Shape and Reflectance from Natural Illumination

Geoffrey Oxholm and Ko Nishino

Department of Computer Science
Drexel University, Philadelphia, PA 19104, USA
{gao25,kon}@drexel.edu

Abstract. We introduce a method to jointly estimate the BRDF and geometry of an object from a single image under known, but uncontrolled, natural illumination. We show that this previously unexplored problem becomes tractable when one exploits the orientation clues embedded in the lighting environment. Intuitively, unique regions in the lighting environment act analogously to the point light sources of traditional photometric stereo; they strongly constrain the orientation of the surface patches that reflect them. The reflectance, which acts as a bandpass filter on the lighting environment, determines the necessary scale of such regions. Accurate reflectance estimation, however, relies on accurate surface orientation information. Thus, these two factors must be estimated jointly. To do so, we derive a probabilistic formulation and introduce priors to address situations where the reflectance and lighting environment do not sufficiently constrain the geometry of the object. Through extensive experimentation we show what this space looks like, and offer insights into what problems become solvable in various categories of real-world natural illumination environments.

1 Introduction

Recovering the geometry of an object from its appearance is a longstanding and challenging computer vision problem. To simplify the problem, traditional shape from shading and photometric stereo methods have made extensive use of controlled lighting environments. Limiting the illumination environment to a single, known, point light source, for example, enables the observed appearance to give strong clues about the surface orientation. Without strong prior knowledge, however, a single observation is insufficient on its own.

Many methods additionally assume that the reflectance of the object is known a priori. It is well known that three illuminating light sources are sufficient to determine the geometry of a Lambertian object [18]. This situation is equivalent to a lighting environment in which one pixel is red, one blue, and one green. Such a lighting environment would be useless, however, in determining the shape of a purely reflective object, as only three orientations would reflect any light at all.

In this work we show that real-world, unconstrained, natural illumination contains within it strongly discerning information that can be used to recover not just the geometry of the object, but its reflectance as well. We show that for certain lighting environments and materials, this problem is sufficiently constrained, and in other conditions ambiguity is impossible to avoid. In between these extrema we find the illumination

A. Fitzgibbon et al. (Eds.): ECCV 2012, Part I, LNCS 7572, pp. 528–541, 2012.

conditions that are most common in the world around us. Unique regions of real-world lighting environments (e.g. stained glass) provide strong clues about surface orientation, while others are less descriptive (e.g. overcast sky). Our primary contribution is a method that flexibly deals with ambiguity by leveraging the structure of the lighting environment.

In a given lighting environment, the appearance of an object is a function of its reflectance and geometry. Since neither of these factors can be estimated without knowledge of the other, we must estimate them jointly. To do so, we derive a probabilistic formulation in which these two conditionally independent factors are linked by their combined role in the image formation process.

In order to recover the underlying geometry and reflectance we adopt an expectation-maximization framework with a maximization step that alternates between estimating one variable while holding the other fixed. Using the directional statistics BRDF model [15, 16], we are able to quickly estimate the reflectance for a current estimate of the surface geometry. The reflectance estimate and the lighting environment together form a reflectance map. By comparing the appearance at a given point with the reflectance map we arrive at a nonparametric distribution of likely surface orientations. On occasion, certain surface patches exhibit color values unique to a specific orientation. More often, however, the distribution contains large regions of possible orientations. By linking neighboring pixels with novel priors in this reflectance map, we are able to propagate the information captured by highly informative observations across less descriptive regions. Further, we introduce a notion of scale by limiting surface normal orientations to discrete intervals on a geodesic hemisphere. By incrementally increasing the resolution of the hemisphere, we mitigate the increased uncertainty that comes from highly specular materials, and featureless lighting regions.

We extensively test our method under various lighting conditions with several materials. In addition to conclusively demonstrating the accuracy and flexibility of our method on both real and synthetic data, we also offer insights into the reflectance-illumination space and show that certain subspaces result in highly discriminative reflectance maps that facilitate longstanding problems.

2 Related Work

Shape from shading and photometric stereo have been popular areas of research for decades. Here we will cover only those methods that relate specifically to the defining characteristics of our work, namely joint geometry and reflectance estimation, and geometry recovery under natural illumination. We recommend surveys by Durou et al. [6] and Zhang et al. [19] for a more thorough history of these problems.

Methods to achieve joint reflectance and geometry estimation have been found in recent photometric stereo papers. Goldman et al. [7], and Alldrin et al. [1, 2], for example, estimate spatially varying reflectance and geometry using many images under controlled illumination. Although our inputs—a single image under uncontrolled, but known illumination—are different from theirs, they note, as we do, that fixing one factor to estimate the other in an iterative scheme enables joint optimization. To further constrain the solution space, Goldman et al. assume an isotropic Ward model of reflectance, and Alldrin et al. require approximately 100 images. We adopt the isotropic

directional statistics BRDF model introduced by Nishino and Lombardi [16], which is compact, expressive, and may be estimated in natural illumination.

