Submitted to ACM Transactions on Graphics, June 2003

Re-rendering from a Sparse Set of Images

Ko Nishino
Columbia University
Katsushi Ikeuchi
The University of Tokyo
and
Zhengyou Zhang
Microsoft Research

We present a method to accomplish photorealistic rendering of real-world objects from their sparsely sampled appearance variation. Using a 3D model and a small set of images of an object, we recover all the necessary photometric information for subsequent rendering from arbitrary viewpoints and under novel lighting conditions. We first extract the diffuse reflection component from the input images as a texture map, and then use the residual images to simultaneously recover the specular reflection parameter and the illumination distribution. The simultaneous estimation of the specular reflection parameter and the illumination distribution is achieved by formulating the specular reflection mechanism as a 2D convolution on the surface of a hemisphere. We then run an iterative algorithm to deconvolve it. Rendering from novel viewpoints and under novel illumination distributions can be accomplished using the estimated three components. Unlike previous approaches, we require less input images and we do not assume anything to be known about the three photometric attributes, namely the diffuse and specular reflection parameters and the lighting condition.

Categories and Subject Descriptors: I.3.7 [**Computer Graphics**]: Three-Dimensional Graphics and Realism–*Color,Shading,Shadowing and Texture*

General Terms: Algorithms, Measurement

Additional Key Words and Phrases: Inverse rendering, BRDF estimation, illumination recovery

1. INTRODUCTION

Rendering real-world objects from the observation of their appearance variation, which we will refer to as *re-rendering*, has been a major research topic in both computer graphics and computer vision for a while. In order to achieve photorealism, one can sample the appearance variation of real objects under various viewpoints and lighting conditions and consider them as samples of the light field. Then, those sampled points can be interpolated to obtain synthetic images from novel viewpoints or under novel lighting conditions [Levoy and Hanrahan 1996; Gortler et al. 1996; Shum and L-W.He 1999]. Furthermore, by knowing the geometry of the object (say, by using a range scanner), we can consider the measured appearance variation as samples in the surface light field, which yields better interpolation and compression schemes to render novel images [Nishino et al. 1999; Chen et al. 2002; Wood et al. 2000]. As both of these approaches are based on the principle of in-

...

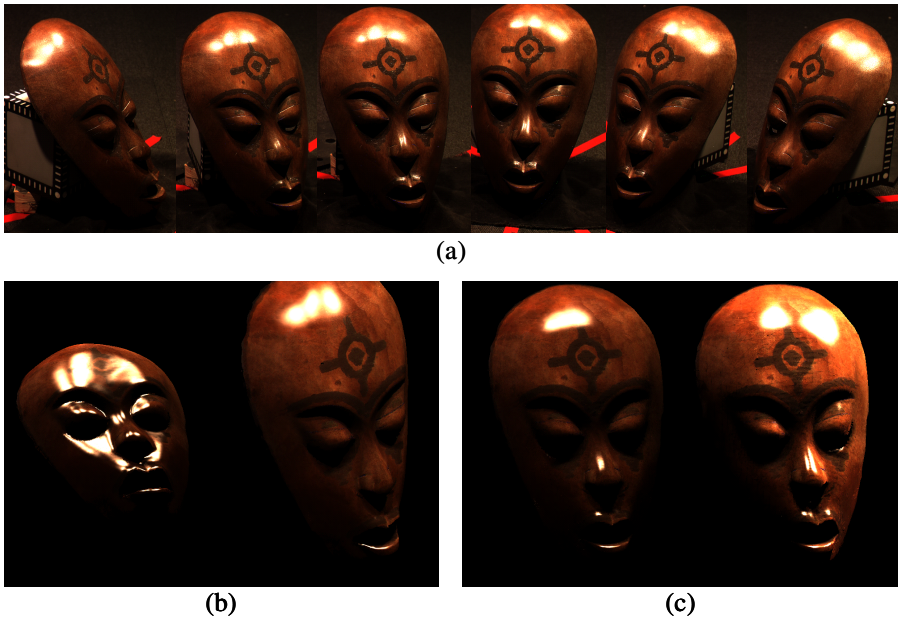

Fig. 1. From a small set of images (a) and a pre-acquired geometric model of the target object, we are able to accomplish rendering from arbitrary novel view points (b) and under arbitrary novel lighting conditions (c).

terpolation, essentially they require a large number of sample points: they require the user to take hundreds or thousands of images. Especially when it comes to re-rendering objects that have high frequency reflection properties, such as specular reflection, the necessary number of images becomes enormous.

An alternative approach for re-rendering is to explicitly recover the necessary set of parameters, which is commonly referred to as *inverse rendering*. In addition to the geometry of the object, one needs to know the material property of the object and the illumination distribution to render novel images. The material property can be described with the Bidirectional Reflectance Distribution Function (BRDF), which can be further separated into diffuse reflection component and specular reflection component. Assuming the geometry of the object is known, many approaches have been proposed to recover the parameters of one or two out of the three components: diffuse and specular reflection and illumination distribution. If the illumination distribution is fully provided by careful calibration or is perfectly controlled, both components of the BRDF can be tabulated [Marschner et al. 2000; Lu et al. 1998; Dana et al. 1999] or the parameters of specific reflection models such as that of [Ward 1992; Torrance and Sparrow 1967], which approximate the components of the BRDF in a simple analytic form, can be recovered through functional fitting [Marschner et al. 2000; Lu et al. 1998; Dana et al. 1999; Sato et al. 1997; Lensch et al. 2001; Nayar et al. 1990; Debevec et al. 2000; Boivin and Gagalowicz 2001; Ikeuchi and Sato 1991]. If the BRDF does not contain any specular components and the object is perfectly diffuse, the illumination distribution can

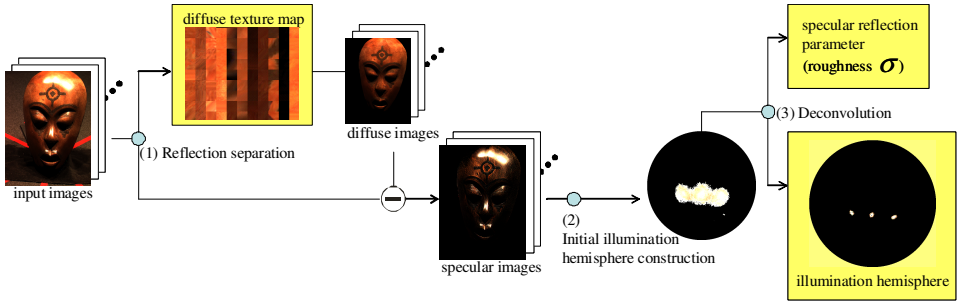

Fig. 2. The pipeline of our method. We first (1) separate the reflection components of the input images and store the diffuse reflection component as one texture map, which we will refer to as the *diffuse texture map*. Next, by using the diffuse texture map, we extract the specular images by simple subtraction, and we then (2) derive a rough estimation of the illumination distribution as a set of point light sources located on a surface of a hemisphere, which we will call the *illumination hemisphere*. Finally, using this initial estimate of the illumination hemisphere, we (3) factor out the specular reflection parameter and the true illumination distribution through iterative deconvolution on the surface of the illumination hemisphere. The specular reflection parameter is the roughness parameter of the Torrance-Sparrow reflection model.

be estimated as a set of basis functions on a hemisphere exploiting the linearity of illumination [Marschner and Greenberg 1997; Sato et al. 1999].

