Image-Based Acquisition of Shape and Spatially Varying Reflectance

Tom Haber and Philippe Bekaert
Hasselt University - tUL - IBBT
Expertise Centre for Digital Media
Wetenschapspark 2, 3590 Diepenbeek, Belgium
firstname.lastname@uhasselt.be

Abstract

The shape and reflectance of complex objects, for use in computer graphics applications, cannot always be acquired using specialized equipment due to cost or pratical considerations. We want to provide an easy and cost-effective way for the approximate recovery of both shape and spatially-varying reflectance of objects using commodity hardware.

In this paper, we present an image-based technique for recovering 3D shape and spatially-varying reflectance properties from a sparse set of photographs, taken under varying illumination. Our technique models the reflectance with a set of low-parameter BRDFs without knowledge of the location of the light-sources or camera. This results an a flexible and portable system that can be used in the field.

We successfully apply the approach to several objects (synthetic and real), recovering shape and reflectance. The acquired information can then be used to render the object with modifications to geometry and lighting via traditional rendering methods.

1 Introduction

Many applications in entertainment, augmented/virtual reality, architecture and digital museums require photorealistic rendering of real-world objects from novel viewpoints and illumination.

This requires the use of realistic models for all components of image synthesis, including geometry, light sources, materials and cameras. These models can each be acquired using specialized equipment such as a 3d-scanner and a gonioreflectometer [10, 11, 14]. But this is generally a time and resource consuming process. We want to provide an easy and cost-effective way for the approximate recovery of both shape and spatially-varying reflectance of objects using commodity hardware. Additionally, we want the setup to be portable and capable of capturing a wide range of objects.

To this end, we present an image-based method that uses a small number of photographs of inanimate objects taken from a fixed viewpoint. In each photograph, the object is lit with an ordinary light source at a different location, however no knowledge about the location and strength of the light sources is required by the system. With such a

set of photographs, we observe the surface points only with a small number of illumination angles. As a result, there is too little data to determine a full bidirectional reflectance function for each surface point. We address this issue in two ways. First, we limit the reflectance to a low-parameter model. Secondly, we use the observation that many objects can be decomposed into a small number of materials [2, 9]. We cluster the acquired samples into groups of similar materials and fit a model to each group.

In our work, we reconstruct both shape and spatially-varying reflectance without making any additional assumptions by alternating between the optimization of geometry, lighting (direction and intensity), diffuse reflectance (albedo) and specular reflectance each time fixing all other unknowns. In particular, our contributions are

- We allow a wide range of objects to be captured using a standard camera with any type of lens.

- The system is flexible and portable enough to be used in the field.

- We present an improved optimization approach suited for our needs.

We require only a relatively small number of photographs (about 10-15 images for one object), thereby speeding up the acquisition phase.

After our reconstruction, many editing methods are made possible. In addition to changes in lighting and viewpoint, the albedo and materials, represented as texturemaps, can be easily edited using conventional image editing software.

2 Related Work

In traditional photometric stereo algorithms [18, 19], the surface is assumed to be lambertian or spatially uniform. More recent work [1–3, 12] applies non-lambertian models to photometric stereo. However, most require calibration of the light source, reference objects [3, 4] or a known reflectance map. Our method works without calibration and prior knowledge.

Georghiades et al. [1] proposed a method that accommodates for the spatial variation of the diffuse reflectance, but assumes the specular properties are constant. Our approach allows both reflectance properties to vary across the surface, permitting the reconstruction of objects comprised of several materials or a mixture of paints with different reflectance characteristics.

Goldman et al. [2] describes a photometric stereo method for surfaces with spatially-varying reflectance, including both diffuse and specular properties. This method decomposes the object's material into a small number of fundamental BRDFs and recovers both shape and weight maps for the materials. However, they assume the lighting directions are known, whereas we recover them from the input images. Additionally, we believe our BRDF optimization scheme is less sensitive to overfitting due to its non-local nature.

