a mixture of (d,e) fundamental materials. Using this representation we can (f) rerender the object under novel lighting.

Abstract

We describe a suite of techniques for extracting shape and materials from multiple photographs of an object captured from the same viewpoint but with differing illumination. Our method extracts per-pixel BRDFs along with 3D shape, under an assumption that the materials can be described as a convex combination of a small number of fundamental materials. We also show examples of interactive lighting and editing operations made possible by our methods, including direct manipulation of material coefficients and material transfer between models.

1 Introduction

Producers of computer graphic content rely heavily upon manual labor for creating digital models of 3D shape and materials. The ubiquity and affordability of high-resolution digital cameras has fueled an interest in acquiring these models from images. For opaque objects, the problem is to reconstruct shape (2D manifold) plus spatially-varying reflectance (6D) from just a few images. Expressed this way, the full problem of reconstructing an 8D function from sparse 2D samples is extremely challenging.

Previous authors have generally limited the scope of the problem by assuming either known shape or known reflectance. For example, laser range scanners work best when acquiring the shape of Lambertian objects of uniform color and intensity. To scan a shiny object with a laser range scanner, one often must coat it with a flat gray paint for scanning [Lensch et al. 2001; Wood et al. 2000; Sato et al. 1997]. Similarly, devices and techniques for acquiring bidirectional reflectance distribution functions (BRDFs) rely on having a flat surface sample, or at the very least a sample of known shape [Sato et al. 1997; Marschner et al. 1999]. However, for real-world objects – which are neither flat or homogenous nor Lambertian – this presents a "chicken or the egg" problem.

Our objective is to devise a very simple scanning technique that requires a minimal amount of equipment and effort but reconstructs both shape and spatially-varying materials for a wide range of real-world objects. We pose this problem in an optimization framework, in which both the objective function and optimization procedure must be carefully chosen. The wrong choice of objective function can lead to *overfitting*, in which the global optimum may match the data perfectly but fail to predict new samples (e.g. new lighting or viewing conditions). Due to the strong nonlinear terms in BRDFs, even a good choice of objective function presents a rocky optimization landscape, so the wrong choice of optimization procedure can end in local optima far from the global optimum. In this paper, we present an objective function and optimization procedure which generally avoids overfitting and finds a plausible optimum solution.

Our approach is based on the observation that many real-world surfaces actually consist of relatively few distinct materials. Under this assumption, we model the object's appearance in terms of a small number of constant *fundamental materials*, each described by the small number of parameters in the isotropic Ward reflectance model [Larson 1992]. The spatial variation of the surface is described by a constrained mixture of these materials across the surface. This model is narrow enough to eliminate many ambiguities of interpretation, but flexible enough to reasonably approximate a wide variety of real-world objects. We have also employed an effective mechanism that exploits user "hints" to resolve intrinsic ambiguities.

Our capture mechanism is inexpensive and fast, requiring only a high-resolution digital camera, tripod, and shutter release. The capture environment is a room with black walls and multiple independently controllable light sources, but, in principle, a closed garage and flashlight will suffice. We require a relatively small number of light sources in arbitrary locations (8 to 10 in our examples), but each additional widely spaced light source improves the quality of the reconstruction. Furthermore, we require no camera calibration, and lighting calibration is easily performed using simple objects of known shape and material.

Using the data acquired by our technique, many novel object editing techniques are made possible. For example, in addition to changes of lighting and limited changes of viewpoint, the fundamental materials of the object can be adjusted to change color, specularity or other parameters of the BRDF. The materials can be painted or adjusted using operations similar to those available in commercial image editing software such as Adobe Photoshop. Finally, a number of surface property transfer operations are made possible by our representation.

Our approach belongs to the class of methods known as photometric stereo. In our experience, photometric stereo methods can produce higher-resolution geometric detail than triangulation-based methods. To date, computer graphics research has almost solely emphasized triangulation for shape reconstruction, but we believe the high availability and low cost of high-resolution digital still cameras makes photometric stereo an appealing alternative for shape capture.

Figure 1 illustrates our model reconstruction and an example of relighting made possible by our approach.

2 Related Work

BRDF acquisition from photographic data has been widely researched in the computer graphics community. Sato et al. [1997] captured shape using a laser range scanner and separately captured varying BRDF measurements over the surface from photographs. Marschner et al. [1999] also took advantage of known shape – either from simple geometry or using laser range scans – to recover high-confidence BRDFs from homogenous surfaces.

More recent works attempt to improve upon this process or capture more BRDF variation across surfaces, but still assume that the geometry is known *a priori*. For example, Matusik et al. [2003] attempt to reduce the number of necessary photographs by determining the best viewpoints for BRDF recovery. A system that recovers BRDFs accounting for interreflections is described by Yu et al. [1999]. Lensch et al. [2001; 2003] have proposed a clustering approach that aggregates material estimates over a known surface. In this work they also refine the scanned geometry using extracted normal maps. However, they require approximate scanned geometry as input. The theoretical limits of factorizing lighting and reflectance have been explored by Ramamoorthi and Hanrahan [2001]. Another interesting variation was presented by Jaroszkiewicz and McCool [2003], in which BRDFs are specified via a painting interface.

Some work has been done to recover reflectance fields without knowing the geometry in advance [Gortler et al. 1996; Matusik et al. 2002], and even without reconstructing it at all [Debevec et al. 2000; Chuang et al. 2000; Levoy and Hanrahan 1996; Malzbender et al. 2001]. These approaches require vast numbers of observations and some do not explicitly capture shape or materials, and therefore have limited flexibility and more expensive representations. However, they can often capture more complex physical phenomena than our method. Gardner et al. [2003] presented a method for extracting both shape and reflectance from a set of photographic measurements for nearly flat objects. Akers et al. [2003] proposed an interactive relighting technique for emphasizing notable features in technical illustrations, by blending pixels from input images of the object under various illumination conditions.

Our work draws most immediately from the photometric stereo literature. In the seminal works [Woodham 1980; Silver 1980; Tagare and deFigueiredo 1991], the surface materials are assumed to be of constant BRDF, but more recent work has begun to extend the range of this approach to surfaces with more complex, spatially-varying BRDFs [Georghiades 2003; Hertzmann and Seitz 2003; Nayar et al. 1990]. In contrast to the approach taken by Hertzmann and Seitz [2003], we recover shape and material without requiring sample objects composed of the same material as the target. This allows us to reconstruct natural objects for which samples of known shape are not always available. Furthermore, our approach reconstructs BRDFs as well as shape, which allows rendering from other illuminants and viewing directions. Georghiades [2003] describes a method that accomodates variation of diffuse reflectance across a surface, but models the specular properties as being constant, whereas our linear combination model allows for variation of all parameters of the material across the surface. Our approach permits reconstruction of more complex surface properties, such as natural objects with variable surface roughness, or mixtures of man-made materials such as metallic and non-metallic paints.