Hertzmann and Setiz [8] also use point light sources, though they do not assume their direction or number is known. A reference sphere is painted with the same material as the target object. Because the sphere has known geometry, it serves as a measured reflectance map for the target object. After a sequence of images has been taken, the surface orientation at each pixel may be estimated by finding the location on the sphere with the most similar sequence of appearances. In our formulation it is rare that the observed appearance in our single image is linked directly to a single orientation. Instead, the reflected light captured at each pixel has a distribution of possible orientations that may explain it. We overcome this ambiguity by probabilistically linking neighboring pixels with novel priors. This allows pixels with tight distributions to pull neighboring regions towards their proper orientation.

Our work is most similar to recent work on shape recovery in natural illumination by Huang and Smith [9], and Johnson and Adelson [11]. The primary distinction is that our work allows for objects of arbitrary isotropic BRDF whereas these authors have assumed a Lambertian material. This assumption allows the authors leverage insights from recent work on spherical harmonics [3] to assume a parametric reflectance map. Huang and Smith estimate the parameters to a first-order approximation yielding a billinear reflectance map, while Johnson and Adelson assume an a priori known second-order approximation, which yields a quadratic reflectance map. Both authors demonstrate that in certain lighting environments the orientation responsible for the observed appearance can be well estimated with a least squares optimization. Our problem is significantly harder. In addition to the added problem of estimating the reflectance, by allowing for arbitrary isotropic BRDFs, we move from a parametric to a non-parametric reflectance map in which many possible orientations may explain an observed appearance. This complex problem becomes tractable through priors carefully derived to leverage the sparse, but essential, orientation clues reflected from the lighting environment.

3 Probabilistic Shape and Material Estimation

In this work we assume that the natural illumination environment $\boldsymbol{L}$ is known, but uncontrolled. Using this information, and a single (orthographic) image of an object $\boldsymbol{I}$, we estimate the other two factors in the appearance of the object—the BRDF parameters $\boldsymbol{\Psi}$, and the object geometry represented by the surface normals $\boldsymbol{N}$.

3.1 The Image Formation Process in an Ideal World

The imaged appearance $\boldsymbol{I}_x$ at a given pixel x depends on the surface orientation at that location $\boldsymbol{N}_x$. This orientation determines the hemisphere of light that is visible at that location; the material $\boldsymbol{\Psi}$ then determines how this light is integrated to form the observed appearance. Taken together, the lighting environment and the material form a reflectance map $\boldsymbol{R}_\Psi$ that gives the appearance (i.e. color vector) for a given surface orientation.

Under ideal circumstances $\boldsymbol{R}_{\Psi}$ will have maximal entropy, that is, each input normal will yield a unique appearance. In such a case, one could recover the orientation at each pixel by exhaustively searching the reflectance map [8],

$$\boldsymbol{N}_x^{\star} = \underset{\boldsymbol{N}_x}{\operatorname{argmin}} \left[\boldsymbol{I}_x - R_{\Psi}(\boldsymbol{N}_x) \right] \ . \tag{1}$$

When the reflectance map is parametric, as is the case for Lambertian materials in typical lighting environments, such a search can be done using a gradient descent method [9, 11]. In our case, the entropy of the reflectance map varies based on the illumination environment and the reflectance making some regions more difficult to estimate than others.

3.2 Probabilistic Shape and Reflectance Estimation

Knowing the appearance and the illumination environment are not enough to recover either the reflectance, or the object geometry on their own; each requires the other to be estimated. This tight-knit relationship requires that we estimate both factors jointly. To do so, we derive a probabilistic formulation similar to a Factorial Markov Random Field (FMRF) as first introduced by Kim and Zabih [12]. In our formulation, the reflectance and the object geometry are represented as latent variables that are linked through their joint contribution on the observed appearance layer. By adopting a probabilistic framework, we enable the use of priors on both factors to incorporate basic observations, and help reduce the search space. The posterior, which we seek to optimize, is then

$$p(\boldsymbol{N}, \boldsymbol{\Psi} | \boldsymbol{I}, \boldsymbol{L}) \propto p(\boldsymbol{I} | \boldsymbol{N}, \boldsymbol{\Psi}, \boldsymbol{L}) p(\boldsymbol{N}) p(\boldsymbol{\Psi}) \ , \tag{2}$$

where we have assumed a uniform prior on the appearance.

3.3 Nonparametric Image Formation Likelihood

We assume the noise inherent in the image $\boldsymbol{I}$ is normally distributed with a common variance σ^2

$$p(\boldsymbol{I} | \boldsymbol{N}, \boldsymbol{M}) = \prod_{x \in \Omega} \mathcal{N}\Big(d\big(\boldsymbol{I}_x, \boldsymbol{R}_{\boldsymbol{\Psi}}(\boldsymbol{N}_x, \boldsymbol{L})\big), \sigma^2 \Big) \ , \tag{3}$$

where Ω is the set of observed pixels for the object and $d(\cdot, \cdot)$ measures the difference between the observed and predicted appearance as

$$d(\boldsymbol{I}_x, \boldsymbol{R}_x) = \| \boldsymbol{I}_x - \boldsymbol{R}_x \|^2 + \left[1 - \frac{\boldsymbol{I}_x \cdot \boldsymbol{R}_x}{\| \boldsymbol{I}_x \| \| \boldsymbol{R}_x \|} \right]^2 \ , \tag{4}$$

where $\| \cdot \|$ where is the Euclidean norm. This term combines the intensity difference, and color difference (computed as a normalized cosine distance), which helps mitigate the effect of inaccurate material estimates.

