Eurographics Symposium on Rendering (2006), pp. 1–10
Tomas Akenine-Möller and Wolfgang Heidrich (Editors)

Sparse Lumigraph Relighting by Illumination and Reflectance Estimation from Multi-View Images

Abstract

We present a novel relighting approach that does not assume that illumination is known or controllable. Instead, we estimate the illumination and texture from given multi-view images captured under a single illumination setting, given object shape. We rely on viewpoint-dependence of surface reflectance to resolve the usual texture-illumination ambiguity. The task of obtaining illumination and texture models is formulated as the decomposition of the observed surface radiance tensor into the product of a light transport tensor, and illumination and texture matrices. We estimate both illumination and texture at the same time by solving a system of bilinear equations. To reduce estimation error due to imperfect input surface geometry, we also perform a multi-scale discrete search on the specular surface normal. Our results on synthetic and real data indicate that we can estimate both illumination and the diffuse as well as specular components of the surface texture map (up to a global scaling ambiguity). Our approach allows more flexibility in rendering novel images, such as view changing, and light and texture editing.

1. Introduction

The images of a scene under varying illuminations and from different viewpoints are highly interrelated, which makes it possible to predict the object's appearance from new viewpoints or under different illuminations. To achieve this, image based rendering and image based re-lighting techniques (e.g. [LH96, GGSC96, WGT*05, DHT*00]) usually require a dense sampling of viewpoints and/or illumination directions. Image based modeling methods (e.g. [YDMH99, LKG*03, RH01, NZI01, NZI05]), on the other hand, attempt to estimate parametric models of the scene appearance. Though these methods require less input data, simultaneous estimation of illumination and texture maps from known geometry and multi-view images is believed to be ambiguous, as noted in [RH01]. Two commonly adopted solutions to this problem are to assume that illumination is known or to control it suitably [HS05, LKG*03, YDMH99]. In real applications, e.g. in outdoor scenes, neither solution is feasible.

In this paper, we point out that the ambiguity in estimating texture and illumination from multi-view images exists only for diffuse material. By exploiting the viewpoint-dependent portion of the object's surface reflectance, one can achieve the effect of varying illumination by just varying viewpoints. We show that by separating the specular radiance from the object appearance and using a parametric specular reflectance model, we can automatically estimate the environment illumination, surface diffuse/specular albedo map, and bump (surface normal) map from multi-view images taken under a single, unknown illumination setting.

Figure 1: *By analyzing multi-view images of the fish, and the Van Gogh statue, captured under fixed unknown illumination (left), we are able to extract the diffuse/specular texture of these objects as well as the environmental illumination map. These results help us perform advanced scene manipulations such as inserting the fish into the Van Gogh scene (right).*

The major contributions of our paper are:

1. We propose a tensor factorization framework where the illumination and spatially varying textures can be solved via bilinear factorization given the object shape and the basis BRDFs.
2. We show that observations of the specular component from multiple viewpoints suffice to solve for both illumination and texture map (up to a global scaling ambiguity). Therefore, we do not need to know, measure (e.g., using a light probe) or manipulate the illumination.
3. We propose a normal map search method during the illumination and texture estimation to handle the imperfect geometric input.
4. We demonstrate the application of our results in different rendering tasks such as view change, illumination and texture editing.

This paper is organized as follows. Sec. 2 gives a review of some related work. Sec. 3 formulates the estimation problem as bilinear tensor factorization. Sec. 4 presents our estimation algorithm. Sec. 5 gives experimental results on synthetic and real datasets and the applications of the estimated results in various rendering tasks. Sec. 6 presents the conclusions and future work.

2. Previous Work

Image Based Rendering (IBR)/Re-Lighting (IBRL). Light field rendering (LFR)/lumigraphh [LH96, GGSC96] have made IBR a great success in computer graphics due to their simplicity of formulation and ability to create photo-realistic images. Later improvements such as Eigen-texture [NSI99], Surface Light Field Compression [WAA*00] and Light Field Mapping (LFM) [CBCG02] use more accurate geometric information and more efficient compression techniques. These methods can create images from novel views. But unlike the method we present in this paper, it is not possible to modify the texture or illumination of the scene. Image based relighting [WGT*05, DHT*00], on the other hand, uses images captured under different illuminations to synthesize images under novel lighting. But the requirement of illumination control is restrictive and not always possible. Furthermore, most of these methods use linear interpolation to generate novel images and therefore need very dense sampling of viewpoints and/or illumination directions. By exploiting both texture and illumination coherency between different viewpoints, our work only need sparse input images for synthesis, in contrast to the Lumigraph [GGSC96] where only texture coherency is considered.

Data Driven BRDF/BTF Modeling. Data driven models have also been used in reflectance modeling of surface with complex textures. Matusik et al. use densely acquired reflectance data to estimate the Bidirectional Reflectance Distribution Function (BRDF) [MPBM03]. Tensor factorization methods have been used for 6-Dimensional Bidirectional Texture Function (BTF) modeling [VT04, WWS*05]. These models serve as a compact representation of densely sampled images for efficient rendering. An image under novel illumination and from a new viewpoint can be generated directly by interpolation in the compact subspace. In these methods, illumination is both known and controlled, and it is very difficult to modify the texture after the capturing.

Parametric Reflectance Modeling under Controlled Illumination. Instead of massive data driven models, parametric models are also used for surface reflectance modeling. One major method is based on the manipulation of illumination. Lensch et al. [LKG*03] extract the spatially varying BRDF parameters of a surface with known geometry and controlled illumination. The extracted BRDF parameters are then used to refine the surface normals. Using a similar setup, photometric stereo methods [HS05, Geo03] also try to estimate the surface reflectance along with shape/normal estimation. The effects of changing illumination can also be achieved by rotating the object for fixed lighting and viewpoint [SWI97, ZCHS03]. By controlling the illumination settings, the texture-illumination ambiguity can be removed even for a purely diffuse object. However, compared to our method, these methods are restricted to a highly controlled environment, and they can only deal with static objects unless very special hardware is used [WGT*05].