Paterson et al. [13] present a system for capturing the geometry of complex materials with varying albedo and BRDF. A digital camera with attached flash is used to sample the geometry from different angles and illumination. The system however requires the geometry to be roughly planar.

Image-based BRDF acquisition has already been widely researched. The proposed methods [10, 11, 14] generally separate the measurement of the BRDF from the shape

estimation: they either use samples with known shape (spherical or planar) or the geometry is obtained using a 3d-scanner. Lensch et al. [9] cluster reflectance samples over a known surface into groups of similar materials and fit a Lafortune [7] model to each group in an iterative way. Their method however requires approximate geometry, whereas our approach reconstructs the surface solely from the images.

We argue that the system we present is simple enough to allow widespread use, depends on a minimum amount of calibration and provides accurate measurements for a variety of materials and objects.

3 Problem Statement

The input of our system is a set of photographs of an inanimate object taken under a single varying distant illuminant. From these images, we seek to reconstruct the shape and reflectance properties of that object as well as the illumination, ignoring effects such as interreflection, transparency and translucency. In general, the contribution from indirect illumination is relatively small and can be ignored.

Our material model is motivated by the observation that real world variations in appearance across a surface are often the result of a composition of a small number of BRDFs [2, 4, 9]. Therefore, the reflectance of a surface point is modeled as a diffuse part and a mixture of the basis BRDFs. The number of BRDFs to solve for is given by the user.

Lastly, in contrast to most related work, we model our camera as perspective. This has many advantages over orthographic cameras: perspective cameras are more common and hence less expensive. Most researchers therefore approximate an orthographic camera using a regular camera with a telephoto lens placed at a distance. This solution is still relatively expensive and also limits the size of objects that can be captured. It is for example impractical to acquire images of a statue. The use of an orthographic model also introduces errors that can be avoided with a perspective model [15]. Perspective cameras allow different kinds of lenses to be used (e.g. wide-angle, macro lenses). The downside however is that a calibration step is required for the camera parameters. On the other hand, this is a well-studied and understood problem and can be achieved with a minimum of effort [17].

In the following sections, we describe the representations used for the geometry and BRDFs.

3.1 Shape Model

The most commonly used representation of the object surface in photometric stereo is a depth function $z(x,y)$ parametrized by the image plane coordinates [1, 2, 8, 19]. These depth values can only be indirectly obtained from the shading by first estimating the normals and integrating. However the normal field is not necessarily integratable. This can be solved by enforcing integrability [2, 6, 19] or by expanding the depth function $z(x,y)$ using basis functions [1, 8]. The latter methods are better suited when dealing with a perspective camera model and non-diffuse reflectance.

Our surface representation is inspired by Lee et al. [8]. We discretize the object surface into a set of triangular elements. We construct these elements by dividing the image domain into triangles T_i with vertices p_i. Then, the surface is approximated by triangular

surface patches S_k with vertices P_i such that S_i and P_i project to T_k and p_i respectively. This surface is uniquely specified by all P_i, or equivalently the depths Z_i associated with all p_i. The surface vertices P_i can be expressed in terms of $p_i = (x_i, y_i)$ and Z_i as

$$(\frac{x_i Z_i}{f}, \frac{y_i Z_i}{f}, Z_i),$$

where f is the focal length of the camera.

3.2 BRDF Model

The surface BRDF is generally represented by a four-dimensional bi-directional reflectance distribution function. These four parameters are the light direction angles (θ_i, ϕ_i) and viewing angles (θ_o, ϕ_o). In this work, we restrict ourselves to the isotropic variant where the angular variation reduces to three dimensions.