To model the reflectance function, we adopt the Directional Statistics Bidirectional Reflectance Distribution Function (DSBRF) model, introduced by Nishino [15], which

models a wide range of isotropic BRDFs while remaining amenable to a strong prior. Assuming a linear camera, the irradiance $\mathbf{R}(\boldsymbol{N}_x, \boldsymbol{L})$ is computed as the reflectance radiance by integrating the incident irradiance modulated by the reflectance over the illumination map

$$\mathbf{R}_{\boldsymbol{\Psi}}(\boldsymbol{N}_x, \mathbf{L}) = \int \varrho(t(\omega_i, \omega_o); \boldsymbol{\Psi}) \mathbf{L}(\omega_i) \max(0, \mathbf{N}_{\mathbf{x}} \cdot \omega_i) d\omega_i, \tag{5}$$

where t is a function which transforms ω_i and ω_o into the alternate BRDF parameterization variables θ_d and θ_h.

The DSBRDF models the reflectance function as a sum of reflectance lobes

$$\varrho^{(\lambda)}(\theta_d, \theta_h; \kappa^{(\lambda)}, \gamma^{(\lambda)}) = \sum_r \exp\left[\kappa^{(r,\lambda)}(\theta_d) \cos^{\gamma^{(r,\lambda)}(\theta_d)}(\theta_h)\right] - 1\,, \tag{6}$$

where the halfway vector parameterization (i.e., (θ_h, ϕ_h) for the halfway vector and (θ_d, ϕ_d) for the difference vector) [17] is used, $\kappa^{(\lambda)}$ and $\gamma^{(\lambda)}$ are functions that encode the magnitude and acuteness of the reflectance, respectively, of lobe r along the span of θ_d for a particular color channel λ. This allows us to express reflectance functions as a collection of κ and γ curves.

As shown by Nishino and Lombardi [15, 16], these curves can be modeled as a log-linear combination of data-driven basis functions,

$$\kappa^{(r,\lambda)}(\theta_d) = \exp\left[b_\mu(\theta_d; \kappa, r, \lambda) + \sum_i \psi_i b_i(\theta_d; \kappa, r, \lambda)\right], \tag{7}$$

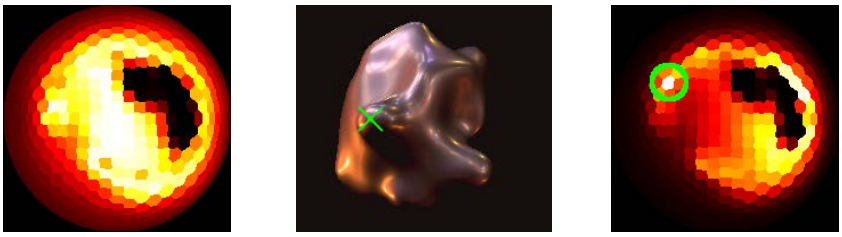
$$\gamma^{(r,\lambda)}(\theta_d) = \exp\left[b_\mu(\theta_d; \gamma, r, \lambda) + \sum_i \psi_i b_i(\theta_d; \gamma, r, \lambda)\right], \tag{8}$$

where b_μ is the mean basis function, b_i is the i^{th} basis function, and ψ_i are the DSBRDF coefficients. We may compute these basis functions from a set of measured reflectance functions using functional principal component analysis (FPCA). This provides the additional benefit of ordering the basis functions by importance so that we can truncate the number of parameters used to achieve a compact representation.

Figure 1 illustrates how a single reflectance map $\boldsymbol{R}_{\boldsymbol{\Psi}}(\boldsymbol{N}, \boldsymbol{L})$ may have various degrees of certainty. The hemispherical distribution in Figure 1a is the likelihood distribution for the pixel indicated with the green X. In these images, brighter values correspond to more likely orientations. Because $\boldsymbol{R}_{\boldsymbol{\Psi}}$ is also a function of the unstructured illumination environment $\boldsymbol{L}$, this distribution is multimodal and nonparametric. In other words, for this combination of illumination environment and reflectance, the shape recovery problem is ill-posed.

