Recovering illumination and texture using ratio images

Alejandro Troccoli
atroccol@cs.columbia.edu

Peter K. Allen
allen@cs.columbia.edu

Department of Computer Science
Columbia University, New York, NY

Abstract

In this paper we consider the problem of factoring illumination and texture from a pair of images of a diffuse object of known geometry. This problem arises frequently in 3D photography applications that use images to acquire photometric properties of a scanned object. Our approach uses the ratio of the images and the geometry information to compute the relative incident irradiance of one image with respect to the other. After the irradiance maps are recovered, we build a spatially varying albedo map, which can then be used to render the object under different illumination conditions. We present two algorithms, one for point-light source illumination, and another one based on spherical harmonics for more general illumination conditions.

1. Introduction

In this paper we consider the problem of factoring illumination and texture from a pair of images of a diffuse object of known geometry. This problem arises frequently in 3D photography: given a 3D model of a diffuse object acquired using a range sensor and a set of registered overlapping images of the object taken under different and unknown (uncalibrated) illuminations, we want to find a factorization of the images into its texture and illumination components. Once the factorization is achieved, it is possible to align all images to the same illumination and create new renderings under new illumination conditions. This approach produces better results than using plain texture mapping of the images on the 3D geometry.

The main contribution of our method is the recovery of an illumination-free texture map from a pair of images and the object geometry without prior recording or calibration of the incident illumination. There are two steps in this process: first we take the ratio of the input images, which for diffuse objects is texture-free, and estimate the illumination in the form of a pair of irradiance maps; second, we use the recovered irradiance to factor out the texture from the shading. Once we have obtained the texture in the form of an albedo map, we can render the object under new illu-

mination conditions. We first analyze the case of an object illuminated by a point light source and show that given two images of the object with the light source at different positions, it is possible to recover these positions from the ratio image and surface normals. Then we extend this case to more general illumination. Diffuse reflection has been well studied [9] and it is has been established that the Lambertian BRDF acts as a low pass filter on the incident illumination. This implies that from the image of a diffuse object only the low frequency components of the incident light field can be recovered. However, these low frequency terms approximate well the incident irradiance.

Our paper falls in the category of inverse-rendering techniques. Measuring surface reflectance of an object of known geometry under controlled and calibrated illumination has proved to produce very good results [3, 6]. But when the illumination is unknown, it is typically assumed that the material properties of the object are homogeneous over the whole surface [5] [9](i.e. the object is texture-less). Placing inverse rendering problems into a theoretical framework, Ramamoorthi and Hanrahan [10] developed a signal-processing approach. In particular they note that lighting and texture can not be factored without resorting to active methods or making prior assumptions of their expected characteristics. In this paper, we can achieve this factorization by requiring the object to be diffuse or mostly Lambertian. The ratio of two images of a convex Lambertian object is texture-free and only depends on the incident illumination.

2. Background and previous work

In inverse-rendering, it is a typical to assume that illumination falling on the scene is distant. Under this condition, and under the assumption that any non-linear effects of the camera gain have been removed, the image of a diffuse Lambertian convex object is:

$$I(x, y) = \rho(x, y)E(\mathbf{n}(x, y)) \qquad (1)$$

where I denotes the observed intensity at pixel (x, y), ρ de-

notes the albedo and E is the incident irradiance parameterized by the surface normal $\mathbf{n}$ at (x, y), and defined as the integral of the product of the incident light and the half-cosine function over the upper-hemisphere:

$$E(\mathbf{n}) = \int_{\Omega_i} L(\theta_i, \phi_i) \cos \theta_i d\omega_i. \qquad (2)$$

When the illumination source is a distant point light source, the above expression simplifies to a dot product of the surface normal and the light direction:

$$E(\mathbf{n}) = max(\mathbf{n} \cdot \mathbf{l}, 0), \qquad (3)$$

where $\mathbf{l}$ is a unit vector in the direction of the light source. Given two different images I_1 and I_2 of the same object acquired from the same viewpoint, the ratio image R is defined as:

$$R(x, y) = \frac{I_1(x, y)}{I_2(x, y)} = \frac{E_1(\mathbf{n}(x, y))}{E_2(\mathbf{n}(x, y))}, \qquad (4)$$

since the albedo terms in the numerator and denominator cancel each other. Hence, in this case, the ratio image is invariant to texture, and can be considered to represent the *irradiance ratio*. Marschner and Greenberg [7] have used these irradiance ratio images to relight a photograph of an object of known geometry and reflectance. Given a photograph of the object, its geometry and its surface reflectance, they developed an inverse rendering system that solves for the illumination of the scene and allows the user to modify the recovered illumination to generate a new renderings. However, this illumination recovery procedure works for non-textured objects only.