Parametric Reflectance Modeling from Multiple Views. Instead of varying illumination, one can also vary the viewpoint to solve the same problem. The advantage of using such a multi-view method is that it does not require the control of lighting. Yu et al. [YDMH99] estimate the diffuse albedo map and a piecewise constant specular map from known geometry and point light sources using an inverse global illumination model. Ramamoorthi and Hanrahan [RH01] propose a signal-processing framework and discuss the feasibility of solving various photometric reconstruction problems in a multi-view setup. They focus mainly on objects with homogeneous BRDFs. To overcome the illumination-texture ambiguity, they also assume knowledge of illumination when solving for spatially varying surface texture maps. Nishino et al. [NZI01, NZI05] solve the illumination, a homogeneous specular BRDF and a diffuse texture map by using blind deconvolution on the specular component. Although their method can also synthesize relighted scene based on the images under a single illumination setting, they assumes homogeneous specular BRDF and restrict light to be point sources. In contrast, our method focused on resolving the illumination-texture ambiguity where a spatially varying specular albedo map can be recovered together with an arbitrary illumination map represented in frequency domain.

Illumination Estimation. Mirror sphere is the standard tool to acquire illumination. Besides that, many methods use an object with uniform or piecewise constant albedo. These

include the methods that use diffuse and specular reflection to estimate multiple directional sources [ZK02], and spherical harmonic decomposition of the illumination [RH01]. Intensity variations due to shadows of a known shape on a uniform or known textured surface are also used to estimate illumination [SSI03, LLLS03]. Our method, on the other hand, can work on objects with arbitrary, unknown texture, as long as there are view-dependent surface radiance variations.

Precomputed Radiance Transfer. PRT, which precomputes most of the non-linear light transport so that fast rendering is possible, is also closely related to our work [SKS02, SLS05, NRH04]. Our method uses a decomposition of the surface radiance similar to the one used in PRT, but our goal is the reverse process, namely to estimate the reflectance and illumination from the images.

3. Problem Formulation: Bilinear Decomposition of Surface Radiance Tensor

In this section, we present the basic formulation of the problem, and explain its relationship to the existing other methods. The observed object appearance is the combined effects of environment illumination, surface geometry, BRDF and texture. Given the surface geometry, the radiance of the object surface from different viewpoints and under different illumination settings can be arranged into a 3D tensor $\mathcal{R} \in \Re^{P \times J \times L}$, where the three dimensions, P, J, L correspond to the surface location, viewpoint and illumination setting †. We formulate the illumination and texture estimation problem as the decomposition of the observed radiance into three components, corresponding to the illumination, light transport (the joint effect of basis BRDFs and surface geometry), and the texture albedo.

3.1. Radiance Tensor Formation Model

Similar to the work on pre-computed radiance transfer [SKS02, SLS05, NRH04], we divide the transformation from incident light to the surface radiance $\mathcal{R}$ into linear and non-linear parts (Fig. 2). Linear components of the incident light first pass through different non-linear light transport processes. The results are linearly weighted by the surface texture to obtain the surface radiance. Mathematically, $\mathcal{R}$ can be written as a product of three tensors:

$$\mathcal{R} = \mathcal{A} \times_1 \mathbf{B} \times_2 \mathbf{H} \tag{1}$$

where $\mathbf{B}^{P \times P}$ is the albedo matrix, and $\mathbf{H}^{L \times K}$ is the illumination matrix, and $\mathcal{A} \in \Re^{P \times J \times K}$ is the light transport tensor (LTT) ‡.

† Each of these three dimensions actually has two degrees-of-freedom. Therefore, the surface radiance tensor also represents a 6-D function that is similar to the Non-local Reflectance Field introduced in [DHT*00].

‡ For occluded surface point the corresponding entry in LTT is zero, meaning no light is transported to that viewpoint.

To simplify the problem, we assume the illumination is distant so that it can be represented as a 2D spherical function. This function can be decomposed, for example, using spherical harmonics or wavelet bases. We assume the object is composed of a base material modulated by an albedo map (texture), so $\mathcal{A}$ is fully specified by the object geometry and a base material's BRDF. We model the surface texture as formed by the triangles in the object's mesh, each having a constant albedo. The light transport tensor records the contribution of each illumination component to each surface triangle. Note the albedo matrix $\mathbf{B}$ becomes a diagonal matrix if we assume no inter-reflection between different triangles.

We can enrich our model's expressiveness by using a linear combination of basis materials. Each basis material i has its own LTT $\mathcal{A}_i$ and the corresponding diagonal albedo matrix $\mathbf{B}_i$. The radiance tensor can then be written as:

$$\mathcal{R} = \sum_i \mathcal{A}_i \times_1 \mathbf{B}_i \times_2 \mathbf{H} \tag{2}$$

Again, the light transport tensor A_i in (2) is known, given the object geometry and the basis BRDFs. The unknowns are the multiple albedo maps $\mathbf{B}_i$ and the illumination matrix $\mathbf{H}$.

In this paper, we use two basis materials, the diffuse and specular materials. We also ignore any inter-reflections between different object surfaces so that $\mathbf{B}_i$ is diagonal. The differences between our approach and that of Lensch et al. [LKG*03] in handling spatially varying materials are that the basis material is shared by the entire surface in our formulation and the spatially variations are introduced by linear albedo maps. This allows the illumination matrix and albedo matrix to be factored out simultaneously. Lensch et al. use more complex surface reflectance models and rely on clustering and illumination manipulation to solve for multiple materials.