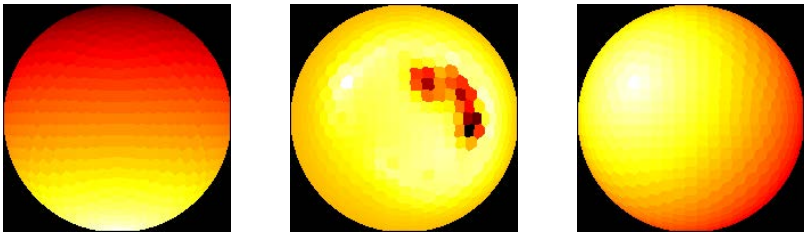
3.4 Orientation Constraints

Figure 1c shows the likelihood distribution for the pixel just above the one denoted with the X. As can be seen by the single bright spot in the distribution, the appearance of this pixel, which lies in a purple highlight, can only be explained by a small range of orientations. This occurrence, though rare, gives valuable context to the area surrounding the



(a) Underdetirmined likelihood (b) Image (c) Well determined likelihood

Fig. 1. Nonparametric likelihood. In this example, there are many possible orientations that may result in the appearance for the pixel denoted with an X. This can be seen in the hemispherical distribution of (b) which has a large bright area corresponding to the likely orientations. The purple pixel just above this one, however, has relatively few likely orientations (c).



(a) Boundary prior (b) Gradient prior (c) Smoothness prior

Fig. 2. Orientation prior examples. In these hemispherical distributions, brighter pixels correspond to more likely orientations. The occluding boundary prior (a) strongly encourages pixels on the boundary to be oriented orthogonally to the viewing direction. The gradient prior (b) encourages the estimated appearance to have the same gradient as the observed image. The smoothness prior (c) penalizes sharp changes in orientation.

well-determined pixel. We introduce two spatial priors to help propagate this information. We also utilize the occluding boundary as a strong unary prior p_b on pixels at the edge of the object. The spatial priors p_g and p_s are formulated pairwise,

$$p(\boldsymbol{N}) = \prod_{x \in \Omega} p_b(\boldsymbol{N}_x) \prod_{y \in \mathrm{ne}(x)} p_g(\boldsymbol{N}_x, \boldsymbol{N}_y) p_s(\boldsymbol{N}_x, \boldsymbol{N}_y) \,. \tag{9}$$

Figure 2 illustrates the role of each of these priors, which we will discuss next.

Occluding Boundary Prior. As first explored by Ikeuchi and Horn [10], surface patches on the occluding boundary of smooth objects must be oriented orthogonally to the viewing direction. As shown in Figure 2a, since the area subtended by each boundary pixel may include a range of orientations, not all of which will be orthogonal to the viewing direction, we formulate this prior as a tight, but non-singular, distribution,

$$p_b(\boldsymbol{N}_x) = \begin{cases} \mathcal{N}(\arccos^2(\boldsymbol{N}_x \cdot \boldsymbol{B}_x), 0.1) & \text{if } x \in \boldsymbol{B} \\ 1 & \text{otherwise} \,, \end{cases} \tag{10}$$

where $\boldsymbol{B}$ is a list of the orthogonal vectors for pixels on the occluding boundary of the object.

Reflected Gradient Prior. Our first spatial prior ensures that the resulting gradient is the same as the observed image gradient

$$p_g(\boldsymbol{N}_x, \boldsymbol{N}_y) = \exp -\Big(\|\boldsymbol{R}_\Psi(\boldsymbol{N}_x) - \boldsymbol{R}_\Psi(\boldsymbol{N}_y)\| - \|\boldsymbol{I}_x - \boldsymbol{I}_y\| \Big)^2 , \tag{11}$$

where $\| \cdot \|$ is the Euclidean norm. An example of this prior is shown in Figure 2b. Here we see how the well-determined purple pixel of Figure 1c influences the area surrounding it.

Smoothness Prior. Although the gradient prior gives sufficient context to pixels in textured areas, in many cases, the reflectance map contains many areas with a similar appearance. Our smoothness prior is designed to allow small changes in orientation (angles less than $\pi/3$) while strongly penalizing sharp changes in orientation. To do so, we formulate this as a logistic function,

$$p_s(\boldsymbol{N}_x, \boldsymbol{N}_y) = \Big(1 + \exp\big(-s(\cos^{-1}(\boldsymbol{N}_x \cdot \boldsymbol{N}_y) + t) \Big)^{-1} , \tag{12}$$

where $t = \pi/3$ is the threshold between weak and strong penalization and $s = 10$ is the speed of the transition. When well-determined points are present in the scene, this prior helps ensure that the geometry in the intermediate area varies smoothly.

3.5 Reflectance Constraints

We utilize a previously developed prior by Lombardi and Nishino [13] to model the space of the $\boldsymbol{\Psi}$, the coefficients which control the influence of the basis functions. By fitting a set of DSBRDF coefficients to measured data, we can model their variation. We adopt the same zero-mean multivariate Gaussian prior to model the space of reflectance,

$$p(\boldsymbol{\Psi}) \sim \mathcal{N}(0, \Sigma_\Psi), \tag{13}$$

where Σ_Ψ is computed from the MERL BRDF database [14]. Note that we exclude the BRDF we try to estimate when conducting the synthetic experimental validation.

4 Optimization and Refinement

Since neither the reflectance nor the geometry can be optimized without knowledge of the other, we employ an iterative, expectation maximization (EM) framework in which one is held fixed as a "pseudo-observable" while the other is estimated in the maximization step, while the expectation step estimates the Gaussian likelihood variance.