We also note the success of Helmholtz stereopsis in shape reconstruction on arbitrary surfaces [Zickler et al. 2002; Tu and Mendonça 2003]. These methods avoid reconstructing BRDFs explicitly, by exploiting the reciprocity property of BRDFs. However, our approach recovers not only shape but also a BRDF representation that can be edited, relit and viewed from different viewpoints. Our method also uses a simpler capture mechanism requiring only a single uncalibrated camera position.

3 Problem Statement

The input to our system is a set of images of a static target object taken from a distant camera using a zoom lens under a different distant illuminant in each image. We assume the lighting is known, and only local illumination effects are present (no cast shadows, interreflections, transparency, or translucency). We assume that the surface does not contain significant depth discontinuities in the camera's view. We also provide the system with hand-drawn mattes of the target object to limit computation to interesting regions of the images. From these inputs, we seek to reconstruct shape and BRDFs with a constrained material model.

Our material model is motivated by the observation that real world variations in BRDF across a surface are often a result of the surface's composition from several different substances. For example, a block of wood with light and dark grain can be viewed as having two different substances that are blended in different amounts across the surface. This decomposition approach is also commonly used in the effects and animation industry to describe complex surfaces as spatially-varying mixtures of uniform materials [Cook 1984; Foley et al. 1990]. We call these substances *fundamental materials*, and the mixtures of these materials at each pixel are specified by *material weight maps*.

To reconstruct the surface and materials using this model, we provide the number of fundamental materials for which the system is to solve, and a few "hints" to constrain the solution and improve convergence, which will be discussed in later sections.

The output of the reconstruction system is a set of BRDF parameters for each of the fundamental materials, and a surface normal and material weights at each pixel. We can then reconstruct the surface by integrating the normal field.

3.1 Model

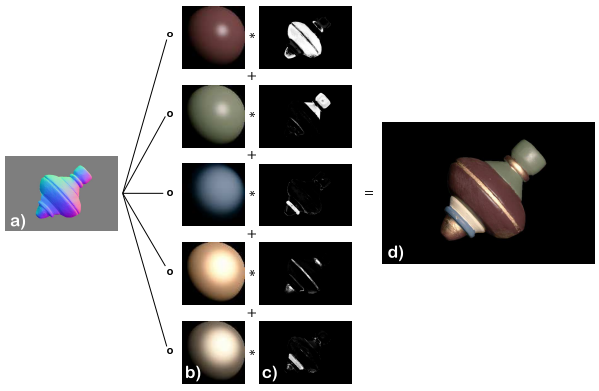

Figure 2: A schematic illustration of our appearance model, in which a normal map (a) is used to remap fundamental materials (b), and the results are modulated by material maps (c) and summed together (d).

We model the lights as distant directional sources, so the lighting direction $\mathbf{L}_i$ is constant over each image. We also model the camera

as orthographic, so the view direction $\mathbf{V}$ is constant for all samples (and is therefore elided in the formulation which follows). We can model the color at pixel p as if generated from a convex combination of fundamental materials:

$$I_{i,p,c} \leftarrow \sum_m \gamma_{p,m} f_c(\mathbf{n}_p, \mathbf{L}_i, \alpha_m) \tag{1}$$

We use the symbol $I_{i,p,c}$ to represent the color of channel c of pixel p of image i. The function f_c represents color channel c of the parameterized lighting model with normal $\mathbf{n}_p$, lighting condition $\mathbf{L}_i$ and BRDF parameter vector α_m; there is one α_m for each fundamental material. In this paper, we have used the isotropic Ward reflectance model because it has low dimensionality, but in theory other parametric models could be substituted.

Note that since our light sources are single distant directional illuminants, the function f_c is simply the BRDF, with a constant viewing direction. But in general we believe our approach could be applied to general distant light sources by allowing f_c to represent an integral over all sources.

The material weights $\gamma_{p,m}$ are subject to pairwise convexity constraints

$$\begin{aligned} 0 \le \gamma_{p,m} \le 1, \\ \textstyle\sum_m \gamma_{p,m} = 1, \\ \exists m_1, m_2 \quad \text{s.t.} \quad \gamma_{p,m_1} + \gamma_{p,m_2} = 1 \end{aligned} \tag{2}$$

Although the fundamental materials are constant over the image, the material weight maps γ are spatially varying. For example, Figure 1a shows a cast iron teapot with speckled green paint. Figure 1c shows two material weight maps (encoded in the red and green channels), and Figures 1d and e illustrate samples of the two fundamental materials, corresponding to the cast iron and green paint respectively.

Based on this model, we formulate the following objective function to solve for shape and materials:

$$Q(\mathbf{n}, \alpha, \gamma) = \sum_{i,p,c} \left(I_{i,p,c} - \sum_m \gamma_{p,m} f_c(\mathbf{n}_p, \mathbf{L}_i, \alpha_m) \right)^2 \tag{3}$$

in which Q is to be minimized with respect to the normals, material parameters, and material weight maps (denoted by $\mathbf{n}$, α and γ without subscripts, respectively).

Although this objective function is easy to express, it can lead to gross overfitting if the material weights γ are left unconstrained. We have found that the use of pairwise convexity constraints for these material weights (Equations 2) largely prevents this overfitting.

Optimization of Equation 3 is non-trivial; most black-box optimization algorithms become easily trapped in local minima due to the nonlinear terms of f and integrability constraints for the normal field $\mathbf{n}$. Our optimization approach, described in Section 4, is tailormade to avoid these problems.

4 Approach

Our approach has six components, each of which is described in detail in the sections which follow.

Light calibration. The light sources' directions and intensities are estimated using diffuse grey and chrome spheres captured under the same illumination as the target object. The calibration method is described in Section 4.1.

Initialize. We propose two different approaches to initialize normals and materials. One is appropriate for objects whose materials are similar and vary continuously, without uniform regions (Section 4.5.1), and the other is appropriate for objects with more widely-varying and distinct regions (Section 4.5.2). Combinations of these strategies are possible for objects of mixed type.

After initialization, we optimize the objective function iteratively by repeating the following three steps:

1. Compute surface normals and material weight maps. To compute the normals and material weight maps, while holding the BRDF parameters constant, we optimize Equation 3 jointly over normals and material weights. The normal optimization is performed as a discrete search, and the material weights are optimized by linear projection (Section 4.2).

2. Enforce integrability. The normals generated in the previous step are not guaranteed to be consistent with a 3D surface, so we enforce integrability by solving a Poisson equation to obtain a least-squares surface reconstruction, and subsequently recompute the normals (Section 4.3).