Among the rest of the paper, Sec. 3.3 discuss the conditioning of solving both illumination and texture given $\mathcal{R}$ and the diffuse and specular LTTs. Sec. 4.1 discuss how to numerically approximate these LTTs. Note that although the diffuse BRDF has no parameters, the specular BRDF usually has additional parameters that are unknown. These parameters are estimated using another non-linear optimization as described in Sec. 4.3.

3.2. Varying Illumination v.s. Varying Viewpoints

As discussed previously, photometric methods (calibrated or uncalibrated) such as [LKG*03, HS05, Geo03] manipulate the illumination so that different transported illuminations are multiplied with the texture albedo in order to separate the two (Fig. 3 (a)). The advantage of controlled illumination is that it can resolve texture-illumination ambiguity even for pure diffuse material. However, manipulating illumination requires a highly controlled environment.

Given the Light Transport Tensor, We separate the texture

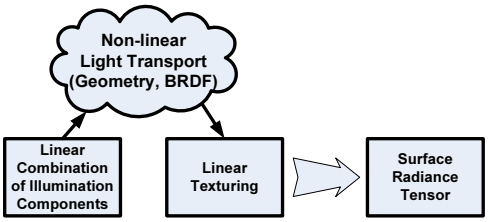


Figure 2: *The radiance tensor formation model.*

albedo and illumination by observing the object from different viewpoints. In order to obtain non-constant transported illumination as in photometric methods, the LTT should have non-constant values in the viewpoint dimension. Because of this constraint, we cannot resolve lighting and texture ambiguity for purely diffuse scenes – the well-known texture-illumination ambiguity [RH01]. However, for specular objects, we can decompose the observed surface radiance into texture albedo and illumination matrix from images of a SINGLE UNKNOWN illumination setting. This method gives us greater flexibility since the illumination no longer needs to be controlled or measured even for textured surfaces.

3.3. Bilinear Factorization

Eq.(1) blends together the two sets of unknowns using the light transport tensor $\mathcal{A}$, which, given the observed surface radiance tensor $\mathcal{R}$, results in an over-determined system of bilinear equations. There is no closed form solution to the least squares problem of such an over-determined bilinear system, but the normal equations of this system can be derived and solved iteratively [CT97]. This is equivalent to iterating between fixing one set of parameters (illumination **H** or texture albedo **B**) and solving a linear least squares problem for the other.

The stability/convergence of the bilinear factorization depends on the structure of the light transport tensor. First, there is a global scaling ambiguity between the illumination **H** and the texture albedo **B**. This can be resolved by normalizing the DC component of the illumination to 1 [§]. After the normalization, since diffuse LTT $\mathcal{A}_d$ does not vary along the viewpoint dimension , observation from different viewing directions does not bring enough independent equations to constrain the unknowns and the estimation is ill-conditioned. For purely specular materials, $\mathcal{A}_s$ has linearly independent entries along the viewpoint dimension. Observing a specular objects from different directions does give us extra equations to solve both texture albedo and illumination. For an object with mixed diffuse and specular reflectance, one can in principle solve for the diffuse and specular albedo, as well as the illumination using the same iterative approach (we will refer to this as the direct method). But in practice, because of the sparse sampling of viewpoints, we will show that it is more stable to remove the under-constrained diffuse component before solving the bilinear system. Diffuse texture can be solved for later after illumination has been estimated.

[§] Our paper handles each color channel separately, so this normalization sets the average color of estimated illumination to white

4. Illumination and Texture Estimation Algorithm

Based on the above formulation, we propose a multi-view texture map and illumination estimation algorithm. A schematic of our algorithm is shown in Fig. 4 (a). Our input is a set of calibrated images of the object and its 3D shape. We first separate the diffuse and specular component of each surface triangle. The diffuse color is computed as the median of its radiance observed from different viewpoints[¶]. It is subtracted from the surface radiance tensor to obtain the specular radiance tensor. Based on the object geometry and the basis BRDF, we are able to compute the specular LTT. The illumination and specular albedo map are first solved for in the specular component analysis. We also perform special non-linear optimization to adjust the specular bump map and specular BRDF parameter σ. The estimated illumination is then used to solve for the diffuse albedo map.

4.1. Compute the Light Transport Tensor

The light transport tensor encodes all the non-linear factors in the light transport. Significant effort has been made to find fast and accurate ways to compute the light transport tensor [SKS02, SLS05, NRH04]. We use spherical harmonics to decompose the environment illumination, which gives us a compact band limited representation. The spherical illumination $L(\Theta)$ can be written as a linear combination of spherical harmonic basis:

$$L(\Theta) = \sum_{m=0}^{N_H} \sum_{n=-m}^{m} h_{mn} Y_{mn}(\Theta) \qquad (3)$$

where N_H is the highest order of harmonic series we use, h_{mn} are the spherical harmonic coefficients and Y_{mn} are the real basis functions [Wei]. To simplify notation we will identify the basis functions using a single index k, so $L(\Theta) = \sum_{k=0}^{K} h_k Y_k$.

[¶] The reason for choosing median value instead of minimum value as the diffuse component is due to the geometric error of the real dataset. When projected to the input images, a triangle might be mapped to the radiances from different parts of the object in different views. The minimum values of these radiances sometimes contains artifacts and are less robust than the median value.