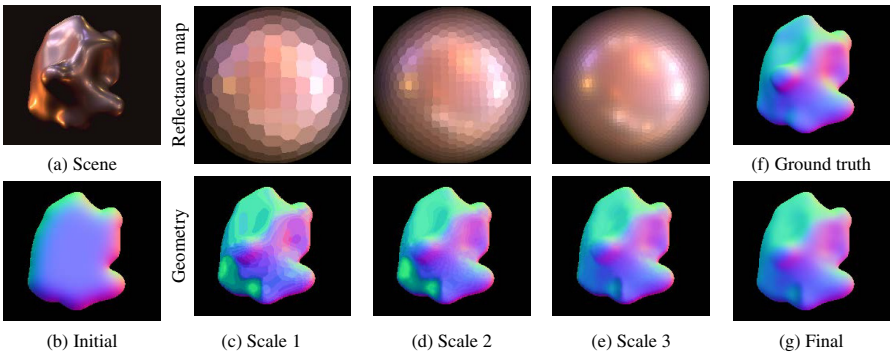


Fig. 3. Multi-scale geometry estimation. We partition the domain of possible surface orientations into a geodesic hemisphere. By incrementally increasing the resolution of this hemisphere we are able to avoid local minima. Our final step is a gradient-descent based minimization to fine-tune the result.

4.1 Initial Estimation

We begin our optimization with a naive geometry estimation. Without an estimate for the reflectance, the likelihood of Equation 3 and the reflectance gradient prior of Equation 11 are meaningless. We may still, however, use the smoothness and occluding boundary priors. As shown in Figure 3b, the result is a normal field that resembles a soap-bubble. Although this is a crude estimate of the object geometry, it serves as a reasonable starting point and allows us to bootstrap the iterative EM optimization.

4.2 Directional Statistics Material Estimation

The Directional Statistics BRDF (DSBRDF) model [16], describes an isotropic BRDF as a mixture of exponential distributions on the unit hemisphere. By doing so, the full range of isotropic BRDFs is well described using few parameters.

In recent work by Lombardi and Nishino [13], a straight-forward optimization method is presented to recover the material parameters under a single point light source. We have extended this framework by treating the illumination map as a dense set of point light sources. This extra information allows for rapid material estimation, the quality of which depends on the accuracy of the orientation information.

4.3 Multi-scale Geometry Estimation

As we have noted, the primary challenge in this problem stems from the non-parametric nature of the reflectance map. Even if a pixel has a unique appearance, the reflectance map is rarely smooth enough that a simple iterative optimization scheme will find this optimal value. To address this, we introduce a novel global optimization algorithm. As shown in Figure 3, we limit the space of possible surface orientations to a set of orientations on the geodesic hemisphere.

There are several benefits to this formulation. First, by limiting the surface orientations to a finite set, we enable the use of rapid global optimization approximation methods. And second, the effect of an inaccurate reflectance estimation is minimized. An inaccurate reflectance estimate may cause the likelihood of Equation 3 to have a unique minimum that is incorrect. By limiting the possible orientations, we force each pixel to align with the *region* it is most similar to, while obeying the constraints of our priors. The size of these regions is gradually decreased allowing the priors to play a more subtle role at finer scales.

For a given material estimate, the surface orientations are estimated at each of three scales by incrementally dividing the geodesic hemisphere. To approximate the optimal set of orientations, we use the alpha-beta swap method of Boykov et al. [4]. To help interpolate from one scale to the next, the orientation estimate from the previous scale is first smoothed with a Gaussian kernel before being re-discretized at the next finer resolution. The progression of this optimization scheme is illustrated in Figure 3.

4.4 Refinement and Integrability

At this point we have a good estimate for the object geometry. In order to ensure that the underlying surface in integrable, and to refine the still somewhat coarse orientation estimate, we apply a gradient descent optimization. Similar to Ikeuchi and Horn [10], we perform this optimization in gradient space. The primary difference in our work is due to the DSBRDF. The final shape estimation for our running example is shown in Figure 3g. Note that this method is only possible because we already have a good estimate through discrete global optimization.

5 Experimental Validation

5.1 Synthetic Images

As illustrated in Figure 4 we evaluated our results on 600 synthetic images. Each row of the figure shows 5 of the 10 different shapes from the Blobby Shapes database [11]. Each of these 10 shapes is rendered under 6 different publicly-available real-world illumination environments [5] with each of 10 different real-world measured BRDFs which were chosen from the MERL database [14] to span a wide variety of reflectances. Objects rendered with the same reflectance are shown along each numbered diagonal. The diversity of appearance across each row, and along each diagonal shows how dramatically appearance varies due to changes in material and illumination.

We compute the DSBRDF prior with the ground-truth BRDF omitted. In other words, the ground-truth BRDF is not a part of the training data for the object being analyzed. In total we achieve a median angular error of 24°. The relative root-mean-squared error for the recovered BRDFs ranges from 0.54 to 3.02 though the median value is 1.1. In order to directly compare with the work of Johnson and Adelson [11], we have also briefly tested our method assuming the ground-truth reflectance map is known. In this case we achieve an median angular errors consistently below 15°. This shows the strength of our model to overcome the inherent ambiguities of a non-parametric reflectance map.