3. Optimize BRDF parameters. We optimize the BRDF parameters while holding the normals and material weights constant (Section 4.4).

Steps one through three are iterated until convergence. In some instances we found that an early result before termination gave slightly better reconstructions than the final converged result.

4.1 Light Calibration

In order to calibrate the directions of the light sources with respect to the camera and their relative intensities, we use two spherical calibration objects: a chrome and diffuse grey sphere. In principle only one of these – a chrome sphere – suffices to recover both direction and intensity of a distant light source: the direction is recovered by reflecting the viewing vector about the normal at the point of greatest brightness, and the relative intensity by integrating the measured radiance.

However, we have found that a large number of exposures may be necessary to obtain an intensity estimate with low variance from an image of a chrome sphere alone. A more economical method is to use a diffuse calibration object to measure the intensity, after the direction has been determined from the image of the chrome sphere. Thus, only one low-dynamic-range exposure for each calibration object is required to reconstruct both intensity and direction.

Given the correct lighting direction $\mathbf{L}_i$, the intensity of the diffuse sphere in image i is $I_{i,p} = \ell_i \rho \mathbf{n}_p^T \mathbf{L}_i$, where ρ is the diffuse albedo of the sphere and ℓ_i is the intensity of the light in image i. So given the known normals, we can solve for the relative light intensity $\ell_i \rho = \sum_p I_{i,p} / \sum_p \mathbf{n}_p^T \mathbf{L}_i$.

(If ρ is known, the absolute intensity can be recovered. However, in this paper we have used the relative intensities $\ell_i \rho$ to solve for materials, so the linear BRDF parameters we have recovered are actually scaled by the unknown constant scale factor ρ.)

These calibration objects may appear similar to those used by Hertzmann and Seitz [2003], but they are used in a different way.

In particular, we do not require or expect that the BRDFs of our target objects are composed of a linear combination of the calibration materials.

4.2 Computing Normals and Material Weight Maps

After initialization, covered in Section 4.5, we jointly optimize the normals ($\mathbf{n}_p$) and material weights ($\gamma_{p,m}$) of Equation 3. We first precompute the function f_c over a discrete sampling of normals $\mathbf{n}$ for each of the lighting samples $\mathbf{L}_i$ and fundamental material parameters α_m. In practice, this simply means rendering a small "virtual sphere" of each fundamental material under each lighting condition.

Given these samples of the appearance functions f, and the pairwise convex combination constraint, weights $\gamma_{p,m}$ are computed by linear projection and brute force search over all normals and all pairwise combinations of fundamental materials. Let $\phi_m^c(\mathbf{n}_p) = f_c(\mathbf{n}_p, \mathbf{L}_i, \alpha_m)$ denote the virtual sphere images for material m. For a given choice of normal $\mathbf{n}_p$ and pair of fundamental materials m_1, m_2, the objective function reduces to

$$Q_p = \sum_{i,c} \left(I_{i,p,c} - \gamma_{p,m_1} \phi^c_{m_1,i}(\mathbf{n}_p) - \gamma_{p,m_2} \phi^c_{m_2,i}(\mathbf{n}_p) \right)^2 \quad (4)$$

If $\gamma_{p,m_2} = 1 - \gamma_{p,m_1}$, this is minimized by differentiating with respect to γ_{p,m_1} and evaluating at zero:

$$\gamma_{p,m_1} \leftarrow \frac{\sum_{i,c} \left(I_{i,p,c} - \phi^c_{m_2,i}(\mathbf{n}_p) \right) \left(\phi^c_{m_2,i}(\mathbf{n}_p) - \phi^c_{m_1,i}(\mathbf{n}_p) \right)}{\sum_{i,c} \left(\phi^c_{m_2,i}(\mathbf{n}_p) - \phi^c_{m_1,i}(\mathbf{n}_p) \right)^2} \quad (5)$$

Since we wish to constrain the solution to convex combinations, we also clamp γ_{p,m_1} to lie between 0 and 1. We solve for these optimal convex weights for each normal and pair of materials, and select the pair and normal with the lowest objective Q_p.

Note that if three or more different materials in the scene are not linearly independent, the fundamental materials are not uniquely defined. For example, the object in Figure 3 has several materials which can plausibly described by many linear combinations. In such cases, we prefer the fundamental materials to correspond to specific materials visible in the scene, i.e. these pixels should have weights of 1 for one material and 0 for all others. To enforce this preference, the user may paint constraint maps which fix the weights of particular pixels throughout the optimization procedure. At such pixels, only the normal is optimized, again by brute force search over the normal samples. The middle and bottom rows of Figure 3 show the fundamental materials solved with and without such material weight constraints, respectively. (See Figure 7b to see the painted constraints for this object.)

4.2.1 Acceleration

Depending on the resolution of the virtual calibration objects, the full brute force normal search described in Section 4.2 can be quite slow; a single pass over the image may take several hours to complete.

In order to accelerate the computation, we limit the brute-force search to normals which lie close to the previous normal for each pixel. This can result in small areas which become "trapped" at the wrong normal values, so after the algorithm converges, we perform a final pass of normal/weight optimization using the full global normal search.

Figure 3: Fundamental materials recovered for the ornament model (top image) without weight constraints (middle row), and with weight constraints (bottom row).

Although in principle this approximation could cause the algorithm to converge to a less-than-optimal solution, in practice we have found it gives good results with dramatically less computation than the full normal search. Most of the examples shown here converged after 10-20 iterations of the full alternating optimization algorithm, in about 5-10 hours on a 2.8GHz Xeon processor.

In addition, we also have found it useful to disable the weight constraints in this final cleanup pass. This removes discontinuities along boundaries between constrained and unconstrained regions.

4.3 Enforcing Integrability

To compute a 3D surface from the estimated surface orientations, given the normal $\{n_x, n_y, n_z\}$ for each point, we solve for the height field $z(x,y)$ that minimizes

$$\Psi(z) = \sum_{x,y} \left(n_z \frac{\partial z(x,y)}{\partial x} + n_x \right)^2 + \left(n_z \frac{\partial z(x,y)}{\partial y} + n_y \right)^2 \quad (6)$$

using the approximations $\frac{\partial z(x,y)}{\partial x} = (z(x+1,y) - z(x,y)), \frac{\partial z(x,y)}{\partial y} = (z(x,y+1) - z(x,y))$ [Forsyth and Ponce 2003; Trucco and Verri 1998]. This gives rise to a large but sparse system of linear equations which can be solved by conjugate gradient or multigrid methods [Press et al. 1992].

The normals are then recomputed from this surface approximation. This step can be viewed as projecting the normal field into the subspace of feasible normal fields.