Recovering the entire BRDF is not feasible due to the large number of samples required. We tackle this problem by imposing some parametric model of the reflectance function. This simplifies to problem to estimating a handful of parameters [1, 2]. The lack of generality of such a parametric model also avoids incorrectly incorporating errors made in the estimation of geometry and lighting. We represent our reflectance by the combination of a lambertian model and a simplified specular Torrance-Sparrow lobe [16]:

$$f_s(\theta_i, \theta_o, \phi) = \alpha_s \frac{e^{-v^2 \theta_a^2}}{\cos(\theta_i) \cos(\theta_o)}$$

with (α_s, v) the specular parameters and θ_a the angle between normal and halfway vector.

3.3 Lighting Model

Since we consider our light sources to be directional, the color of a pixel $I(x)$ due to light source L_i can be expressed as

$$I(x) = [\alpha(x) + \sum_m \omega_m f_s(\alpha_m, n(x), v(x), L_i)] * n(x)^T L_i \tag{1}$$

We use α to denote the diffuse reflectance (albedo) for each pixel. The function f_s represents the Torrance-Sparrow BRDF with parameters α_m, with associated weight ω_m. The normal is denoted as $n(x)$ and viewing direction as $v(x)$.

4 Shape and Reflectance Reconstruction

We acquire k images of the object with different illumination, denoted as I_k. As the images might contain shadows and saturations, which are not in our model, we need to mark such pixels as invalid. Saturations are simply determined by thresholding, while shadowed pixels are found by thresholding the ratio between the pixel intensity and the albedo (similar to Georghiades et al. [1]). If the variance of a pixel's reflectance samples is low, all the samples are removed. This effectively makes that pixel part of the background.

The initial albedo is obtained by sorting the measured reflectance values for each pixel, resulting in images P_i where P_0 is the dimmest and P_{n-1} the brightest and selecting the image P with no highlight nor shadow. For a set of 20 images, we found that P_{12} is a robust choice.

Based on the described model, we solve the following minimization problem to determine shape and reflectance properties

$$\min_{\alpha,\alpha_m,S,L_i,\omega_m} \sum_{k,x} |I_k(x) - I(x,\alpha,\alpha_m,Z,L_i,\omega_m)|^2 \qquad (2)$$

Optimization of this problem is non-trivial: many optimization algorithms can easily become trapped in local minima and the high dimensionality makes standard algorithms intractable. We solve this problem in an iterative fashion by decomposing it into a set of smaller problems first. Such an optimization is also performed by Georghiades et al. [1] and Goldman et al. [2]. Although this strategy still does not guarantee to converge to the optimal solution, we have found it gives good results with some modifications.

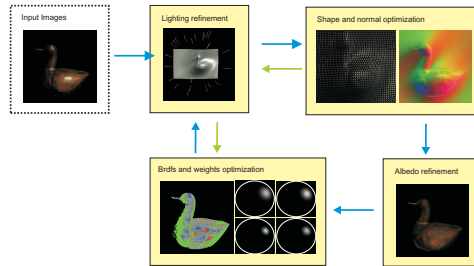Our optimization approach is outlined in the next section.

5 Algorithm

Figure 1: An overview of the algorithm. The blue arrows show the flow between the different optimization components. The green arrows indicate the tight coupling between some components.

Our algorithm is composed of several components, each of which is described in detail in the following sections. Figure 1 gives an overview of these components.

Initialization. In this phase, we compute an initial estimate for the diffuse albedo as described in §4 and initialize all depth variables to a arbitrary constant and the specular BRDF parameters to 0.

We obtain a decent estimate of the light directions from lambertian photometric stereo [18, 19]. This part is further described in section 5.1.

After initialization, the system optimizes the objective function by repeating the following steps:

1. Refine shape. The depth values for our shape model are optimized while holding the light directions and material attributes constant. We extended the algorithm by Lee et al. [8] to incorporate BRDFs and added smoothness constraints.

After this optimization, we optimize the normals further per pixel.

2. Update the lambertian albedo. While the other optimization parameters are kept fixed, we compute a new albedo for each pixel in a linear least squares fashion.

3. Refine the light source directions. We refine the directions and strengths of the light sources, while holding the other parameters constant. The light source directions are refined using Newton's method and their strengths using linear least squares.