Recently, [Ramamoorthi and Hanrahan 2001] have proposed a signal processing framework for inverse rendering. They formulate the reflection mechanism as a convolution in the angular space and accomplish inverse rendering through a spherical harmonic representation. They provide solutions for two scenarios, (a) where the spatially varying diffuse reflection (texture) and the lighting location are known but the spatially homogeneous specular reflection parameter and the radiance of lighting are unknown, and (b) where the lighting is known but the texture and the homogeneous specular reflection parameter are unknown. In all the approaches mentioned above, the problem of recovering all three sets of unknowns from images have not been tackled.

Ultimately, we would like to let users use their digital cameras to "casually" take several snapshots of an object and turn those images into an efficient representation of the object's appearance which enables them not only to view it from any direction but also to place it under new lighting conditions. We would like to lessen the burden in capturing the input images; we would like to reduce the necessary number of input images; and more importantly, we would *not* like to measure the material property of the object surface and the lighting distribution of the environment. In order to meet this goal, we present a method to recover the necessary parameters to accomplish subsequent rendering of real-world objects from a small set of images and a pre-acquired geometric model of the object.

Figure 2 depicts the flow of our method. We extract the necessary information for re-rendering as the *diffuse texture map*, the specular reflection parameter (roughness

parameter of the Torrance-Sparrow reflection model [Torrance and Sparrow 1967]) and the *illumination hemisphere*. Re-rendering from novel viewpoints and under novel illumination conditions can be accomplished using these three final estimates depicted by the yellow boxes in Figure 2.

The key insights are that (a) we do not try to explicitly estimate the diffuse reflection parameters, since a lighting dependent texture map is enough for subsequent re-rendering, and (b) we represent the illumination distribution as a set of point light sources on a hemisphere which enables us to formulate the specular reflection mechanism as a 2D convolution and successfully deconvolve it on its surface. We are able to derive a very compact representation to accomplish view-dependent rendering and relighting of the object and more importantly we recover all three of the necessary components from a small set of images, typically about ten images, making the whole system flexible and handy.

In the remainder of the paper, we first clarify the assumptions we make in Section 2. We explain how we separate the reflection components in Section 3. In Section 4, we will formulate the specular reflection mechanism as a 2D convolution on a surface of a hemisphere and solve a 2D deconvolution problem to estimate the illumination distribution and the specular reflection parameter. In Section 5 we present our results. We discuss limitations of our method in Section 6 and conclude in Section 7.

2. ASSUMPTIONS

In our framework, we make the following assumptions, while trying not to lose the generality of the scenario that we are considering.

Static high frequency illumination environment, static object and moving camera
We consider a scenario where the user takes several snapshots of an object while moving around the object. The object and the light sources do not move during the image capturing period; only the camera moves. To enable stable recovery of both the reflection parameters and the illumination distribution, we capture the input images under high frequency lighting [Ramamoorthi and Hanrahan 2001]: under point light sources. Also we assume all the light sources have the same color.

Pre-acquired geometry of the target object
We assume we know the precise geometry of the target object. By using a laser range scanner, we acquire several range images of the target object and go through the geometric modeling pipeline: 3D-3D registration [Turk and Levoy 1994], integration [Curless and Levoy 1996] and simplification [Garland and Heckbert 1997]. The final result is a 3D mesh model.

Known intrinsic camera parameters
The camera that is used to capture the input images can be pre-calibrated easily, using techniques like [Zhang 1999].

Known extrinsic camera parameters
The motion of the camera can be pre-estimated by applying low-level computer vision techniques: structure from motion, bundle adjustment, etc. In the experiments we conduct, we simultaneously acquire a range image using a light stripe

range finder [Sato and Inokuchi 1987] as well as capturing color images for each viewpoint. We then use those range images to estimate the viewpoint via 3D-3D registration against the mesh model we mentioned earlier.

Homogeneous specular reflectance property

We assume the specular reflectance property of the object surface is the same for all points on it. Specifically, we assume $(\mathbf{K}_S, \sigma)$ and the Fresnel reflectance coefficient in the Torrance-Sparrow reflection model that we describe in Section 3.1 are the same for the object surface.

3. REFLECTION SEPARATION

We begin by separating the reflection components in the input images (Figure 2 (1)): the diffuse reflection and the specular reflection.

3.1 Reflection Model

We consider real world objects whose reflection mechanism at their surfaces can be approximated by a dichromatic reflection model [Shafer 1985; Klinker et al. 1990], where the reflected light at a surface point is explained as a linear combination of two reflection components: diffuse reflection and specular reflection. As the image irradiance (pixel value) can be considered to be proportional to scene radiance [Horn 1986], we will forget about the constant multiplier that relates scene radiance to image irradiance; instead, we will consider that pixel values in the images are directly associated to scene radiance. Under this assumption, the dichromatic reflection model tells us:

$$\mathbf{I} = \mathbf{I}_D + \mathbf{I}_S \tag{1}$$

where $\mathbf{I}$ is the pixel value corresponding to a particular object surface point, and $\mathbf{I}_D$ and $\mathbf{I}_S$ denote the diffuse and specular reflection radiance, respectively. Bold characters denote three-dimensional color vectors, such that $\mathbf{I} = [I_R \; I_G \; I_B]^T$.

Diffuse reflection can be explained as the result of the light that penetrated into the object medium and radiated back to the air after internal scattering due to the small particles in the object surface layer. Because of this internal scattering, diffuse reflection can be approximated with the Lambertian model [Lambert 1760]. If we consider a local spherical coordinate system with its origin set to the object surface point in interest, the radiance of diffuse reflection can be described as:

$$\mathbf{I}_D = \max[0, \mathbf{K}_D \int_\Omega L_i(\theta_i, \phi_i) \cos\theta_i d\omega_i] \tag{2}$$

where $\mathbf{K}_D$ is a three band color vector which is determined by both the light source color and the surface color. Also, θ_i and ϕ_i denote the altitude and azimuth coordinate of the light source with radiance L_i, respectively, and ω_i denotes the solid angle of the light source L_i. Although recent studies on the diffuse reflection mechanism have revealed that, when the object surface has a high macroscopic roughness, the diffuse reflection becomes view-dependent [Nayar and Oren 1995], we assume that the objects we handle have a diffuse reflection that can be approximated with a Lambertian reflection model.