4.4 BRDF Parameter Optimization

The BRDFs of the fundamental materials, denoted as f in Equation 3, are generally nonlinear functions of their parameters α. We optimize the objective function over all α simultaneously, using the Levenberg-Marquardt nonlinear optimization algorithm. [Press et al. 1992]

To keep the solution space highly constrained, we use the isotropic Ward model [Larson 1992] as our parametric reflectance model.

Using the original notation:

$$\rho_{bd,iso}(\theta_i,\phi_i;\theta_r,\phi_r) = \frac{\rho_d}{\pi} + \frac{\rho_s}{\sqrt{\cos\theta_i \cos\phi_r}} \cdot \frac{\exp[-\tan^2\delta/\alpha^2]}{4\pi\alpha^2} \tag{7}$$

where ρ_d and ρ_s are the diffuse and specular reflectance coefficients, α is a measure of roughness, and δ is the angle between $\mathbf{n}$ and the halfway vector $\mathbf{h} = (\mathbf{V}+\mathbf{L})/||\mathbf{V}+\mathbf{L}||$. This model has a very small number of parameters – seven, if ρ_d and ρ_s are RGB values – and is thus well-suited to our problem.

Also, in order to keep the fundamental materials from diverging too far from observed materials in the scene, we have added a small spring term between each pair of fundamental materials to our objective function. This takes the form

$$Q_s(\alpha) = \varepsilon \sum_{i \neq j} ||\alpha_i - \alpha_j||^2 \tag{8}$$

where i and j loop over all pairs of materials.

4.5 Initialization

For all our datasets, the first step is to initialize the normals using Lambertian photometric stereo [Woodham 1980]. Since this method fails in the presence of specular highlights, the user chooses intensity threshholds to reject shadow and specular highlight pixels from consideration. For particular configurations of lights and shiny objects, some pixels may have fewer than 4 inlier samples with which to estimate a normal, and for these pixels we simply choose an arbitrary plausible normal.

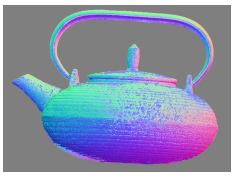

Figure 4: RGB-encoded normals acquired using the Lambertian photometric stereo method, with threshholding to exclude shadows and highlights.

Although the resulting normal map is quite poor (see Figure 4) it suffices as an initial guess.

The next stage of initialization treats large regions as if they were homogenous but non-Lambertian. At this point we have slightly different strategies for different types of objects. For objects with continuous variation of material across their surface, we find that the materials are often similar enough that an initial approximation as a single homogenous material can be successful. But for objects with more widely differing materials, we segment the images into multiple homogenous regions.

4.5.1 Similar Materials

For surfaces with multiple similar and intermixed materials, we initially constrain the solution to use just a single constant material, and apply the optimization algorithm described in Sections 4.2 through 4.4. This produces an approximation of an "average" material over the surface. Then we manually perturb the parameters of that single "average" material in order to generate two or more slightly different materials, and resume the optimization using the normals generated in this initialization step.

For example, the normals and material estimates for the single-material estimate of our teapot example are shown in Figure 5. We note that the final materials shown in Figure 1 are actually quite different from each other, suggesting that a fairly large class of objects may be initialized in this way.

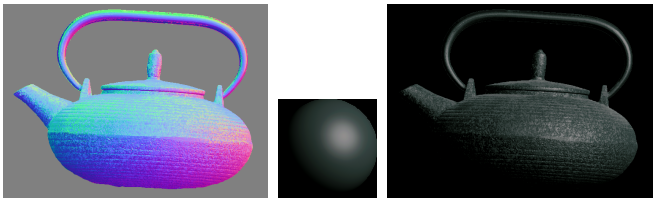

Figure 5: Normals (left), virtual sphere (middle) and reconstruction (right), using only a single material to initialize.

4.5.2 Widely Varying Materials

For widely varying material regions, we first segment the object into regions of constant or nearly-constant material. If the material boundaries are simple enough, the segmentation is performed manually as in the candlestick example shown in Figure 6. If they are more complex, we apply an automatic segmentation algorithm to an estimate of diffuse reflectance, assuming that the diffuse colors are distinct between material regions.

Many approaches to estimating diffuse reflectance and automatic segmentation can be found in the computer vision literature, and we believe that our algorithm is not especially sensitive to the quality of this initial segmentation. However, for the examples used in this paper, we have used the intrinsic images algorithm proposed by Weiss [2001] to obtain our estimate of diffuse reflectance, and an EM optimization for mixtures-of-Gaussians to segment the intrinsic image into separate regions [Yamazaki 1998; Belongie et al. 1998]. To help guide this clustering we also apply cluster constraints, in a similar manner to the weight constraints described in Section 4.2.

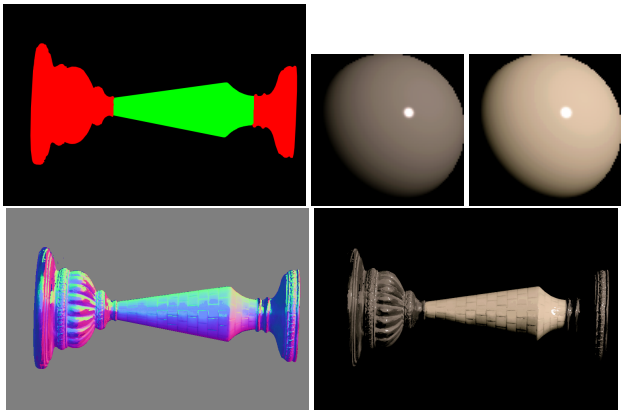

Figure 6: Initial segmentation (upper left), materials (upper right), normals (lower left) and reconstruction (lower right), using two fundamental materials.

We use this segmentation to initialize the material weights, and apply the optimization algorithm described in Sections 4.2 through 4.4 while holding these material weights fixed. After a few iterations of

the alternating optimization, we free the weights to be optimized, or sparser weight constraints as in the example of Figure 3.

Figure 7 shows the intrinsic image, constraints, and initial segmentation for the ornament dataset.

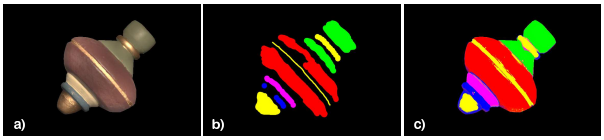

Figure 7: Ornament: a) Intrinsic image, b) constraints c) initial segmentation, using five fundamental materials.

5 Applications and Results

A few of our reconstructions, with different viewpoints and comparisons to new sources, are shown in Figures 8, 9, 10, and 11. Note that, for each object, the algorithm estimates a detailed normal map, accurately segments the surface materials, estimates the reflectance properties of the materials, and faithfully reproduces the input imagery. A few artifacts still occur, however, in regions of the surface that were had highlights in most of the views, such as the frontal portion of the candlestick, and the lower part of the leaf. The presence in highlights in all of the images causes the algorithm to over-esimate the diffuse component. These artifacts do not have a significant impact on rerendering and relighting of these objects.