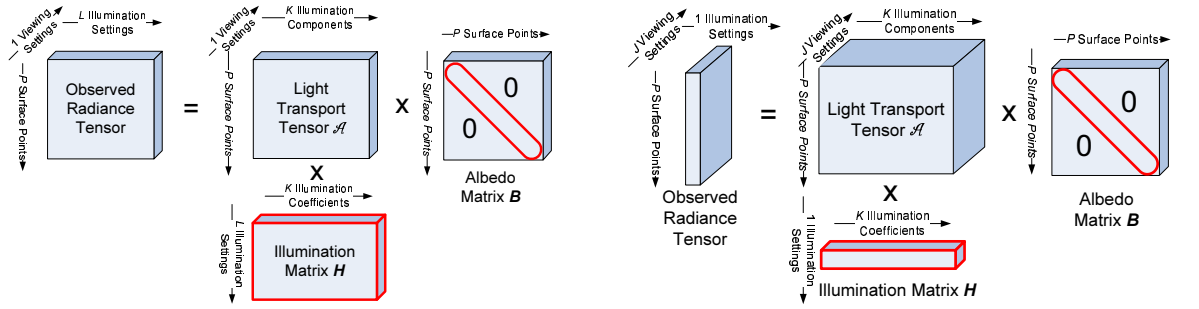


Figure 3: *Comparison of (a) varying illumination (photometric) methods, and (b) our method (multi-view with unknown illumination) in terms of surface radiance tensor decomposition. Red contours indicate unknowns.*

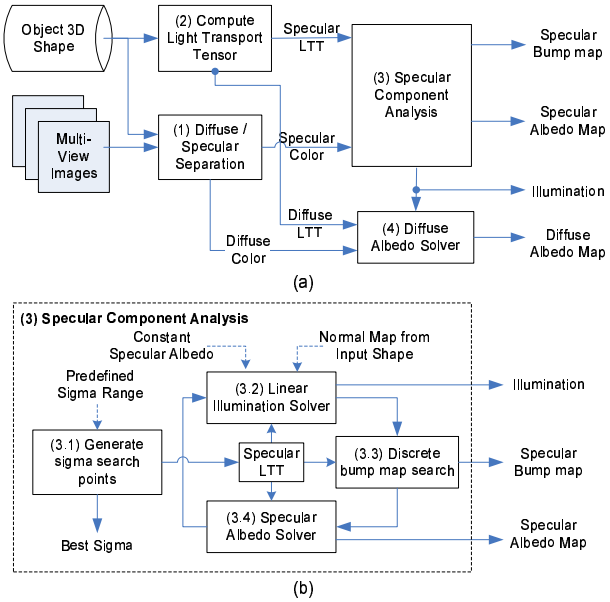


Figure 4: *(a) Flow diagram of the illumination and reflectance estimation algorithm. (b) Details of the Specular Comonent Analysis module. Dashed arraw represents the initialization.*

For diffuse material, the light transport is view independent and can be computed using the lambertian law:

$$\mathcal{A}_d(p,j,k) = \frac{1}{\pi}\int_{\Gamma} s(\Theta_l)Y_k(\Theta_l)(\vec{l}\cdot\vec{n}_p)\,d\Theta_l \qquad (4)$$

where Γ is the upper hemisphere about the surface normal of point p, and $s(\Theta)$ is the shadow function specifying whether p is in shadow with respect to the light incident from direction Θ.

For specular BRDF, we use a gaussian filtered mirror model. The specular reflection is computed by first filtering

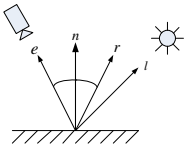


Figure 5: *The viewing direction $\vec{e}$, surface normal $\vec{n}$, light direction $\vec{l}$ and the reflected viewing direction $\vec{r}$.*

the illumination function $L(\Theta)$ with a circular gaussian filter and then reflecting it:

$$\mathcal{A}_s(p,j,k) = \frac{1}{4\pi\sigma^2}\int_{\Gamma} exp\left(-\frac{\Phi_{lr}}{2\sigma^2}\right) Y_k(\Theta_l)s(\Theta_l)\,d\Theta_l \qquad (5)$$

where Φ_{lr} is the angle between the light direction $\vec{l}$ and the reflected viewing direction $\vec{r}_{pj}$, as illustrated in Fig. 5. σ is similar to the roughness parameter in Torrance-Sparrow model that controls the blurry effects of the BRDF to the illumination. Note that this model does not handle the Fresnel effects so the errors near the grazing viewing angle will be larger.

The shadow function $s(\Theta)$ is completely determined by the object geometry. We evaluate this function on a uniform grid of directions placed on the unit sphere. We use discrete integral approximation to compute the diffuse and specular light transport tensor based on these sampling directions.

4.2. Specular Component Analysis

The Specular Component Analysis module implements the bilinear factorization algorithm discussed in Sec. 3.3, where two linear system solvers are alternated to solve both illumination and specular albedo map. Since we use the specular reflection to estimate the illumination, the reflection has to be observable in the input images. Similarly, for a surface point having non-zero specular albedo, specular reflection has to be observed in at least one input images. These requirement are easier to be met when the number of light

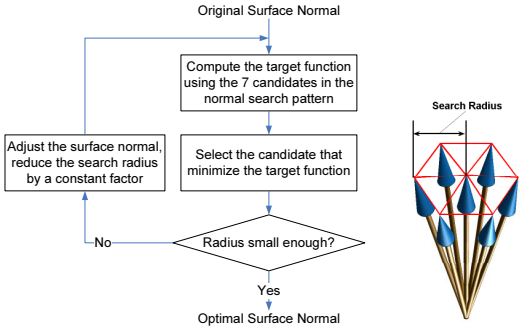


Figure 6: *Discrete normal map search algorithm and the hexagon normal search pattern (right).*

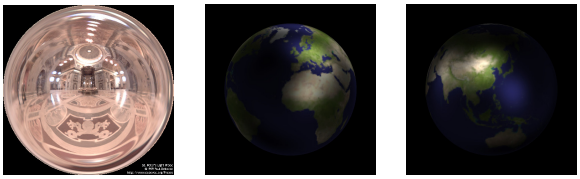


Figure 7: *Light probe data used for synthesis (left) and two of the 24 synthesized images of the earth data set.*

sources and viewpoints increase. To handle the geometric errors from the input shape (camera calibration error, shape to camera registration error, and lack of geometric details) and the unknown σ parameter in the specular basis BRDF, we add additional discrete normal map search and 1D σ search into the algorithm (Fig.4 (b)).