Specular reflection is the light directly reflected at the interface between the air and the object surface. We use the Torrance-Sparrow reflection model [Torrance and Sparrow 1967] which is well known to describe the physical mechanism of specular reflection well with a simple numerical expression [Nayar et al. 1991]. The Torrance-Sparrow reflection model expresses the "lobe" of specular reflection as light reflected at object surface with microscopic roughness; small V-shaped grooves which are lined with flat mirrors called micro-facets. The distribution of the orientations of micro-facets is approximated with a Gaussian distribution as follows:

$$\mathbf{I}_S = \int_\Omega \frac{\mathbf{K}_S F G}{\cos\theta_r} L_i(\theta_i, \phi_i) \exp[-\frac{\alpha^2}{2\sigma^2}] d\omega_i \tag{3}$$

where $\mathbf{K}_S$ is the color vector of the reflection which accounts for the normalization factor of the exponential function and the reflectivity of the surface, F is the Fresnel reflectance coefficient which depends on the incident angle of the light and the refraction index of the object, G is the geometrical attenuation factor, θ_r is the angle between the viewing direction and the surface normal, α is the angle between the surface normal and the bisector of the viewing direction and the light source direction, and σ represents the surface roughness.

3.2 Diffuse Texture Map

As we vary the viewpoint with fixed object position and lighting, (θ_i, ϕ_i) in Equation (2) does not vary in the input image sequence. Therefore, it can be easily observed that the diffuse reflection component at each surface point has a constant value for all images in the input image sequence and that the pixel value is equal to or larger than the diffuse reflection color. Hence, we are able to represent the diffuse reflection with a single RGB color vector for each surface point.

To extract this diffuse reflection vector for each object surface point, we examine the scene radiance (image irradiance) variation of each surface point throughout the image sequence, and use the pixel value with minimum magnitude (norm of the color vector) as the diffuse reflection color. As we have only a few images, and because the object surface point may be occluded in some of those images, taking the minimum brightness pixel value is the best strategy we can employ to obtain the constant diffuse reflection vector. If we do have more images as the input, more sophisticated strategies may work better, for instance, taking the median value like Wood et al. [Wood et al. 2000] or fitting a Gaussian distribution to the histogram of pixel values.

Instead of keeping these diffuse reflection color vectors for each object surface point, we extract them in a triangular patch-based manner. By considering grids in each triangular patch of the geometric mesh model, and projecting those grids onto each input image and extracting the minimum brightness pixel value for each grid point, we are able to construct a texture map that represents the constant diffuse reflection. We will refer to this texture map as the *diffuse texture map*. Note this diffuse texture map captures the diffuse reflection under the lighting condition of the input images and is different from an albedo map. We do not try to obtain an lighting-independent texture map, not only because it cannot be done without knowing the illumination distribution, but also because we can relight the diffuse

tb

(a) (b) (c)

Fig. 3. Separation of reflection components. (a) original input image, (b) diffuse reflection image (diffuse-texture-mapped image), (c) specular reflection image.

reflection component without accomplishing this factorization once we recover the illumination distribution as described latter.

3.3 Illumination Hemisphere

By texture mapping the diffuse texture map for each viewpoint in the input image sequence (Figure 3(b)) and by subtracting those diffuse-texture-mapped images from each input image (Figure 3(a)) one by one, we obtain a set of residual images (Figure 3(c)). These residual images mainly consist of scene radiance resulting from specular reflection at the object surface and some interreflection and noise inherited from the diffuse texture map extraction procedure. In our work, we ignore interreflection and simply consider it as noise. We will refer to these residual images as specular images.

As described in Section 3.1, the specular reflection approximated with the Torrance-Sparrow reflection model has an intensity peak slightly off from the perfect mirror direction: the direction where the surface normal becomes the bisector of the incident light vector and viewing direction vector. To obtain a rough approximation of the illumination distribution to be used as an initial estimate for subsequent finer estimation (Figure 2 (2)), we shoot back each pixel value in each specular image in the perfect mirror direction and map those values to a hemisphere covering the 3D object model (Figure 4). The radius of the hemisphere will be pre-determined (We will discuss on this later). We call this representation of the illumination distribution the *illumination hemisphere*[1]. Although we should use a sphere to account for all the possible illumination directions in general, we use this hemispherical representation for simplicity in this paper.

The illumination hemisphere generated from each specular image covers only partial regions of the true illumination environment. This is because all light sources will not necessarily contribute to highlights in each image. Hence, we need to com-

[1]In the figures, the illumination hemisphere is depicted with its projection to its base.

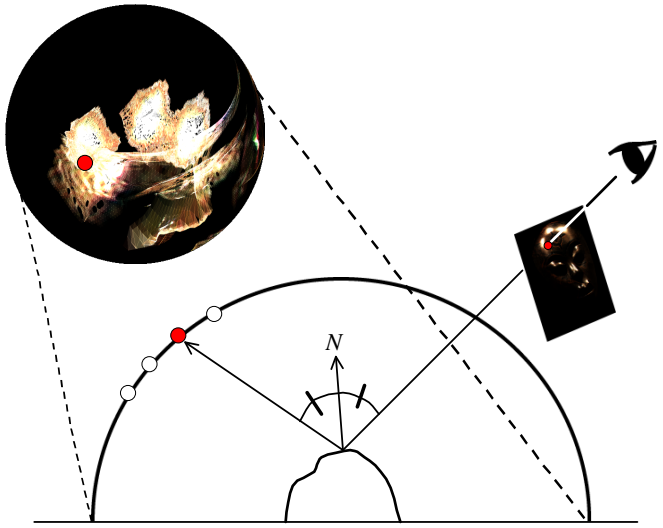

Fig. 4. We shoot back each specular pixel with regards to its perfect mirror direction to construct each partial illumination hemisphere.

bine these partial illumination hemispheres to make a full illumination hemisphere that approximates the real lighting environment. To deal with noise, we make a separate mask hemisphere that counts how many times each point on the illumination hemisphere was taken into account while making each partial illumination hemisphere. We then adopt hemisphere points that have counts close to the total number of images. Since some light sources may be occluded in some partial illumination hemispheres, we set the threshold less than the total number of images. Only those points that pass this check will be mapped onto the final illumination hemisphere; in this case, we take the mean of the intensity values from the partial illumination hemispheres as the intensity value. This means of consistency checking also reduces errors in estimating the illumination, e.g., those introduced by 3D-3D misalignment for viewpoint estimation. These errors are not view-dependent and would not stay in a particular region on the illumination hemisphere. Interreflection will also be faded out, since it can be considered to be the reflected light of moving light sources.

4. ILLUMINATION AND SPECULAR REFLECTION PARAMETER ESTIMATION

Next, using the combined illumination hemisphere and the Torrance-Sparrow reflection model, we will decouple the surface specular reflectance property from the illumination distribution in the specular images (Figure 2 (3)).