5.1 Capture Mechanism

To capture our source images, we programmed a Lutron lighting control system and Canon 10D camera with a 400mm zoom lens to automatically capture multiple exposures of multiple lighting samples. (Although our algorithm assumes an orthographic camera and parallel light rays, we have obtained good results using this long focal length, with distances from target to camera and lights of only 5 feet.) Our images are captured at the full resolution of the camera (3072×2048), but most of the examples in this paper were computed at a downsampled resolution of 768×512. A typical capture session with 12 light sources takes about 25 minutes, most of which is simply the time to download the high-resolution images to disk via the camera's USB 1.0 interface.

The multiple exposures of each lighting sample are then combined into high-dynamic range images using the technique of Debevec and Malik [1997].

Since we use multiple fixed light sources, we can capture our calibration images once for an entire capture session. Figure 12 shows several images from the teapot capture session.

5.2 Editing Operations

Direct Manipulation. For our reconstructions we have used the Ward BRDF model, which has a small number of parameters that can be directly manipulated to change one or more of the fundamental materials without modifying the others.

For example, in Figure 13, we have manually edited the BRDF parameters of the green paint of our teapot to appear as gold leaf.

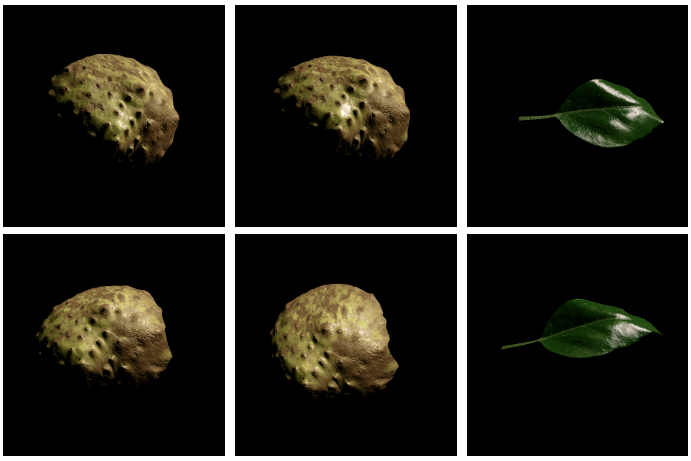

Figure 9: Alternate viewpoints of our reconstructions of a cherimoya (a tropical fruit) and a leaf.

Figure 10: Left column: Models under novel lighting conditions. Right column: Photographs of real object under the same lighting condition.

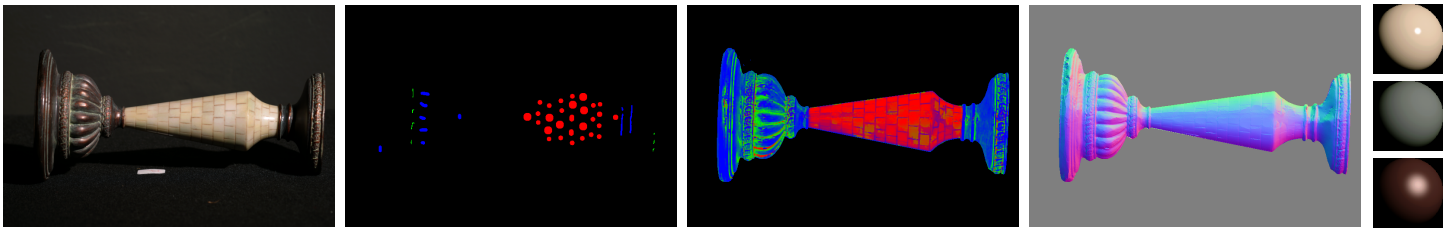

Figure 8: Input, model, and reconstructions for candlestick data. a) Source image (1 of 10). b) Material weight constraints, in false color. c) Recovered material weights, in false color. d) Recovered normal map, RGB-encoded. e) Virtual spheres for fundamental materials.