Fig. 4. Sampling of 600 synthetic images. The bulk of our synthetic data are formed by rendering 10 shapes (5 shown per row) [11] with 10 real-world BRDFs (numbered diagonals) [14] under 6 real-world natural illumination environments (one per row) [5].

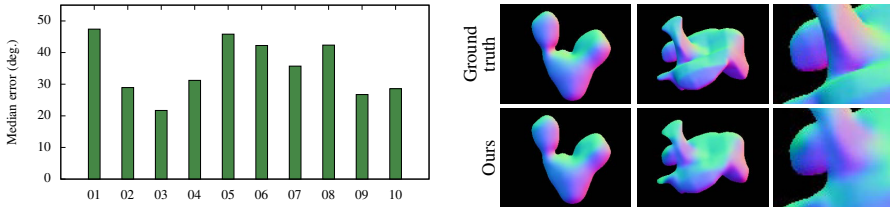
The increased error when the reflectance is not known is due to the added challenge. We will focus the rest of this discussion on that case.

As shown in Figure 5, some shapes are more challenging than others to estimate. In (a) we graph the median angular error, in degrees for each of the 10 shapes averaged over the 60 different lighting and reflectance combinations. In (c) we show the most challenging shape. Note that this shape has several self-occlusions which violate our assumptions of integrability and smoothness. In (b) we show a shape that is easier to estimate. Taken together, these 10 shapes constitute a range of characteristics. For the remainder of our evaluation we will present averages over all 10 shapes.

Figure 6 illustrates the accuracy of our geometry recovery for each of the 60 different illumination environment and reflectance combinations averaged over the 10 different shapes. The heat map (a) shows the median angular error with brighter colors used to indicate increased accuracy. In this figure the rows and columns correspond to the labels in Figure 4. Several examples are included to illustrate the results.

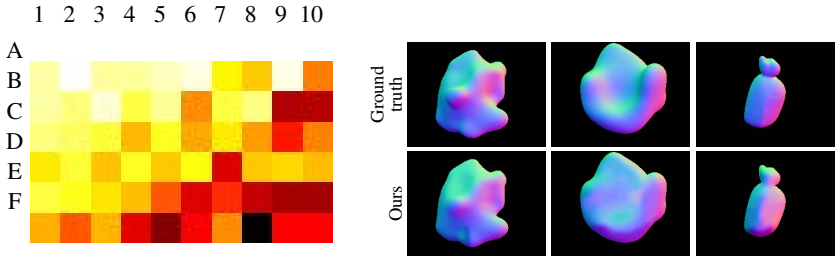
Some reflectances, such as GRAY PLASTIC (3) and SPECIAL-WALNUT (2) consistently yield more accurate results. These reflectances, which are the most matte, are more successful when large regions of the lighting environment differ strongly from each other. The notable exceptions to this occur in the eucalyptus forest lighting environment (F) and in the St. Peter's environment (E). These two lighting environments have high-frequency texture whose details become nearly uniform under these diffuse reflectances. To test the limits of our method we chose an anisotropic material—WHITE FABRIC (5). As expected, the reflectance estimates and consequently, the geometry estimates for this reflectance are consistently lower than the others.

The outdoor scene from the city of Pisa (B) is shown to facilitate reflectance and geometry estimation. This is likely due to the smooth gradients present in the scene.



(a) Geometry error for each blob (b) BLOB03 (c) BLOB5 (d) Detail

Fig. 5. Shape accuracy. Shapes with self occlusions, like BLOB05 (c) and BLOB01 present challenges to our algorithm, while simple shapes like BLOB03 (b) are more reliably estimated



(a) Surface normal accuracy (b) B3 Detail (c) D6 Detail (d) F10 Detail

Fig. 6. Geometry recovery. The rows and columns of (a) correspond to the labeled illumination maps and reflectances (respectively) in Figure 4. The brightest value corresponds to the reflectance and lighting environment combination that yielded the most accurate geometry estimation, with a median error of 14°. The darkest value corresponds to a median error of 49°.

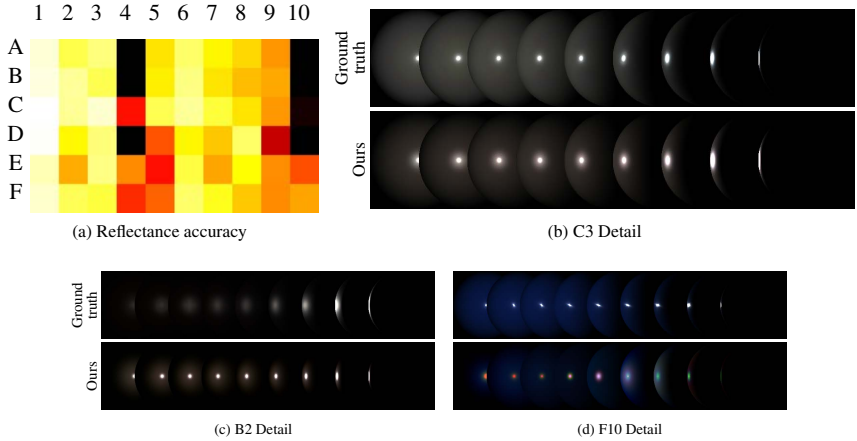
The result is that, whether specular or diffuse, the reflected appearance is integrating a unique hemisphere of light. On the other hand, environments such as St. Peter's (E) and the Ennis-Brown House (C) have repeated patterns that are less distinctive.