To compensate for the geometric error, we perform a multi-scale greedy search in the neighborhood of the normal computed from the input shape to reduce the sum of squared residue of the bilinear system in Eq. (1) [‖]. The search algorithm is outlined in Fig. 6. During the discrete normal map search, both the albedo map and illumination are fixed. The estimated surface normal can be used as a specular bump map to compute the LTT for the subsequent albedo and illumination solver until a new bump map search is performed. Each bump map search always starts from the same normal computed from the input shape, which will prevent the normal map from diverging far away from the original shape as the number of iterations increases.

For the unknown σ parameter in the specular basis BRDF, we perform a 1-D golden section search to find an optimal value. A sequence of σ values are generated according to the golden bisection rules and for each σ the tensor factorization residue is computed (after the normal map search and the illumination and albedo estimation are finished). The one that minimize the residue is selected as the optimal parameter. Note that previous research [RH01] indicates that estimating σ and illumination at the same time is not well-conditioned. Nishino et al. [NZI01, NZI05] rely on strong regularizations to resolve this illumination-BRDF ambiguity, which restricts the light to be point sources and no shadowing effects. One source of information that can be used to constraint the illumination is the sharp shadow. Our LTT incorporates the self-shadow information from object geometry which can be used to dis-ambiguous the illumination and BRDF. But even if the scene does not have any shadow and the estimation problem is not well-conditioned, golden section search will converge to some intermediate value and will not diverge.

[‖] This does not solve all the problems though, since the geometric errors also give inaccurate image correspondences which cannot be compensated by adjusting the surface normal

4.3. Other Implementation Details

We use OpenGL to compute the shadow function in 1164 sampling directions. For each sample direction, we render the surface mesh orthographically onto a plane perpendicular to that direction. The shadow function is set to 1 for all the visible triangles in that direction, and set to 0 for the rest. To represent illumination, we use spherical harmonics up to the 10^{th} order , which result in an 121×1 illumination matrix H.

The iterative algorithm in the specular component analysis is initialized by a constant specular albedo for all the triangles and surface normal equal to the original normal from the input shape. The search range of σ is set to be large enough to include the typical values used in our experiments. Due to the diagonal structure of the albedo matrix, the surface albedo, given the illumination, can be solved for independently at each surface location. On the other hand, illumination, given the surface albedo, needs to be solved for using equations from all surface points.

We implement our algorithm in mixed matlab and c++ program. To represent the texture and geometric details, the input geometry is subdivided to have around 200 300K triangles. For each σ value we perform 4 iterations in the specular component analysis. It takes about 12 hours for the program to search a ten value σ sequence and estimate all other unknowns (on a P4 2.8Ghz computer with 1 GB memory). Given the pre-computed shadow function, the rendering takes about 1-2 minute/frame.

5. Experimental Results

We perform experiments on one synthetic dataset and two real datasets. We also demonstrate the application of our estimated results in various rendering tasks.

5.1. Estimation Results

We use a synthetic experiment to show that we can solve for both texture and illumination unknowns from only multi-

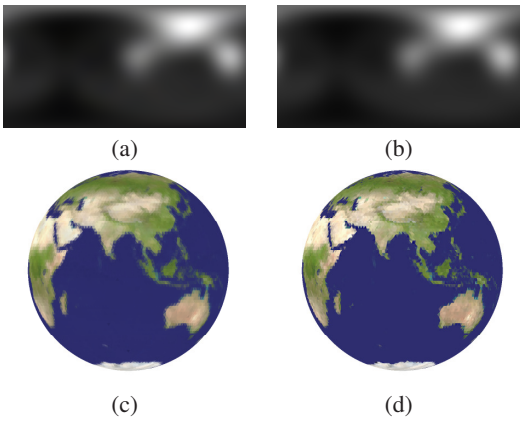


Figure 8: *Estimated illumination (top, in longitude $[0,2\pi]$ - latitude $[0,\pi]$ plot) and specular texture (bottom) for the earth data set. (a) Estimated results . (b) Ground truth. Ground truth illumination is filted by the same phone BRDF and DC is normalized to 1. The root mean square error (RMSE) for estimated illumination is 0.0026 (DC normalized to 1) and the RMSE for estimated specular texture is 0.032.*

view specular radiance given the basis light transport. We render 24 different images of the sphere with a pure specular BRDF modulated by a texture map of the earth using the Phong model (Fig. 7). The Phong exponent used in rendering is 10 for the entire sphere and the illumination is a set of 40 directional sources obtained by clustering the light probe data. [**] These light sources are converted to white color during rendering because we cannot recover the global scaling factor between different color channel.