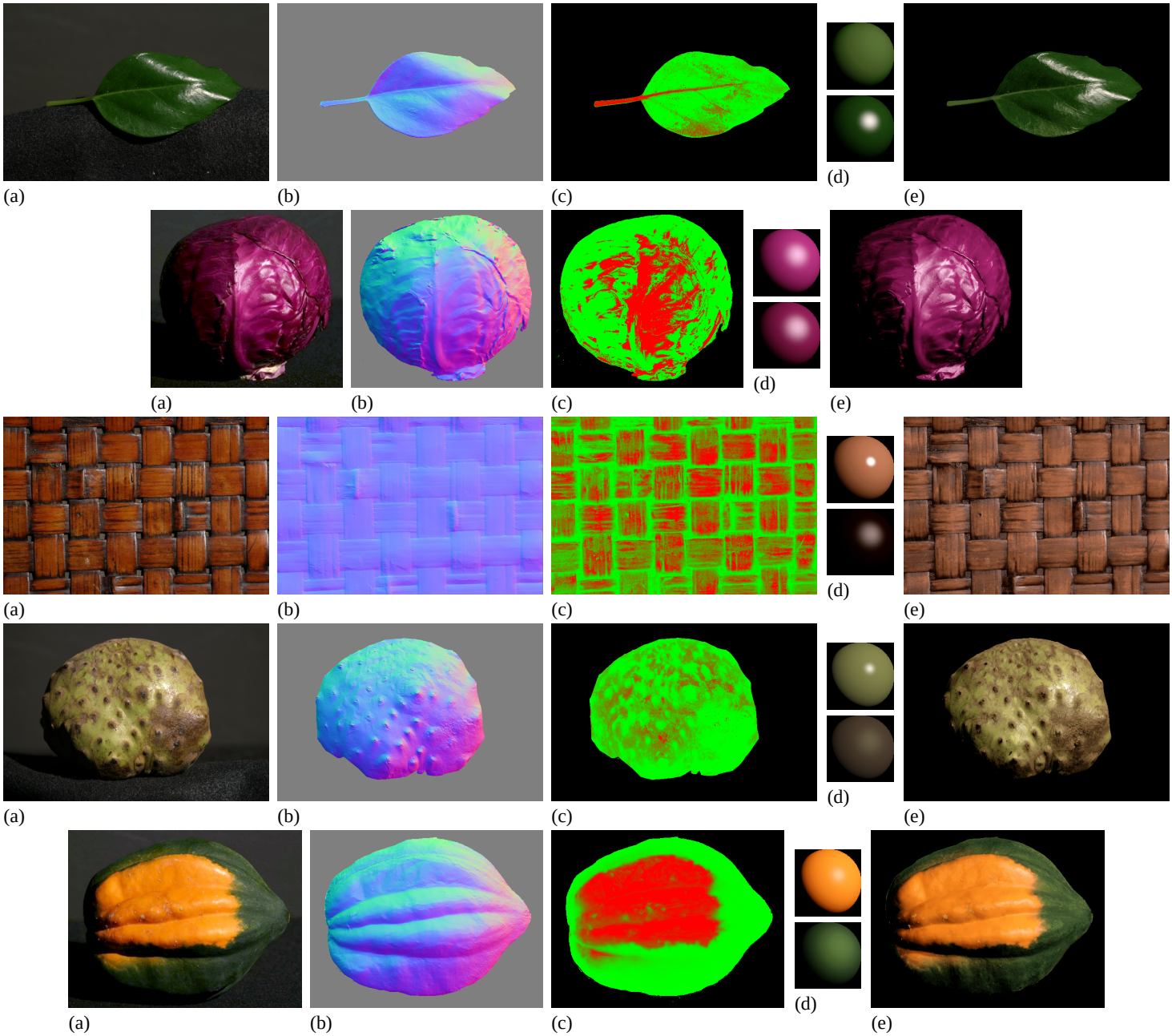

Figure 11: Input, model, and reconstructions for leaf, cabbage, woodweave, cherimoya, squash. a) Source image (1 of 10). b) Recovered normal map, RGB-encoded. c) Recovered material weights, in false color. d) Virtual spheres for fundamental materials. e) Model under original lighting condition.

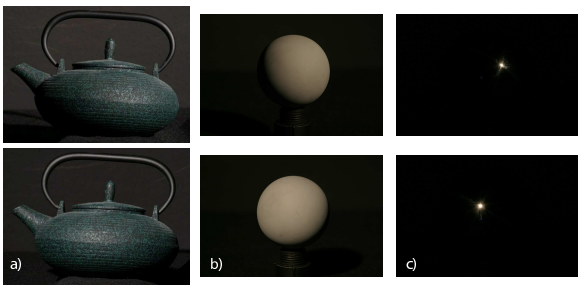

Figure 12: Two of ten target, gray and chrome calibration images.

Figure 13: A pot of gold, created by manually editing the BRDF parameters of the teapot to appear as gold leaf.

Material Transfer. BRDF parameters captured from one object can easily be transferred to another.

Figure 14 illustrates materials captured from a lacquered wood tabletop, applied to the rerendering of our teapot.

Note that in both Figures 13 and 14, the normals and material weights are unchanged from the original, retaining the pattern of paint splotches from the real teapot.

Material/shape extraction and transfer. One application of our method is to extract surface texture and detail for transfer to new objects; we show an example using a portion of the cherimoya in Figure 11. We first cropped a portion of the image containing roughly uniform texture, and applied the following steps.

Our first goal was to remove large-scale variations in shape, while keeping the fine-scale "textural" detail. We applied Gaussian smoothing to the depth map $z(x,y)$ to produce a *base surface* $z'(x,y)$. We additionally define a target surface, in our case a flat plane $z''(x,y) = 0$. Since the source patch is relatively flat, we did not need to reparameterize it, although reparameterization could be used for more highly-curved surfaces. A conventional approach would be to transfer displacement vectors from their local coordinate frames on the base mesh to the corresponding coordinate frames on the target mesh (e.g. [Biermann et al. 2002]); this change-of-basis amounts to a rotation for each displacement vector. We chose an alternative approach of instead transferring surface normals, i.e. applying to each surface normal a change-of-basis from the base surface to the target surface. This gives us a new normal field with global shape variations removed; a new depth map is obtained by surface integration as in Section 4.3 (the target surface is discarded). In this example, we found that our approach preserved more high-frequency detail than transferring displacements. The material weight map is unchanged by this process.

Figure 14: A lacquered wooden teapot, created by applying the BRDF parameters from the wood table top to the shape and weights of the teapot.

This gives us a planar patch of surface texture that may be rendered, or used as source data for texture synthesis. We used texture synthesis to produce a much larger patch from just the normal map (Figure 16f), and then used Image Analogies [Hertzmann et al. 2001] to produce a corresponding weight map from the normal map (Figure 16g). (An alternative approach would be to perform texture synthesis by directly comparing pixel neighborhoods between the source and target surfaces, similar to [Tong et al. 2002; Turk 2001; Wei and Levoy 2001; Ying et al. 2001].)

Finally, the synthesized normals are integrated again to form a depth map which can be used for bump or displacement mapping. In Figure 18, we applied much larger maps synthesized in this fashion to a polygon mesh obtained from a Cybwerware scan, and rendered the resulting model using Pixar's Photorealistic RenderMan with a 70-line surface shader.

6 Discussion and Future Work

We have demonstrated a method which acquires both shape and spatially-varying BRDFs from a set of photographs under varying illumination. Although our shape and material reconstructions are lower fidelity than those attainable using methods that assume one or the other is given and use large numbers of observations, we can nonetheless acquire a wide range of models which can be reused under various lighting and viewing conditions. The spatially varying BRDFs that we acquire enable a host of useful and interesting editing operations. We believe this represents an important step towards acquisition and reconstruction of both shape and material from a single set of photographic data.

In particular, the increasing resolution of consumer-grade digital cameras suggests that very high-resolution shape reconstruction may be attainable at low cost using photometric stereo methods, especially when one considers that we have used only one-sixteenth of the available pixels in our images. To illustrate this point, Figure 15 shows a submillimeter surface reconstruction for one of our data

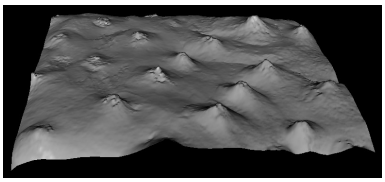

Figure 17: The normal map from Figure 16(c), integrated and viewed from the side.

sets. Although this reconstruction has extraordinary resolution, it has some geometric distortions due primarily to the inaccuracy of the distant light source and orthographic camera assumptions. We would like to remove these assumptions in future work.

Our approach is able to capture shape and BRDFs of reflective objects using a small number of photos and without modification of the objects for scanning. However, our method still requires more user input than we would like. We hypothesize that an objective function which includes a term that prefers sparsity for material weights may reduce the need for manually painted weight constraint maps.

Also, the objective function in Equation 3 is still subject to some overfitting, usually in the case where pixels appear in specular highlight in most image samples. In such cases, these pixels may be assigned a brighter diffuse material instead of the correct specular material. This type of artifact appears as noise in the material maps for the leaf and cabbage datasets of Figure 11. For some inputs, these overfit pixels begin to dominate the fundamental material optimization, skewing towards brighter, washed out colors, as in the woodweave datasets in Figure 11. We suspect that adding a smoothness term for material weight maps to our objective function may reduce this problem, but such an objective function will require a new optimization approach.