4.1 Problem Formulation

First we assume that the object has a specular reflection property that obeys the Torrance-Sparrow reflection model, except that we assume the refraction index is constant (1.5 for the later experiments) and use that value to compute Fresnel reflectance coefficient F. As we represent the illumination distribution with the

illumination hemisphere described in Section 3.3, and each point light source's position on the nodes of the illumination hemisphere can be parameterized with its altitude θ_l and azimuth angle ϕ_l, the pixel value of surface point v will be,

$$\mathbf{I}_S(v) = \frac{\mathbf{K}_{S,v}}{cos\theta_r} \sum_{l}^{N_L} \omega_l F_{l,v} G_{l,v} L_l(\theta_l, \phi_l) \exp[-\frac{\alpha^2}{2\sigma_v^2}] \tag{4}$$

where L_l stands for the radiance of each point light source. Also, θ_r and α are still the angle between the viewing direction and the surface normal and the angle between the surface normal and the bisector of the viewing direction and the light source direction, respectively, however computed in the global coordinate system of the illumination hemisphere. ω_l is the solid angle for each point light source and becomes equally $\frac{2\pi}{N_L}$ where N_L stands for the number of nodes in the geodesic hemisphere.

The color vector direction of the specular reflection is the same as that of the light source. Thus, if we assume that all light sources in the scene have the same color, we can concentrate on the relationship between the radiance of the light sources L_l and the image irradiance $I_S(v)$. Therefore, we use the average color, $\bar{\mathbf{L}}$, of the initial illumination hemisphere as the color of the point light sources. Also, we assume that the target object has a homogeneous specular reflection property, so that σ_v can be represented with one value σ. Similarly $K_{S,v}$ will be represented with one value K_S. With these assumptions, we can rewrite Equation (4) as,

$$\mathbf{I}_S(v) = I_S(v)\bar{\mathbf{L}} \tag{5}$$

$$I_S(v) = \frac{2\pi}{N_L} \frac{K_S}{cos\theta_r} \sum_{l}^{N_L} F_{l,v} G_{l,v} L_l \exp[-\frac{\alpha^2}{2\sigma^2}]. \tag{6}$$

Note that all parameters in Equation (6) are scalar values, and L_l is now the magnitude of the color vector of each point light source placed on the illumination hemisphere.

Now, we will write the intersection of the line connecting the surface point v and the viewpoint E with the illumination hemisphere as (θ_E, ϕ_E), and the intersection of the vector stemming from the surface point v pointing to the perfect mirror direction with respect to the surface normal at v with the illumination hemisphere as $(\theta_{E,v}, \phi_{E,v})$ (Figure 5). Then, we can rewrite Equation (6) as follows.

$$I_S(v) = \frac{2\pi}{N_L} \frac{K_S}{cos\theta_r} \sum_{l}^{N_L} F_{l,v} G_{l,v} L_l g_{\frac{1}{2\sigma^2}}(\frac{\|(\theta_l, \phi_l) - (\theta_{E,v}, \phi_{E,v})\|}{2}) \tag{7}$$

where g_γ is a Gaussian function:

$$g_\gamma(\phi) = \exp[-\gamma\phi^2] \tag{8}$$

From Equation (7), we can clearly see that given a viewpoint E and a surface point v with its known surface normal, a unique point on the surface of the illumination hemisphere $(\theta_{E,v}, \phi_{E,v})$ is determined, and a skewed[2] Gaussian weighted

[2]Because of the foreshortening factor etc. in front of the summation. Note, however, this deformation factor is known (can be computed).

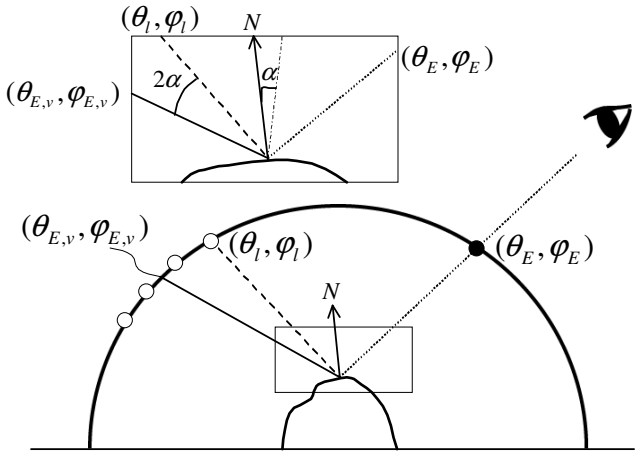

Fig. 5. The angles in Equation (7).

integration on the point light sources surrounding that point is accomplished to determine the intensity of that surface point v. Note Equation (7) is an approximation since the relationship between angles in the rectangle of Figure 5 does not hold when (θ_l, ϕ_l) is not coplanar to (θ_E, ϕ_E) and $(\theta_{E,v}, \phi_{E,v})$. Hence, in reality, the Gaussian applied to each point on the surface of the illumination hemisphere changes its shape from point to point. However, note that we have all the information to compute the shape of this filter. Hence, it is worthwhile illustrating the specular reflection as a convolution on the illumination hemisphere surface which leads us to the deconvolution algorithm described in the next section. Since K_s works as a scaling factor to the result of the integral, we are not able to recover its actual value. Nevertheless, we can relight the object specifying the new lighting distribution with intensity values relative to the original. Therefore, we simply set K_s to 1.0. Note, as a result, everything that deforms the pure Gaussian shape is known, and hence recovering the specular reflection parameter (σ) and the lighting distribution ($L_l \ : \ l = 0, ..., N_L$) simultaneously becomes a 2D blind deconvolution problem[3]. The advantage of treating specular reflection as a 2D convolution on the surface of a hemisphere instead of convolution in angular space [Ramamoorthi and Hanrahan 2001] is that we can apply powerful blind image deconvolution algorithms as we will see in the next section.

4.2 Alternating Minimization

If we view the surface of the illumination hemisphere as an image of point light sources, we can consider the problem of estimating the lighting distribution and the specular reflection parameter as an blind image deconvolution problem. To make the following argument simple, we will describe the surface of the illumination hemi-

[3]This deconvolution is blind in a sense that we have to estimate the variance of the Gaussian kernel σ as well as the true illumination hemisphere.

sphere as $u(x, y)$, the convolution filter as $h(x, y)$ and the specular image $z(x, y)$[4]. Then, the task of estimating the reflectance parameters and the true illumination distribution becomes finding u, h that best describes the following equation,

$$z = h * u + \eta \tag{9}$$

where η is the noise mainly added while observing through an imaging system. This problem is typically referred to as a blind image deconvolution problem, restoring both the image u and the blurring function h while given only the observed image z and probably some statistics of the noise η. The problem is well studied in the image processing community. The difficulty of solving this blind deconvolution is that it is ill-posed with respect to the image (the illumination hemisphere) and the blurring function (the Gaussian of the Torrance-Sparrow reflection model).