Figure 7 illustrates the accuracy of the reflectance estimation. In (b) we show an estimate for GRAY-PLASTIC (3). This reflectance was among the most reliably estimated, due to the smoothly varying appearance resulting from its matte finish. In (d) we show an estimate for BLUE-ACRYLIC (10). This material has subsurface scattering, which our model cannot explain well, causing the global illumination effect to be absorbed in erroneous surface normal and reflectance estimates.

5.2 Real-World Images

As shown in Figure 8, we have also acquired images, and aligned ground-truth geometry for several objects in both out-door and in-door real-world scenes[1]. The ground-truth geometry was acquired using a Canon VIVID 910 light-stripe range-finder. Illumination environments were acquired using a reflective sphere. Although real-world data comes

[1] This data set is available at `http://www.cs.drexel.edu/~kon/natgeom`



(a) Reflectance accuracy

(b) C3 Detail

(c) B2 Detail

(d) F10 Detail

Fig. 7. Reflectance accuracy. The rows and columns of (a) correspond to the labeled illumination maps and reflectances (respectively) in Figure 4. The brightest value corresponds to the most accurate reflectance estimation with a relative RMS error of 0.54. The darkest regions correspond to an error of greater than 3.0.

with added sources of noise and inaccuracy, our method is able to recover the shape of the objects quite well in each of the environments.

In each of the four sections of the figure we show the illumination environment along with the image, recovered normal field and ground-truth normals for each of the objects in the scene. The top left section shows a bear figure and a horse figure in a large lobby. The dominant illumination from the overhead skylights is relatively uniform. This leads to some inaccuracy along the top of the objects. The walls of the lobby, however, contain many different points of illumination which contribute to the smoothly varying geometry recovered along the middle of the objects. In this scene the bear has a mean angular error of $24°$ and the horse has a mean angular error of $17°$.

The garden scene in the top right is illuminated primarily by a cloudy sky. This smoothly varying illumination gives rise to more accurate estimations in the top portions of the objects. In the picnic area scene of the lower left we have a similar environment. This one, however, contains flowers and has more texture in the grass at the bottom of the scene. These two features give extra orientation clues that the garden scene does not have. The effect can be seen most dramatically in the cheek of the bear. Since the lower portion of the garden scene is essentially uniform, points oriented to the left and right have very similar reflected appearances. The objects in the garden scene have mean errors of $19°$ and $26°$ whereas the objects in the picnic area scene both have mean angular errors of $17°$.

In the entrance hall scene (bottom right) we see the most accuracy along the top of the objects. This is due to variation in the overhead lighting. On the other hand, the pattern in the walls is relatively uniform throughout, giving rise to less accurate detail around the middle of the objects. In this scene the mean error for the objects are $20°$ for the shell and $16°$ for the milk bottle.

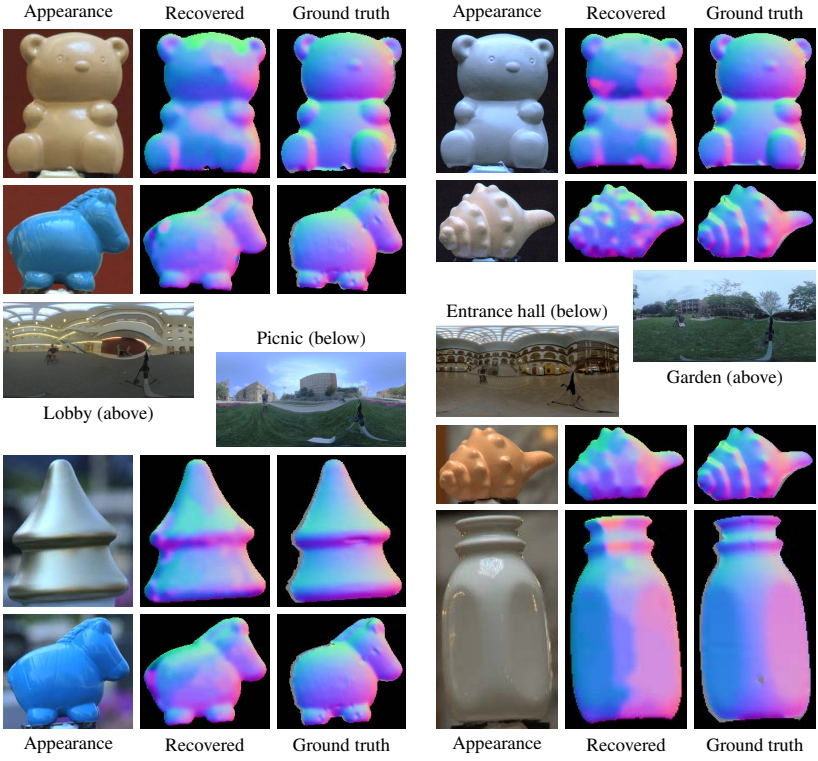


Fig. 8. Real-world results. We captured several objects in four different natural lighting environments, and aligned ground-truth normal maps. As we discuss in the text, differences in the lighting environments have a clear impact on the accuracy of the recovered geometry.