Finally, because we employ a local reflectance model, our acquisitions are confounded by the presence of shadows, interreflections and subsurface scattering. Few real-world objects are free of these effects, therefore future work must address techniques to compensate for their appearance.

Despite these limitations, we believe our method will enable more rapid acquisition of computer models suitable for use as background or set dressing objects, from source material at a range of scales which are inaccessible to laser range scanners. We have also demonstrated editing tools made possible by this unique surface representation.

Figure 18: A head model textured using material maps and bump maps synthesized from a cherimoya skin (as shown in Figure 16).

References

AKERS, D., LOSASSO, F., KLINGNER, J., AGRAWALA, M., RICK, J., AND HANRAHAN, P. 2003. Conveying Shape and Features with Image-Based Relighting. In *IEEE Visualization 2003*.

BELONGIE, S., CARSON, C., GREENSPAN, H., AND MALIK, J. 1998. Color- and texture-based image segmentation using EM and its application to content-based image retrieval. In *Proceedings of the Sixth International Conference on Computer Vision*.

BIERMANN, H., MARTIN, I., BERNARDINI, F., AND ZORIN, D. 2002. Cut-and-paste editing of multiresolution surfaces. *ACM Transactions on Graphics 21*, 3 (July), 312–321.

CHUANG, Y.-Y., ZONGKER, D. E., HINDORFF, J., CURLESS, B., SALESIN, D. H., AND SZELISKI, R. 2000. Environment matting extensions: Towards higher accuracy and real-time capture. In *Proceedings of ACM SIGGRAPH 2000*, ACM Press / ACM SIGGRAPH / Addison Wesley Logman, Computer Graphics Proceedings, Annual Conference Series, 121–130. ISBN 1-58113-208-5.

COOK, R. L. 1984. Shade Trees. In *ACM Computer Graphics (Proceedings of SIGGRAPH 84)*, 223–231.

DEBEVEC, P. E., AND MALIK, J. 1997. Recovering High Dynamic Range Radiance Maps from Photographs. In *Proc. SIGGRAPH 97*, 369–378.

DEBEVEC, P., HAWKINS, T., TCHOU, C., DUIKER, H.-P., SAROKIN, W., AND SAGAR, M. 2000. Acquiring the Reflectance Field of a Human Face. In *Proceedings of SIGGRAPH 2000*, ACM Press / ACM SIGGRAPH, New York, K. Akeley, Ed., Computer Graphics Proceedings, Annual Conference Series, ACM, 145–156.

FOLEY, J. D., VAN DAM, A., FEINER, S. K., AND HUGHES, J. F. 1990. *Computer Graphics: Principles and Practice*. Addison-Wesley, Reading, MA.

FORSYTH, D. A., AND PONCE, J. 2003. *Computer Vision: A Modern Approach*. Prentice Hall.

GARDNER, A., TCHOU, C., HAWKINS, T., AND DEBEVEC, P. 2003. Linear Light Source Reflectometry. *ACM Transactions on Graphics 22*, 3, 749–758.

GEORGHIADES, A. S. 2003. Recovering 3-D Shape and Reflectance From a Small Number of Photographs. In *Eurographics Symposium on Rendering: 14th Eurographics Workshop on Rendering*, 230–240.

GORTLER, S. J., GRZESZCZUK, R., SZELISKI, R., AND COHEN, M. F. 1996. The lumigraph. In *Proc. SIGGRAPH 96*, 43–54.

HERTZMANN, A., AND SEITZ, S. M. 2003. Shape and Materials by Example: A Photometric Stereo Approach. In *Proc. Computer Vision and Pattern Recognition*, vol. 1, 533–540.

HERTZMANN, A., JACOBS, C. E., OLIVER, N., CURLESS, B., AND SALESIN, D. H. 2001. Image analogies. In *SIGGRAPH 2001, Computer Graphics Proceedings*, ACM Press / ACM SIGGRAPH, E. Fiume, Ed., 327–340.

JAROSZKIEWICZ, R., AND MCCOOL, M. D. 2003. Fast Extraction of BRDFs and Material Maps from Images. In *Graphics Interface*, to appear.

LARSON, G. J. W. 1992. Measuring and Modeling Anisotropic Reflection. In *Computer Graphics (Proceedings of SIGGRAPH 92)*, vol. 26, 265–272.

LENSCH, H., GOESELE, M., KAUTZ, J., HEIDRICH, W., AND SEIDEL, H. 2001. Image-Based Reconstruction of Spatially Varying Materials. In *Proceedings of the 12th Eurographics Workshop on Rendering*, Springer, Eurographics, 104–115.

LENSCH, H. P. A., KAUTZ, J., GOESELE, M., HEIDRICH, W., AND SEIDEL, H.-P. 2003. Image-based reconstruction of spatial appearance and geometric detail. *ACM Trans. Graph. 22*, 2, 234–257.

LEVOY, M., AND HANRAHAN, P. 1996. Light field rendering. In *Proc. SIGGRAPH 96*.

MALZBENDER, T., GELB, D., AND WOLTERS, H. 2001. Polynomial texture maps. In *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, ACM Press, 519–528.

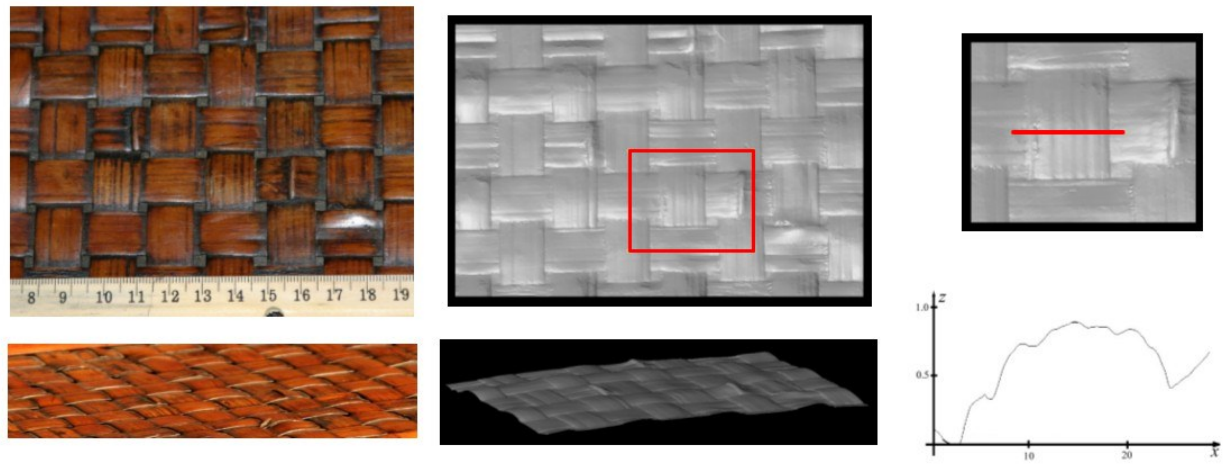

Figure 15: Top left: wooden tabletop with ruler for scale (in centimeters). Top center: plastic shaded reconstruction. Bottom left: oblique view photograph. Bottom center: oblique view reconstruction. Top right: closeup of inset. Bottom right: depth profile of red line in inset, with z-axis scaled up for clarity (axes units are millimeters).