Recently, [You and Kaveh 1996] proposed a joint regularization technique to regularize both u and h to solve this problem:

$$\min_{u,h} f(u, h) = \min_{u,h} ||h * u - z||^2_{L^2(\Omega)} + \rho_1 R(u) + \rho_2 R(h) \tag{10}$$

where Ω is the domain of u and h, R is the regularization function.

The objective function $f(u, h)$ in Equation (10) as a two-variables function is not convex and hence can have multiple solutions. [You and Kaveh 1996] observed that, for a fixed h (respectively u), $f(\ \cdot\ , h)$ (respectively $f(u,\ \cdot\ )$) is a convex function of u (respectively h) and they proposed the Alternating Minimization (AM) algorithm which can be run in the following steps.
Given u^0:
1. Find minimizer h^k of $f(u^{k-1}, h)$ by solving $\nabla_h f(u^{k-1}, h) = 0$
2. Find minimizer u^k of $f(u, h^k)$ by solving $\nabla_u f(u, h^{k-1}) = 0$
3. If not converged, go to 1.

As for the regularization function, we adopt the *Total Variation* (TV) norm [Chan and Wong 1998] which is known to provide faster convergence compared to other regularization terms such as the H^1 and preserves edges better. Convergence of AM algorithm with TV regularization is proved [Chan and Wong 1998]. Furthermore, it is not necessary to explicitly impose the numerical conditions that u and h has to satisfy, e.g., $u \geq 0$. These conditions can be embedded in the minimization steps of the AM algorithm, for instance, by simply setting $u(x, y) = 0$ when its estimate is negative. Hence, by enforcing negative light sources to have zero values at each step of updating light source intensities in the iteration, we can safely constrain the point light source on the illumination hemisphere to have positive values.

When setting up a joint regularization minimization formula to solve the blind deconvolution problem of our case, we do not have to impose a regularization term for the blurring function, since we already know that the blurring function is a one dimensional Gaussian and because it is smooth with respect to its parameter. Thus, the minimization problem to estimate the reflectance parameter σ and the illumination distribution represented as a discretized hemisphere can be formulated

[4]The convolved result gets mapped back to the 3D surface points on the object and then gets projected on to the image plane

as follows.

$$\min_{L,\sigma} f(L,\sigma) \equiv \min_{L,\sigma} \sum_{k}^{N_K} \sum_{s,t}^{N_S,N_T} \sum_{l}^{N_L} ||I_S(s,t,k) - I(s,t,k)||_{L^2}^2 + \rho \sum_{l}^{N_L} |\nabla L(\theta_l, \phi_l)| \quad (11)$$

where (s,t) are the image coordinates and k is the image number. To solve Equation (11), we apply the AM algorithm, starting with the initial illumination hemisphere (u^0)which we derived in Section 3.3, with two special considerations described in the following sections.

4.2.1 *M-estimator for Robustness.* As we have obtained the observed values of specular reflection, $I(s,t,k)$, by subtracting the diffuse-texture-mapped images (estimated diffuse images) from the original input images, we have inherited errors generated in the diffuse reflection estimation. Thus, when estimating both the illumination distribution and the reflectance parameters from the specular images, we will have to ensure that the estimation procedure is robust against noise. For this purpose, we adopt the robust estimation technique based on M-estimator [Press et al. 1992; Gill et al. 1981]. Consequently, we minimize Equation (11) for each step in the AM algorithm through conjugate gradient with Lorentzian function [Press et al. 1992; Gill et al. 1981] as the M-estimator and golden section search for line minimization.

4.2.2 *Initial Estimate.* As observed in [Chan and Wong 1998] with numerical results, estimating blurring functions without edges, e.g., Gaussian, with the AM algorithm happens to converge slowly as compared with estimating point spread functions with edges, e.g., out-of-focus blur. One way to speed up this convergence rate is to give a good initial estimate of the blurring function. In our case, we would like to utilize a good initial estimation of σ, instead of setting σ^0 very small to simulate a delta function. Then, the question is what is σ_{true}? Fortunately, we know that most shiny objects that can be modeled with the micro-facet BRDF reflection model have a σ around the order of 0.1. So it is fair to start with $\sigma^0 = 0.1$.

Also, ρ in Equation (11) is set to a small value (1.0×10^{-12}), not to impose too much smoothness on the illumination distribution because we are recovering spatially high frequency light sources. Note, as we used small lamps with shades as the light sources in our experiments (Figure 9), imposing smoothness terms on the illumination distribution itself does make sense.

4.3 Simulation Result

We ran a simulation test on the simultaneous estimation of the specular reflection parameter and the illumination distribution. We rendered six images of a sphere without any texture under the light sources depicted in Figure 6(a) with $\sigma = 0.06$, and ran the estimation algorithm with $\sigma^0 = 0.1$ and with the initial illumination hemisphere (Figure 6(b)) derived in the way mentioned in Section 3.3. After eight iterations, the illumination distribution converged to Figure 6(c) and $\sigma = 0.0586$. From Figure 6, we can see that the method can handle area light sources to some extent. As regularization is applied on the illumination distribution, the estimate

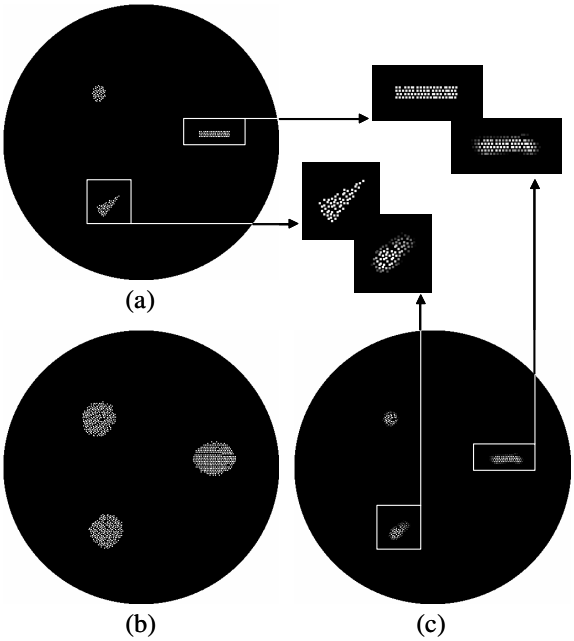

Fig. 6. Simulation result. (a) Ground truth illumination distribution. (b) Initial estimate. (c) Estimated illumination distribution. Each point light source is splatted for better visualization.

becomes a smoothed version of the ground truth lighting, where the intensity gradually falls off in the edge regions. Nevertheless, the shape of the illumination distribution is recovered quite well, and the roughness parameter (σ) is estimated accurately with only 2% difference.