4. Optimize BRDF parameters. The BRDF parameters are optimized using an iterative clustering and fitting procedure, similar to Lensch's et al. [9] Further details can be found in Section 5.5.

5. Termination Repeat steps 1-4 until the estimate converges. Each step is guaranteed to decrease the objective function.

5.1   Initialization

Yuille et al. [19] outline a method to decompose the intensity matrix of M images into the M light directions and the product of normal and albedo per pixel up to an arbitrary transformation. Since we are only interested in the light directions, we can use a simpler, more efficient technique: we perform an eigenvalue analysis on JJ^T and obtain the light directions from the three eigenvectors with the biggest eigenvalues. This however can not be expected to work in the presence of shadows, ambient light, specularities and noise. We therefore sift out all the columns of the intensity matrix which contain invalid measurements as described in Section 4.

The results obtained from lambertian photometric stereo are subject to the so-called bas-relief ambiguity [18,19]. At this point, we can introduce some prior knowledge about the lighting conditions: most of the light sources will be placed in front of the object. We therefore rotate the obtained light directions so that the average direction coincides with the optical axis of the camera. We also introduce a user specified rotation around this axis to resolve further ambiguities.

5.2   Shape Reconstruction

We now only solve for shape and fix all other parameters. Note that since the normals and viewing directions are all expressed in terms of the depth at the vertices $Z_i, i = 0, \ldots, M$, the reflectance function $I(x)$ is only a function of the depths Z_i, Z_j and Z_l of the supporting triangle:

$$\sum_x \sum_k |I_i(x) - I(x, S, L_k)|^2 = \sum_x \sum_k |I_k(x) - I(Z_i, Z_j, Z_l)| \tag{3}$$

We minimize this equation by successively linearizing the nonlinear function $I(Z_i, Z_j, Z_l)$ around the previously estimated solution, following Lee et al. [8]. The resulting linear equation is then efficiently solved using conjugate gradient. To ensure a smooth solution, we altered the functional to additionally minimize the laplacian at each vertex.

The shape model gives a fairly coarse representation of the geometry and normals, therefore we apply an additional normal refinement per pixel using a multi-scale grid search: the error is successively computed for a set of normals in a certain radius around the estimated normal. Each time the best normal is kept and the radius is decreased.

5.3 Albedo Refinement

Note that the optimization is now linear with respect to the albedo and each albedo value is decoupled from all the others. Hence we can easily solve for it using linear least squares. This boils down to

$$\alpha(x) = \frac{\sum_k I_k(x) - \sum m \omega_m f_s(\alpha_m, n(x), v, L_k)}{\sum_k n(x)^T L_k}$$

5.4 Light source Refinement

Note that light source of each frame can be updated independently. We utilize Newton's method since the direction of light source is only two dimensional. We found the error function sufficiently smooth for Newton's method to work. To avoid converging to a local minima in the early steps of the algorithm, we added a prior to the objective function such that the pixels, that are believed to be due to a specular highlight, have a reflection direction close to the viewing direction. We determine these pixels by thresholding the difference between the real reflectance and our current estimate of the diffuse reflectance. This prior creates a smoother error function and aids in a faster convergence.

The strengths are determined in a least squares fashion similar to the albedo refinement.

5.5 BRDF Optimization

The BRDF optimization consists of the split-recluster-fit and project process as presented by Lensch et al. [9] with some minor modifications: instead of the Lafortune model, we employ a Torrance-Sparrow model. In the fitting step, we also weight the reflectance samples based on the product of incident and outgoing angles and filter out any samples with grazing angles (> 80 degrees) to avoid fitting errors. In each iteration, only one split-recluster-fit step is executed to prevent overfitting.

We believe this procedure to be less sensitive to overfitting than the segmentation based method of Goldman et al. [2] because it is inherently a non-local optimization.