6 Conclusion

In this work we introduced a method to jointly estimate the reflectance and geometry of an object in natural illumination. We have shown how the BRDF of the object and the illumination environment together determine the accuracy of such scene analysis. The accuracy of our method shows that problems previously unexplored involving arbitrary reflectance and unstructured natural illumination become tractable when the rich contextual information embedded in the lighting is exploited. We believe this opens the way to bringing the success of methods traditionally confined to the laboratory out into the real world.

Acknowledgments. The authors are grateful to Stephen Lombardi and Gabriel Schwartz for their important technical contributions. This work was supported in part by the Office of Naval Research grant N00014-11-1-0099, and the National Science Foundation awards IIS-0746717 and IIS-0964420.

References

1. Alldrin, N.G., Kriegman, D.J.: Toward Reconstructing Surfaces With Arbitrary Isotropic Reflectance: A Stratified Photometric Stereo Approach. In: IEEE Int'l Conf. on Computer Vision, pp. 1–8 (2007)
2. Alldrin, N.G., Zickler, T., Kriegman, D.J.: Photometric stereo with non-parametric and spatially-varying reflectance. In: IEEE Int'l Conf. on Computer Vision and Pattern Recognition, pp. 1–8 (2008)
3. Basri, R., Jacobs, D., Kemelmacher, I.: Photometric Stereo with General, Unknown Lighting. Int'l Journal of Computer Vision 72(3), 239–257 (2006)
4. Boykov, Y., Veksler, O., Zabih, R.: Fast Approximate Energy Minimization via Graph Cuts. IEEE Trans. on Pattern Analysis and Machine Intelligence 23(11), 1222–1239 (2001)
5. Debevec, P.: Light Probe Image Gallery (2012), `http://www.pauldebevec.com/Probes/`
6. Durou, J.-D., Falcone, M., Sagona, M.: Numerical Methods for Shape-from-Shading: A New Survey with Benchmarks. Computer Vision and Image Understanding 109(1), 22–43 (2008)
7. Goldman, D.B., Curless, B., Hertzmann, A., Seitz, S.M.: Shape and Spatially-Varying BRDFs from Photometric Stereo. IEEE Trans. on Pattern Analysis and Machine Intelligence 32(6), 1060–1071 (2010)
8. Hertzmann, A., Seitz, S.M.: Example-Based Photometric Stereo: Shape Reconstruction with General, Varying BRDFs. IEEE Trans. on Pattern Analysis and Machine Intelligence 27(8), 1254–1264 (2005)
9. Huang, R., Smith, W.: Shape-from-Shading Under Complex Natural Illumination. In: IEEE Int'l Conf. on Image Processing, pp. 13–16 (2011)
10. Ikeuchi, K., Horn, B.K.P.: Numerical Shape from Shading and Occluding Boundaries. Artificial Intelligence (1981)
11. Johnson, M.K., Adelson, E.H.: Shape Estimation in Natural Illumination. In: IEEE Int'l Conf. on Computer Vision and Pattern Recognition, pp. 1–8 (2011)
12. Kim, J., Zabih, R.: Factorial Markov Random Fields. In: Heyden, A., Sparr, G., Nielsen, M., Johansen, P. (eds.) ECCV 2002, Part III. LNCS, vol. 2352, pp. 321–334. Springer, Heidelberg (2002)
13. Lombardi, S., Nishino, K.: Single Image Multimaterial Estimation. In: IEEE Int'l Conf. on Computer Vision and Pattern Recognition, pp. 238–245 (2011)
14. Matusik, W., Pfister, H., Brand, M., McMillan, L.: A data-driven reflectance model. ACM Trans. on Graphics 22(3), 759–769 (2003)
15. Nishino, K.: Directional Statistics BRDF Model. In: IEEE Int'l Conf. on Computer Vision, pp. 476–483 (2009)
16. Nishino, K., Lombardi, S.: Directional Statistics-based Reflectance Model for Isotropic Bidirectional Reflectance Distribution Functions. Journal of Optical Society America, A 28(1), 8–18 (2011)
17. Rusinkiewicz, S.: A New Change of Variables for Efficient BRDF Representation. In: Eurographics Workshop on Rendering, pp. 11–22 (1998)
18. Woodham, R.J.: Photometric Method for Determining Surface Orientation from Multiple Images, vol. 19. MIT Press (1989)
19. Zhang, R., Tsai, P.-S., Cryer, J.E., Shah, M.: Shape from Shading: A Survey. IEEE Trans. on Pattern Analysis and Machine Intelligence 21(8), 690–706 (1999)