5. RESULTS

We applied our framework to model the appearance of real objects from a sparse set of images, and rendered new views from novel viewpoints and under novel illumination conditions.

5.1 Inputs

The objects we used in our experiment are a mask made in Bali and a ceramic ornament of a woman. For each object, several range images were scanned with Vivid 900 [Minolta 2001]. They were registered and fused to produce a 3D mesh model. Additionally, we took six and twelve high dynamic range (HDR) images [Debevec and Malik 1997; Mitsunaga and Nayar 1999] for each object respectively, and used them as the input image sequence. For each image sequence, three light bulbs were placed above the object, albeit in different locations. A HDR image is obtained by combining several photographs taken with different shutter speeds. Figure 1(a) and Figure 7 shows the input images with the shutter speed approximately set to $\frac{1}{30}$.

To estimate the viewpoint of each input image, we captured a range image at the

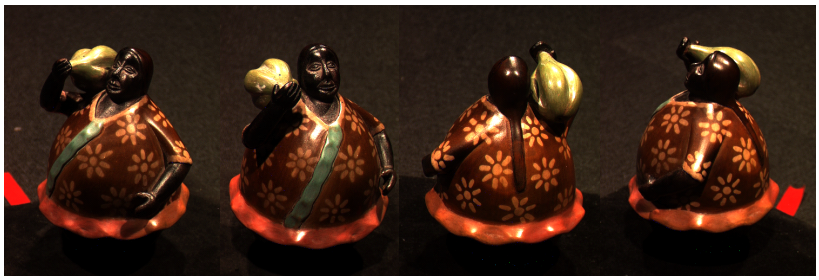

Fig. 7. Four out of twelve input HDR images for the woman object.

same viewpoint with a light-stripe range finder, and carried out 3D-3D alignment between the range data and the 3D mesh model.

5.2 Estimated parameters

Figure 8 shows a diffuse-texture-mapped image for each object viewed from one of the viewpoints in each image sequence. Note that the specular reflection is successfully removed, except slightly remaining on the tip of the woman's nose. This is because the nose of the woman object has high curvature, and thus the specular tends to remain at the same location in many input images.

Figure 9 depicts the initial and estimated final illumination hemisphere for the mask object. Note that it is impossible to manually specify the locations of the light sources from the initial estimation of the lighting environment (Figure 9:Left). To compare the estimated illumination distribution with the true illumination distribution, we captured the lighting condition using a CCD camera with fish-eye lens mounted, as shown in Figure 9. Note even if we neglect the glow around the light sources, because of the shades, the point light sources span a small area rather

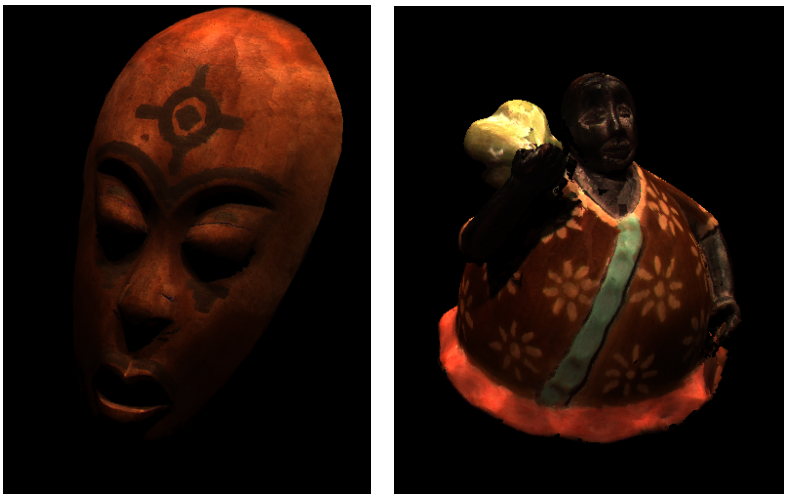

Fig. 8. Diffuse-texture-mapped images.

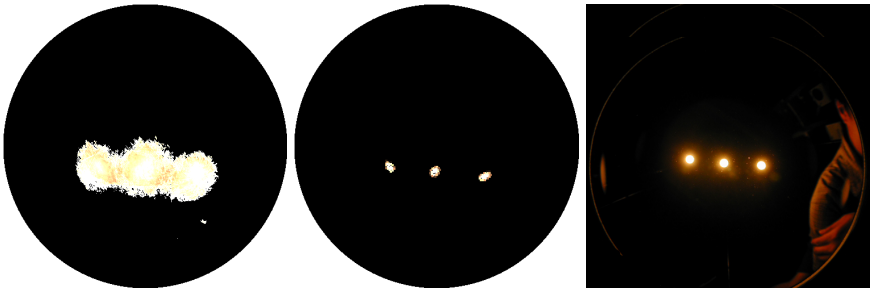

Fig. 9. Estimated illumination distribution for the mask sequence. Left: Initial illumination hemisphere. Middle: Estimated illumination hemisphere (splatted), Right: A photograph of the true illumination distribution captured through a fish-eye lens.

than being a real point. As we did not do any geometric calibration to enable direct comparison between the two images of the illumination distribution, we cannot see whether they match exactly. However, it is obvious that the estimated illumination hemisphere has an illumination distribution very close to the positions of the three point light sources. The intensity of the brightest point light source in the illumination hemisphere was 1370.83 and that of the darkest was 0.65. Note that these values are a combination of K_s and L_i.

By starting with 0.1 for σ, we obtained 0.0668 as its final estimated value. For the woman object, final estimate of σ was 0.0731.

5.3 View-dependent Rendering

With the estimated illumination hemisphere, the specular reflection parameter and the diffuse texture map, we are able to accomplish photorealistic view-dependent rendering from arbitrary viewpoints. Figure 10 shows the result of rendering the mask from a novel viewpoint. The left image in Figure 10 is a real photograph taken from the same viewpoint but not used in the input image sequence. As can be seen, the result is almost indistinguishable from the real photograph. The slight difference is mainly due to lack of perfection in the 3D model and consequent difference in computed normals. Figure 1(b) and Figure 11 show other examples of rendering from novel viewpoints.

5.4 Relighting

In addition to rendering from arbitrary viewpoints, we can render images under arbitrary novel illumination conditions. As we have estimated the original illumination distribution, we can compute the ratio of irradiance between the original and novel illumination distribution for each surface point. Hence, we can relight the diffuse texture without explicitly estimating $\mathbf{K}_{D,v}$ in Equation (2) [Marschner 1998; Levoy et al. 2000].

For instance, we can easily obtain illumination hemispheres corresponding to each original light bulb by masking the estimated illumination hemisphere. Figure 12 shows a real photograph and a rendered synthetic image of the woman with one of the three point light bulbs turned on. Despite the fact that we have some errors in the geometry which leads to slight differences in the two images, nevertheless, we

Fig. 10. Comparison between a synthetic image from a novel viewpoint and a real photograph. Left: real photograph, Right: synthetic image.