5.6 Tight coupling

After some experiments, we observed that a tighter coupling between the light refinement and both normal and BRDF optimization results in a better convergence and further helps the reconstruction to avoid local minima. The tight coupling is obtained by placing the light refinement within the error functions of the normal and BRDF optimizations. This way, the optimizations are performed on the error under the best possible light estimate.

6 Results

To capture our input images, we used a Canon 20D camera with a 17-55mm lens and a single exposure for each lighting direction. We used a small halogen light source for the lighting. The images are captured at full resolution of the camera (3504x2336), but the results were computed at a down-sampled resolution (640x480). A typical capture session consists of 12 lighting directions and takes about 10 min.
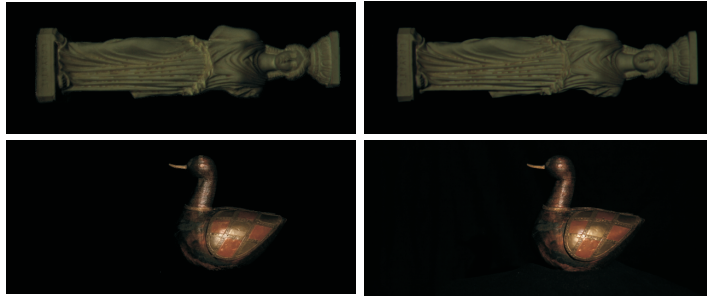
Figure 2: Comparison to ground truth. Left: models rendered under novel lighting conditions not in training set. Right: images under approximately the same lighting condition.
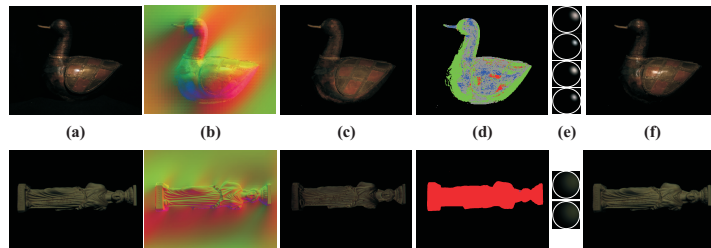

(a)　　(b)　　(c)　　(d)　　(e)　　(f)

Figure 3: Input, model and reconstruction for some objects. a) Source Image. b) Recovered normal map. c) Recovered albedo. d) Recovered weight map (false color). e) Spheres shaded with basis BRDFs. f) Model rendered under original illumination.

Most of the examples converged after 8-10 iterations, in about 2-4 hours on a 3Ghz Pentium 4. Some steps of the algorithm make use of the GPU to accelerate shading and error computation.

Since our goal is to produce plausible reconstructions, we show a few objects under different viewpoints and lighting conditions in Figures 4(a) and 2. The relighting results are compared with real photos taken under similar lighting conditions not present in the input set. Note the similarity of these images.

Figure 3 shows the different components of the obtained model for a range of objects. Note that the algorithm estimates detailed normal maps, reflectance properties and lighting conditions for each object. A few artifacts still occur in areas where highlights are observed in most of the input images, such as the neck of the duck. The presence of highlights in all of the images causes the shape and normal estimation to overshoot the ideal solution. The impact of these artifacts on the rerendering and relighting of these objects is however minor.

The obtained material reconstruction also makes many editing techniques possible. The obtained BRDFs can be altered, an example of this is shown in Figure 4(b). Additionally, texture synthesis [5] techniques can be applied on the albedo and weightmaps.

<div style="text-align:center">(a)   (b)</div>

Figure 4: a) Synthesized views of our reconstructions. b) The ducks material properties have been changed to a much more specular BRDF by direct manipulation of the parameters.

7 Discussion and Future Work

We have demonstrated a method that acquires shape and spatially varying BRDFs from a set of photographs under varying illumination. The fidelity of the reconstructions are lower than those obtained using methods with more observations or knowledge about some of the unknowns (lighting, materials or geometry). Nonetheless, we are able to acquire a range of models that can be reused under different viewing and lighting conditions.