Since we have the perfect geometry and sphere doesn't have any shadow, we don't perform σ parameter or normal search on this data set and only use the perfect mirror light transport model. The estimated illumination is therefore the combined effect of the actual illumination and the Phong BRDF. Fig. 8(a) and (b) show the texture map and illumination along with the ground truth illumination (after it is filtered by the same Phong BRDF and normalized to have unit DC component) and the ground truth specular texture map. The estimated illumination closely matches the filtered ground truth illumination, and the estimated albedo map correctly shows the texture variation with the exception of some blur. This is mostly due to the truncation errors in the spherical harmonic expansion of the illumination, and the surface granularity we chose in the estimation. This experiment shows that given the basis light transport, we can extract the arbitrary specular texture map and illumination up to a global scaling ambiguity.

[**] http://www.debevec.org/Probes/.

Figure 9: *Three of the 30 input images of the Fish data set.*

Figure 10: *Structured light scanned fish model (left) and the estimated diffuse color.*

In the second experiment, we use the fish data set captured by Wood et al. [WAA*00] with structured light scanned geometric shape [††]. We select 30 input images evenly distributed around the fish out of the original more than 600 images (Fig. 9) [‡‡].

We separate the diffuse and specular component by using the median color method mentioned earlier. Fig. 10 shows the fish shape we use and the separated diffuse color, where specular highlights are correctly removed. Fig. 12 compares the estimation result of directly solving the bilinear system for diffuse and specular albedo (the direct method) - solving for the specular albedo after specular texture separation - and the specular texture separation algorithm with bump map optimization. The direct method shows texture overfitting problem on the tail and the fish head area, where strong specularities occur (Fig. 12(a)). Specular texture separation helps regularize the texture solution so the diffuse texture becomes more uniform (Fig. 12(b)). With specular bump map optimization, those fine details not captured in the input shape (e.g., the fish scales) can be modeled and the specular texture become more smooth (Fig. 12(c)). Fig. 12 also shows the synthesized images by different estimation algorithms for one input view. The direct method shows artifacts in some region. The diffuse-specular separation helps to remove these artifacts, and the bump map optimization helps to fit the correct highlight region. Overall, the estimation al-

[††] We perform addition camera-geometric shape registration refinement based on the object silhouettes to improve the camera calibration accuracy of the original dataset.

[‡‡] Due the low intensity of the original images, all the images of the fish data set under original illumination are scaled by a factor of 1.5 for display purpose.

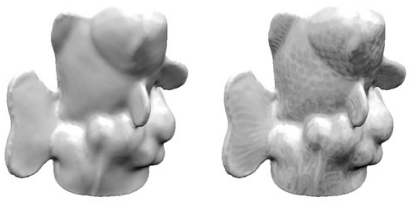


Figure 11: *Original fish shape (left) compared with the estimated bump map (right); geometric details not modeled by the input shape are estimated.*

gorithm correctly decomposes the surface radiance into illumination and diffuse/specular albedo. The estimated illumination map shows two strong clusters of light which correspond to the two point light sources used for capturing. The estimated specular map also shows correct spatial variation of the texture, e.g. the golden lines on the fish. Note that due to this light configuration, part of the fish is under complete shadow and we cannot recover any texture information from those parts. The synthesized images appears to be a little blur than the original image, which is largely due to the calibration and geometric registration error. Since our algorithm estimate the surface albedo by combining information from different viewpoints, any mis-correspondence between different input images can result in blurriness. In Fig. 12(c), we can find some bright yellow specular region on the fish body near the tail. That's the estimation error due to the inter-reflection from the tail. In our algorithm it is mistaken as a very bright specular albedo. Fig. 11 compares the estimated specular bump map compared with the original shape. Geometric details not modeled in the input shape are recovered. We also measure the Root-Mean-Squared-Error (RMSE) of synthesized images compared with the 30 input images. The average RMSE is 0.0465 on a scale of [0, 1].

We also apply our algorithm to the Van Gogh data set [CBCG02]. Fig. 13 (a) shows two of the 30 input images [§§]. Fig. 13 (b) shows the estimated illumination, diffuse and specular albedo map, and bump map. The estimated illumination shows four light spots in the upper hemisphere (left part in the figure), and the white reflection from the white desktop. Strictly speaking, the reflections are not distant illumination to the object, so estimates in that region are noisier. The inter-reflection on the statue makes the diffuse estimate appears brighter especially in the shadowed regions (e.g. surface areas facing downwards). Ideally, the specular albedo map should be uniform, but the actual estimated results are quite noisy, which suggests they are used to compensate the various errors in the calibration, geometric registration and light transport models. Fig 13 (d) plots the RMSE for the synthesized result in all the 30 input images and 22 novel images.

[§§] Original Van Gogh data set used in [CBCG02] has more than 300 images.

5.2. Applications in Rendering

Due to the limit amount of data available (only images under a single illumination setting), our goal is not to recover physically accurate surface reflectance properties, but a meaningful decomposition of all the input data (the surface radiance tensor) guided by some simple light transport models. This decomposition can be used to generate plausible new images in many rendering tasks, such as novel views of the object, relighting the object in a new environment, or changing the material of the object. For example, Fig. 14 (a) shows the synthesized image for the fish data set from a novel viewpoint. Fig. 14(b) shows the result of modifying the diffuse material in a part of the fish. In Fig. 14 (c) we render the fish model after rotating the illumination, and under a new illumination captured by light probe.

Fig. 14 (d) shows the rendered Van Gogh model in a novel view compared with the ground truth image. Fig. 14 (e, f) demonstrate the effect of light editing. We can modify individual light sources in the illumination map to different colors. Note that only part of the highlight area changes color. Fig. 14 (g) shows synthesized images under illumination from the Fish data set and under a new illumination captured by light probe.