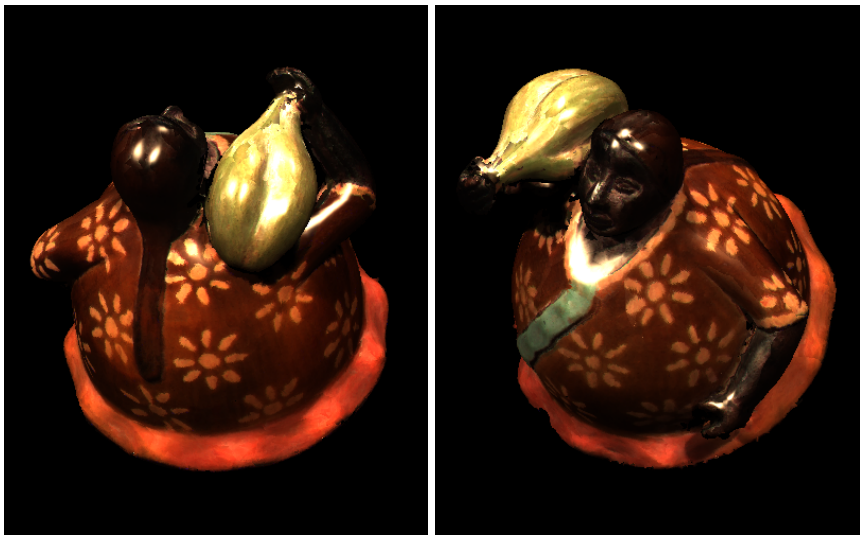

Fig. 11. Synthetic images rendered from novel viewpoints.

can see the diffuse texture is relighted accurately and the highlights are correct in shape and location. Figure 1(c) and Figure 13 show other example images rendered under novel lighting conditions.

Fig. 12. Comparison between a photograph and a synthetic image of the woman object captured/rendered under one light bulb. Left: real photograph, Right: synthetic image.

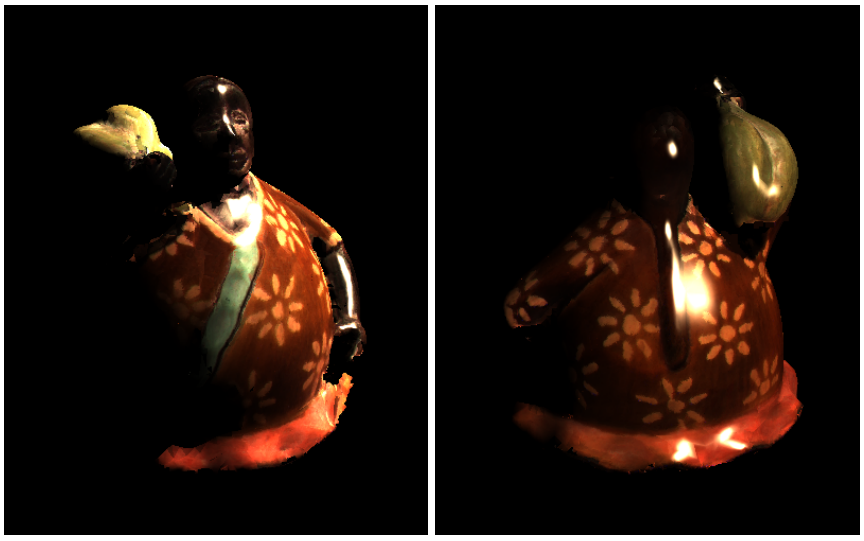

Fig. 13. Synthetic images rendered under novel lighting conditions.

6. DISCUSSION

In this section, we briefly discuss the limitations of our method.

The number of input images and the illumination color

As we construct the diffuse texture by extracting the minimum intensity value for each surface point, the necessary number of input images is determined by whether

at least one of the input images contain diffuse only reflection for each surface point. This, in fact, depends not only on the illumination distribution and the roughness of the object surface but also on the surface geometry. In other words, the consequent diffuse texture map will contain specular reflection where the illumination is fairly diffusive, or when the surface has a large σ value or where the object surface is extremely sharp. To overcome these problems, we are currently investigating a color-based reflection component separation approach. This will not only help extracting diffuse texture maps in difficult scenarios but also to recover illumination distributions containing different colored light sources which will let us relight the target object with different colors. The remaining problem, however, is that it is difficult to apply color-based (or color-constancy) methods to highly textured surfaces such as the object of a woman we used.

Radius of the illumination hemisphere

Currently we estimate the illumination distribution as point light sources mapped on a hemisphere with fixed radius. However, the light sources in place when capturing the input images were not necessarily placed in a uniform distance from the object. We empirically found the final illumination distribution estimate on the hemisphere represents cross sections of cones having their apex at the actual point light source position and holds the object inside tightly. As a result, estimates of a far point light source dominates a smaller region on the hemisphere compared to a closer point light sources. We are currently investigating a formal explanation for this.

Furthermore, by incorporating the fall off effects of lighting (the $\frac{1}{d^2}$ effect where d is the distance between a surface point and a light source) we might be able to recover the spatial distribution of light sources with different distances. This remains as our future work.

Diffuse light sources

We have shown that we are able to recover area light sources as well. However, it is obvious that we would not be able to recover area light sources that do not produce highlights. For instance, ambient and diffuse light will be embedded in the diffuse texture map and subsequent relighting of those lights would not be possible.

7. CONCLUSION

We have shown a method to accomplish photorealistic view-dependent rendering and relighting from a sparse set of images and a pre-acquired geometric model of the target object. We take an approach similar to inverse rendering, however, in a significantly different manner: we estimate all three components necessary for rendering in a representation that is compact but contains enough information, which includes the diffuse texture map, the specular reflectance parameter and the illumination hemisphere. The key idea is to first separate the reflection components to obtain a texture map that represents the diffuse reflection and images whose pixel values' dominant factor is specular reflection. The latter images are then used to recover the illumination distribution and reflectance parameters of the Torrance-Sparrow reflection model via 2D blind deconvolution with special care taken for the inherent ambiguity of lighting and reflectance. We described the method in theory, and showed its effectiveness through simulation and experiments with real-world

objects. As discussed in Section 6, future work includes recovery of the spatio-spectral illumination distribution and the distances to light sources.

REFERENCES

Boivin, S. and Gagalowicz, A. 2001. Image-based rendering of diffuse, specular and glossy surfaces from a single image. In *ACM SIGGRAPH 01*. 197–116.

Chan, T. and Wong, C. 1998. Total Variation Blind Deconvolution. *IEEE Transactions on Image Processing 7*, 370–375.