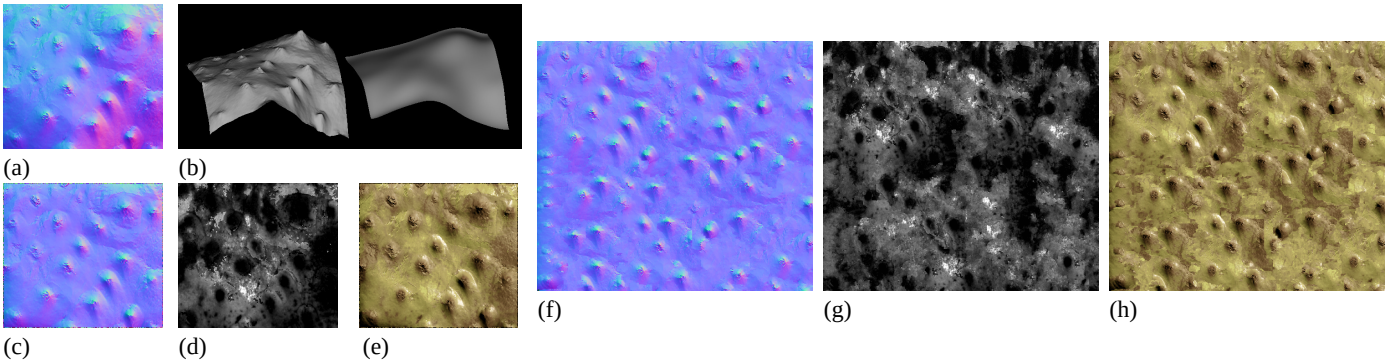

Figure 16: Intermediate steps in the construction of Figure 18. (a) Normals extracted using our methods are integrated to construct (b) depths, and smoothed. The smoothed normals are used to create a "flattened" normal map (c) which retains high-frequency texture. This is combined with (d) the recovered weight map to render (e) a flattened swatch of cherimoya skin. (f) Larger normal maps and (g) weight maps can be constructed using texture synthesis to create (h) novel renderings.

MARSCHNER, S. R., WESTIN, S. H., LAFORTUNE, E. P. F., TORRANCE, K. E., AND GREENBERG, D. P. 1999. Image-based BRDF Measurement Including Human Skin. In *Proceedings of 10th Eurographics Workshop on Rendering*, 139–152.

MATUSIK, W., PFISTER, H., ZIEGLER, R., NGAN, A., AND MCMILLAN, L. 2002. Acquisition and Rendering of Transparent and Refractive Objects. In *Eurographics Workshop on Rendering (EGRW)*, Eurographics, 267–278.

MATUSIK, W., PFISTER, H., BRAND, M., AND MCMILLAN, L. 2003. Efficient Isotropic BRDF Measurement. In *Eurographics Symposium on Rendering: 14th Eurographics Workshop on Rendering*, Eurographics, 241–248.

NAYAR, S. K., IKEUCHI, K., AND KANADE, T. 1990. Determining shape and reflectance of hybrid surfaces by photometric sampling. *IEEE Trans. on Robotics and Automation 6*, 4, 418–431.

PRESS, W. H., TEUKOLSKY, S. A., VETTERLING, W. T., AND FLANNERY, B. P. 1992. *Numerical Recipes in C: The Art of Scientific Computing (2nd ed.)*. Cambridge University Press.

RAMAMOORTHI, R., AND HANRAHAN, P. 2001. A Signal-Processing Framework for Inverse Rendering. In *SIGGRAPH 2001, Computer Graphics Proceedings*, ACM Press / ACM SIGGRAPH, E. Fiume, Ed., 117–128.

SATO, Y., WHEELER, M. D., AND IKEUCHI, K. 1997. Object shape and reflectance modeling from observation. In *Proc. SIGGRAPH 97*, 379–387.

SILVER, W. M. 1980. *Determining Shape and Reflectance Using Multiple Images*. Master's thesis, MIT, Cambridge, MA.

TAGARE, H., AND DEFIGUEIREDO, R. 1991. A theory of photometric stereo for a class of diffuse non-lambertian surfaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (February), 133–152.

TONG, X., ZHANG, J., LIU, L., WANG, X., GUO, B., AND SHUM, H.-Y. 2002. Synthesis of bidirectional texture functions on arbitrary surfaces. *ACM Transactions on Graphics 21*, 3 (July), 665–672.

TRUCCO, E., AND VERRI, A. 1998. *Introductory Techniques for 3-D Computer Vision*. Prentice Hall.

TU, P., AND MENDONÇA, P. R. S. 2003. Surface Reconstruction via Helmholtz Reciprocity with a Single Image Pair. In *Proc. CVPR*, vol. 1, 541–547.

TURK, G. 2001. Texture synthesis on surfaces. In *Proceedings of ACM SIGGRAPH 2001*, Computer Graphics Proceedings, Annual Conference Series, 347–354.

WEI, L.-Y., AND LEVOY, M. 2001. Texture synthesis over arbitrary manifold surfaces. In *Proceedings of ACM SIGGRAPH 2001*, Computer Graphics Proceedings, Annual Conference Series, 355–360.

WEISS, Y. 2001. Deriving intrinsic images from image sequences. In *ICCV 2001*, 68–75.

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

WOODHAM, R. J. 1980. Photometric method for determining surface orientation from multiple images. *Optical Engineering 19*, 1, 139–144.

YAMAZAKI, T. 1998. Introduction of EM algorithm into color image segmentation. In *Proceedings of ICIPS'98*, 368–371.

YING, L., HERTZMANN, A., BIERMANN, H., AND ZORIN, D. 2001. Texture and shape synthesis on surfaces. In *Rendering Techniques 2001: 12th Eurographics Workshop on Rendering*, 301–312.

YU, Y., DEBEVEC, P., MALIK, J., AND HAWKINS, T. 1999. Inverse Global Illumination: Recovering Reflectance Models of Real Scenes From Photographs. In *Siggraph99, Annual Conference Series*, Addison Wesley Longman, Los Angeles, A. Rockwood, Ed., ACM, 215–224.

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 (September/October), 215–227.