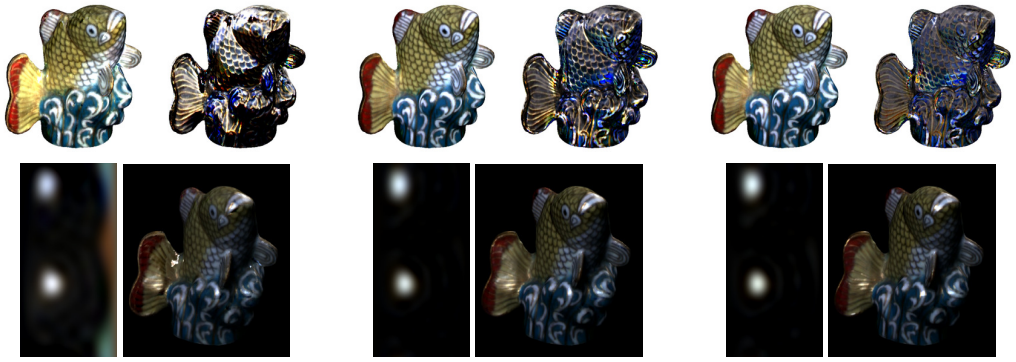
6. Conclusions and Future Work

In this paper, we have considered the problem of simultaneously estimating the illumination, diffuse/specular texture map, and bump map from multi-view images under fixed illumination. We show that view dependent light transport can be used to resolve the texture-illumination ambiguity. The illumination and the texture albedos are estimated by iteratively solving a bilinear system of equations. Our algorithm also performs a non-linear optimization of specular bump map. Experimental results show that our algorithm can be used to estimate both the texture map and illumination, and also refines the surface geometry. The estimated result can be used to render the object for novel views, novel illumination and after texture change.

Following are some possible directions for further work. The estimated bump map can be integrated back into the geometric model to recover more consistent shape details. More complex light transport model than the current diffuse-specular model can be used, but this should be done without making the estimation ill-conditioned. Inter-reflection between object parts can be modeled if the albedo matrix has non-zero off-diagonal entries.

References

[CBCG02] CHEN W.-C., BOUGUET J.-Y., CHU M. H., GRZESZCZUK R.: Light field mapping: efficient representa-



(a) (b) (c)

Figure 12: *Estimated diffuse albedo (top left), specular albedo (top right), environment illumination (bottom left), and synthesized image from an input view (bottom right) for (a) direct method (RMSE = 0.0548), (b) specular texture separation method (RMSE = 0.0441), and (c) specular texture separation with bump map optimization (RMSE = 0.0331). The ground truth image is shown in Fig. 9 right.*

(a) (b) (c) (d)

Figure 13: *(a) Two of the 30 input images of the Van Gogh data set. (b) Estimated illumination, diffuse albedo, and specular albedo. (c) Estimated specular bump map compared with the original surface shape (right). (d) RMSE of the synthesized images in all 30 input views, (blue line, average RMSE = 0.0507) and 22 novel views (red line, average RMSE = 0.0619).*

tion and hardware rendering of surface light fields. *ACM Trans. Graph. 21*, 3 (2002), 447–456.

[CT97] COHEN S., TOMASI C.: *Systems of Bilinear Equations*. Tech. rep., Stanford, CA, USA, 1997.

[DHT*00] DEBEVEC P., HAWKINS T., TCHOU C., DUIKER H.-P., SAROKIN W., SAGAR M.: Acquiring the reflectance field of a human face. In *SIGGRAPH '00: Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques* (2000), pp. 145–156.

[Geo03] GEORGHIADES A.: Incorporating the torrance and sparrow model of reflectance in uncalibrated photometric stereo. In *Proceedings of the 9th IEEE International Conference on Computer Vision* (2003), vol. 2, pp. 816–823.

[GGSC96] GORTLER S. J., GRZESZCZUK R., SZELISKI R., COHEN M. F.: The lumigraph. In *SIGGRAPH '96: Proceedings of the 23rd annual conference on Computer graphics and interactive techniques* (New York, NY, USA, 1996), ACM Press, pp. 43–54.

[HS05] HERTZMANN A., SEITZ S. M.: Example-based photometric stereo: Shape reconstruction with general, varying BRDFs. *IEEE Trans. PAMI 27*, 8 (2005), 1254–1264.

[LH96] LEVOY M., HANRAHAN P.: Light field rendering. In *SIGGRAPH '96: Proceedings of the 23rd annual conference on Computer graphics and interactive techniques* (New York, NY, USA, 1996), ACM Press, pp. 31–42.

[LKG*03] LENSCH H. P. A., KAUTZ J., GOESELE M., HEIDRICH W., SEIDEL H.-P.: Image-based reconstruction of spatial appearance and geometric detail. *ACM Trans. Graph. 22*, 2 (2003), 234–257.

[LLLS03] LI Y., LIN S., LU H., SHUM H.-Y.: Multiple-cue illumination estimation in textured scenes. In *Proceedings of the 9th IEEE International Conference on Computer Vision* (Washington, DC, USA, 2003), p. 1366.

[MPBM03] MATUSIK W., PFISTER H., BRAND M., MCMILLAN L.: A data-driven reflectance model. *ACM Trans. Graph. 22*, 3 (2003), 759–769.

[NRH04] NG R., RAMAMOORTHI R., HANRAHAN P.: Triple product wavelet integrals for all-frequency relighting. *ACM Trans. Graph. 23*, 3 (2004), 477–487.

[NSI99] NISHINO K., SATO Y., IKEUCHI K.: Eigen-texture method: appearance compression based on 3d model. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR'99)* (1999), pp. 618–624.