Chen, W.-C., Bouguet, J.-Y., Chu, M., and Grzeszczuk, R. 2002. Light Field Mapping: Efficient Representation and Hardware Rendering of Surface Light Fields. *ACM Trans. on Graphics 21,* 3, 447–456.

Curless, B. and Levoy, M. 1996. A volumetric method for building complex models from range images. In *Computer Graphics Proceedings, ACM SIGGRAPH 96*. 303–312.

Dana, K., Ginneken, B., Nayar, S., and Koenderink, J. 1999. Reflectance and texture of real-world surfaces. *ACM Transactions on Graphics 18,* 1 (Jan.), 1–34.

Debevec, P., Hawkins, T., Tchou, C., Duiker, H.-P., Sarokin, W., and Sagar, M. 2000. Acquiring the Reflectance Field of a Human Face. In *Computer Graphics Proceedings, ACM SIGGRAPH 00*. 145–156.

Debevec, P. and Malik, J. 1997. Recovering High Dynamic Range Radiance Maps from Photographs. In *Computer Graphics Proceedings, ACM SIGGRAPH 97*. 369–378.

Garland, M. and Heckbert, P. 1997. Surface Simplification using Quadric Error Metrics. In *Computer Graphics Proceedings, ACM SIGGRAPH 97*. 209–216.

Gill, P., Murray, W., and Wright, M. 1981. *Practical Optimization*. Academic Press.

Gortler, S., Grzeszczuk, R., Szeliski, R., and Cohen, M. 1996. The Lumigraph. In *Computer Graphics Proceedings, ACM SIGGRAPH 96*. 43–54.

Horn, B. 1986. *Robot Vision*. McGraw-Hill Book Company.

Ikeuchi, K. and Sato, K. 1991. Determining Reflectance Properties of an Object Using Range and Brightness Images. *IEEE Trans. Pattern Analysis and Machine Intelligence 13,* 11 (Nov.), 1139–1153.

Klinker, G., Shafer, S., and Kanade, T. 1990. A Physical Approach to Color Image Understanding. *International Journal of Computer Vision 4*, 7–38.

Lambert, J. 1760. *Photometria sive de mensura de gratibus luminis colorum et umbrae*. Augsburg.

Lensch, H., Kautz, J., Goesele, M., Heidrich, W., and Seidel, H.-P. 2001. Image-Based Reconstruction of Spatially Varying Materials. In *Eurographics Rendering Workshop*. 104–115.

Levoy, M. and Hanrahan, P. 1996. Light Field Rendering. In *Computer Graphics Proceedings, ACM SIGGRAPH 96*. 31–42.

Levoy, M., Pulli, K., Curless, B., Rusinkiewicz, S., Koller, D., Pereira, L., Ginzton, M., Anderson, S., Davis, J., Ginsberg, J., Shade, J., and Fulk, D. 2000. The Digital Michelangelo Project: 3D Scanning of Large Statues. In *Computer Graphics Proceedings, ACM SIGGRAPH 00*. 131–144.

Lu, R., Koenderink, J., and Kappers, A. 1998. Optical properties (bidirectional reflection distribution functions) of velvet. *Applied Optics 37,* 25, 5974–5984.

Marschner, S. 1998. Inverse Rendering for Computer Graphics. Ph.D. thesis, Cornell University.

Marschner, S. and Greenberg, D. 1997. Inverse Lighting for Photography. In *Proc. of IS&T/SID Fifth Color Imaging Conference*. 262–265.

Marschner, S., Westin, S., Lafortune, E., and Torrance, K. 2000. Image-Based Measurement of the Bidirectional Reflectance Distribution. *Applied Optics 39,* 16, 2592–2600.

Minolta. 2001. *Vivid 900*. http://www.minolta-rio.com/vivid/.

Mitsunaga, T. and Nayar, S. 1999. Radiometric Self Calibration. In *Proc. of Computer Vision and Pattern Recognition '99*. 374–380.

Nayar, S., 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.

Nayar, S., Ikeuchi, K., and Kanade, T. 1991. Surface reflection: physical and geometrical perspectives. *IEEE Transactions on Pattern Analysis and Machine Intelligence 13,* 7 (Jul.), 611 – 634.

Nayar, S. and Oren, M. 1995. Generalization of the Lambertian Model and Implications for Machine Vision. *International Journal of Computer Vision 14*, 227–251.

Nishino, K., Sato, Y., and Ikeuchi, K. 1999. Eigen-Texture Method: Appearance Compression based on 3D Model. In *Proc. of Computer Vision and Pattern Recognition '99*. Vol. 1. 618–624.

Press, W., Teukolsky, S., Vetterling, W., and Flannery, B. 1992. *Numerical Recipes in C, Second Edition*. Cambridge University Press.

Ramamoorthi, R. and Hanrahan, P. 2001. A Signal-Processing Framework for Inverse Rendering. In *Computer Graphics Proceedings, ACM SIGGRAPH 01*. 117–128.

Sato, I., Sato, Y., and Ikeuchi, K. 1999. Illumination distribution from brightness in shadows: adaptive estimation of illumination distribution with unknown reflectance properties in shadow regions. In *Proc. of Seventh International Conference on Computer Vision ICCV '99*. Vol. 1. 875–882.

Sato, K. and Inokuchi, S. 1987. Range-imaging system utilizing nematic liquid crystal mask. In *First International Conference on Computer Vision*. 657–661.

Sato, Y., Wheeler, M., and Ikeuchi, K. 1997. Object shape and reflectance modeling from observation. In *Computer Graphics Proceedings, ACM SIGGRAPH 97*. 379–387.

Shafer, S. 1985. Using Color to Separate Reflection Components. *COLOR Research and Application 10,* 4, 210–218.

Shum, H.-Y. and L-W.He. 1999. Rendering with Concentric Mosaics. In *Computer Graphics Proceedings, ACM SIGGRAPH 99*. 299–306.

Torrance, K. and Sparrow, E. 1967. Theory for off-specular reflection from roughened surfaces. *Journal of Optical Society of America* 57, 1105–1114.

Turk, G. and Levoy, M. 1994. Zippered polygon meshes from range images. In *SIGGRAPH 94*. 311–318.

Ward, G. 1992. Measuring and modeling anisotropic reflection. In *Computer Graphics Proceedings, ACM SIGGRAPH 92*. 265–272.

Wood, D., Azuma, D., Aldinger, K., Curless, B., Duchamp, T., Salesin, D., and Stuetzle, W. 2000. Surface Light Fields for 3D Photography. In *Computer Graphics Proceedings, ACM SIGGRAPH 00*. 287–296.

You, Y. and Kaveh, M. 1996. A regularization approach to joint blur identification and image restoration. *IEEE Transactions on Image Processing 5*, 416–428.

Zhang, Z. 1999. Flexible Camera Calibration By Viewing a Plane From Unkown Orientations. In *Proc. of Seventh International Conference on Computer Vision ICCV '99*. 666–673.