Our approach is able to capture shape and BRDFs using a small number of photos and without the need for specialized equipment. We only require a camera and a light source. This results in a compact and portable setup.

The reconstruction depends on the lighting conditions used during acquisition. The optimization is subject to overfitting if the light sources are too similar or when highlights are always observed in the same area. Additional smoothness terms in the objective function might ameliorate these artifacts.

In future work, we want to investigate methods that take shadows and interreflections into account. Both contain information that can be used during shape and material reconstruction.

Acknowledgements

References

[1] Athinodoros S. Georghiades. Recovering 3-d shape and reflectance from a small number of photographs. In *Proceedings of EGRW '03*, pages 230–240, 2003.

[2] Dan B Goldman, Brian Curless, Aaron Hertzmann, and Steven M. Seitz. Shape and spatially-varying brdfs from photometric stereo. In *Proceedings of ICCV 2005*, 2005.

[3] A. Hertzmann and S. Seitz. Shape and materials by example: A photometric stereo approach, 2003.

[4] A. Hertzmann and S.M. Seitz. Example-based photometric stereo: Shape reconstruction with general, varying brdfs. *PAMI*, 27(8):1254–1264, August 2005.

[5] Aaron Hertzmann, Charles E. Jacobs, Nuria Oliver, Brian Curless, and David H. Salesin. Image analogies. In *ACM SIGGRAPH 2001*, pages 327–340, 2001.

[6] B.K.P. Horn and M.J. Brooks. *Shape from Shading*. Book, 1989.

[7] Eric P. F. Lafortune, Sing-Choong Foo, Kenneth E. Torrance, and Donald P. Greenberg. Non-linear approximation of reflectance functions. *Computer Graphics*, 31(Annual Conference Series):117–126, 1997.

[8] K.M. Lee and C.C.J. Kuo. Shape from shading with a generalized reflectance map model. *CVIU*, 67:143–160, 1997.

[9] Hendrik P. A. Lensch, Jan Kautz, Michael Goesele, Wolfgang Heidrich, and Hans-Peter Seidel. Image-based reconstruction of spatial appearance and geometric detail. *ACM Trans. Graph.*, 22(2):234–257, 2003.

[10] S. R. Marschner, S.H Westin, E. P. F. Lafortune, and K. E. Torrance. Image-based brdf measurement including human skin. In *EGRW '99*, pages 139–152, 1999.

[11] Wojciech Matusik, Hanspeter Pfister, Matt Brand, and Leonard McMillan. A data-driven reflectance model. *ACM Trans. Graph.*, 22(3):759–769, 2003.

[12] S.K. Nayar, K. Ikeuchi, and T. Kanade. Determining shape and reflectance of hybrid surfaces by photometric sampling. *RA*, 6(4):418–431, August 1990.

[13] James A. Paterson, David Claus, and Andrew W. Fitzgibbon. BRDF and geometry capture from extended inhomogeneous samples using flash photography. *Computer Graphics Forum*, 24(3):383–391, 2005.

[14] Yoichi Sato, Mark D. Wheeler, and Katsushi Ikeuchi. Object shape and reflectance modeling from observation. In *SIGGRAPH '97*, pages 379–387, 1997.

[15] Ariel Tankus and Nahum Kiryati. Photometric stereo under perspective projection. In *ICCV '05*, pages 611–616, 2005.

[16] K.E. Torrance and E.M. Sparrow. Theory for off-specular reflection from roughened surfacesw. *JOSA*, 57:1105–1114, 1967.

[17] R.Y. Tsai. A versatile camera calibration technique for high-accuracy 3d machine vision metrology using off-the-shelf tv cameras and lenses. *RA*, 3(4):323–344, 1987.

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

[19] A. Yuille and D. Snow. Shape and albedo from multiple images using integrability. In *CVPR 97*, pages 158–164, 1997.