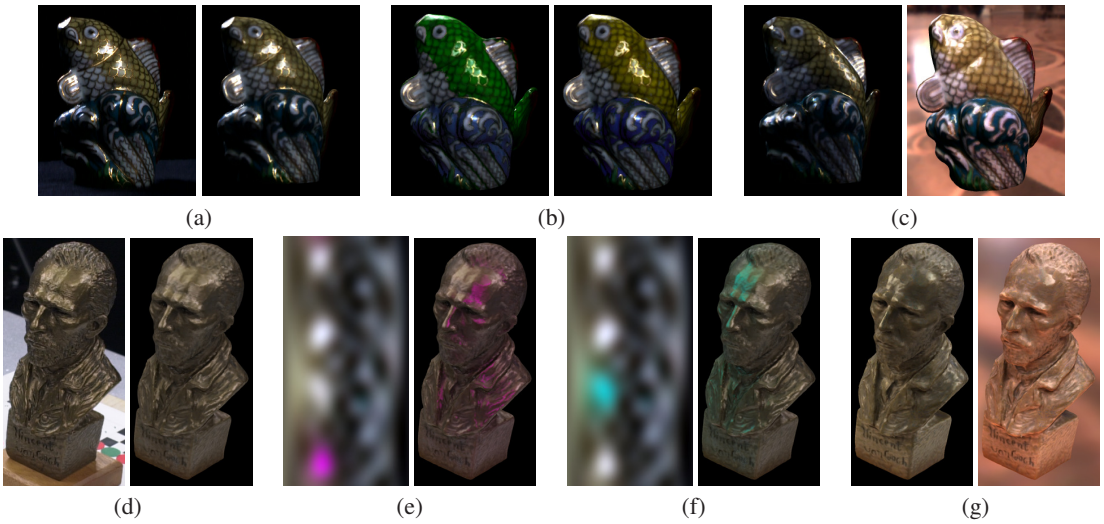


Figure 14: *Synthesizing novel images from the estimation results. Fish data set: (a) Ground truth compared with the synthesized image from a novel viewpoint (RMSE = 0.0553). (b) Fish model with part of the diffuse texture modified; note the specular part is kept constant so the highlight is unchanged (c) Synthesized fish with illumination rotated by* 60° *and under a new environment map. Van Gogh data set: (d) Ground truth compared with the synthesized image from a novel viewpoint (RMSE = 0.0531). (e) Modifying one light to purple. (f) Modified another light to cyan. (g) Synthesized images under fish data set's illumination and a new environment map.*

[NZI01] NISHINO K., ZHANG Z., IKEUCHI K.: Determining reflectance parameters and illumination distribution from a sparse set of images for view-dependent image synthesis. In *Proceedings of International Conference on Computer Vision (ICCV 2001)* (2001), pp. 599–606.

[NZI05] NISHINO K., ZHANG Z., IKEUCHI K.: *Re-rendering from a Sparse Set of Images*. Tech. Rep. DU-CS-05-12, Department of Computer Science, Drexel University, 2005.

[RH01] RAMAMOORTHI R., HANRAHAN P.: A signal-processing framework for inverse rendering. In *Proceedings of SIGGRAPH* (2001), pp. 117–128.

[SKS02] SLOAN P.-P., KAUTZ J., SNYDER J.: Precomputed radiance transfer for real-time rendering in dynamic, low-frequency lighting environments. In *SIGGRAPH '02: Proceedings of the 29th annual conference on Computer graphics and interactive techniques* (New York, NY, USA, 2002), ACM Press, pp. 527–536.

[SLS05] SLOAN P.-P., LUNA B., SNYDER J.: Local, deformable precomputed radiance transfer. *ACM Trans. Graph. 24*, 3 (2005), 1216–1224.

[SSI03] SATO I., SATO Y., IKEUCHI K.: Illumination from shadows. *IEEE Trans. Pattern Anal. Mach. Intell. 25*, 3 (2003), 290–300.

[SWI97] SATO Y., WHEELER M., IKEUCHI K.: Object shape and reflectance modeling from observation. In *Computer Graphics Proceedings, Annual Conference Series* (1997), pp. 379–388.

[VT04] VASILESCU M., TERZOPOULOS D.: Tensortextures: Multilinear image-based rendering. *ACM Transactions on Graphics 23*, 3 (2004), 334–340.

[WAA*00] WOOD D. N., AZUMA D. I., ALDINGER K., CURLESS B., DUCHAMP T., SALESIN D. H., STUETZLE W.: Surface light fields for 3D photography. In *SIGGRAPH '00: Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques* (New York, NY, 2000), pp. 287–296.

[Wei] WEISSTEIN E. W.: *MathWorld–A Wolfram Web Resource*. MathWorld.

[WGT*05] WENGER A., GARDNER A., TCHOU C., UNGER J., HAWKINS T., DEBEVEC P.: Performance relighting and reflectance transformation with time-multiplexed illumination. *ACM Trans. Graph. 24*, 3 (2005), 756–764.

[WWS*05] WANG H., WU Q., SHI L., YU Y., AHUJA N.: Out-of-core tensor approximation of multi-dimensional matrices of visual data. *ACM Trans. Graph. 24*, 3 (2005), 527–535.

[YDMH99] YU Y., DEBEVEC P., MALIK J., HAWKINS T.: Inverse global illumination: Recovering reflectance models of real scenes from photographs. In *Proceedings of SIGGRAPH* (1999), pp. 215–224.

[ZCHS03] ZHANG L., CURLESS B., HERTZMANN A., SEITZ S. M.: Shape and motion under varying illumination: Unifying structure from motion, photometric stereo, and multi-view stereo. In *The 9th IEEE International Conference on Computer Vision* (Oct. 2003), pp. 618–625.

[ZK02] ZHOU W., KAMBHAMETTU C.: Estimation of illuminant direction and intensity of multiple light sources. In *Proceedings of the 7th European Conference on Computer Vision-Part IV* (2002), pp. 206–220.