Irradiance ratio images can also be used for relighting diffuse objects if we have *two* overlapping images captured under different illumination. This idea has been applied in 3D photography by Beauchesne and Roy [2] to relight two partially overlapping images and align both of them to the same lighting, without the need to solve explicitly solve for the illumination. The irradiance ratio is computed from the area of overlap and applied to the non-overlapping regions. This requires that the overlapping regions samples all surface normals that one expects to find in the non-overlapping region. In this paper we extend this idea one step further by solving for the illumination and computing and albedo map.

Ratio images have also been used for recognition tasks, with particular emphasis in face recognition under different illumination. However, ratio images for recognition are defined in a slightly different manner. Since faces of different subjects can be considered to have the same geometry but different albedos, Shashua and Riklin-Raviv [11] define the quotient image Q of two different faces a and b:

$$Q(x, y) = \frac{\rho_a(x, y)}{\rho_b(x, y)} \qquad (5)$$

In their work, Shashua and Riklin-Raviv show how this quotient image can be computed from an image of face a and three images of face b illuminated by three non-collinear point light sources. Their recognition algorithm does solve for the illumination of the test image a, but it is different to ours in aim and context: it's main purpose is to obtain an illumination invariant signature of the face, it makes an implicit assumption that the geometry is the same but it does not make explicit use of the geometry, and requires a database of images to bootstrap the process. In later work, Wang et al. [12] take this method a step further and generalize it to images illuminated by non-point light sources.

3. Methodology

The definition of our problem is as follows. Given

1. $\mathcal{G}$ the geometry of an object.

2. $\mathcal{I} = \{I_1, I_2, \ldots, I_n\}$ a set of photographs of the object captured under unknown illumination conditions. The images overlap.

3. $\mathcal{P} = \{P_1, P_2, \ldots, P_n\}$ the set of camera projection matrices that relates $\mathcal{G}$ with $\mathcal{I}$.

we want to factor shading from texture and recover an albedo map of the scene. The images need not to be taken from the same viewpoint. Since we have the geometry of the scene and the projection matrices, we can warp any two overlapping images to the same viewpoint. To simplify the discussion that follows, we will only consider a single pair of images and assume that these have been warped to the same view.

The first step of our method is to recover the relative illumination. For this, we consider two different cases: 1) an object illuminated by a point light source, 2) generalized illumination. The point light source algorithm recovers the position of the light sources that generated each of the input images. The generalized illumination algorithm, which is based on spherical harmonics, will recover a low frequency representation of the illumination. In both cases, the input to the illumination recovery procedure is a ratio image $R(x, y)$, which we compute by taking the quotient of the two input images, and a normals image $\mathbf{n}(x, y)$ which gives the normal for each pixel and that is generated by raytracing the geometry of the object.

3.1. Point light source illumination

When an object is illuminated by a directional point light source with direction $\mathbf{l}_1 = (l_{1.x}, l_{1.y}, l_{1.z})$, the irradiance for a pixel with associated normal $\mathbf{n}$ is:

$$E(\mathbf{n}) = L \max(\mathbf{n} \cdot \mathbf{l}_1, 0), \qquad (6)$$

where L denotes the source intensity and $\cdot$ the vector dot product. Putting this expression into (4) for a pair of images acquired under point light sources $\mathbf{l}_1$ and $\mathbf{l}_2$ we obtain:

$$R(x,y) = \frac{L_1 \max(\mathbf{n}(x,y) \cdot \mathbf{l}_1, 0)}{L_2 \max(\mathbf{n}(x,y) \cdot \mathbf{l}_2, 0)}, \tag{7}$$

defined only for non-zero values of the denominator and numerator. It should be clear at this point that there is an inherent ambiguity in the above expression since multiplying the numerator and denominator by the same constant does not affect the final result. Therefore, it will not be possible to recover the absolute intensities of the light sources. Hence, we can fix L_1 to unity and solve for $\mathbf{l}_1$ and $\mathbf{l}_2$ scaled by L_2. To simplify the notation, we will drop the (x,y) coordinates in $R(x,y)$ and $\mathbf{n}(x,y)$ and use a single subindex to enumerate all pixels. Then, we setup the following system of equations:

$$\begin{bmatrix} \mathbf{n}_0^T & -R_0\mathbf{n}_0^T \\ \vdots & \vdots \\ \mathbf{n}_k^T & -R_k\mathbf{n}_k^T \end{bmatrix} \begin{bmatrix} \mathbf{l}_1 \\ L_2\mathbf{l}_2 \end{bmatrix} = 0. \tag{8}$$

Thus we have a linear system of the form $\boldsymbol{A}\mathbf{x} = 0$. The solution we are looking for is the one dimensional null-space of $\boldsymbol{A}$. When the dimension of the null-space of $\boldsymbol{A}$ is greater than one it will not be possible to solve uniquely for the light directions. This condition will arise if the distribution of the imaged surface normals is small: e.g. if the scene is a flat wall. Given the null-space vector $\mathbf{x}$ (a 6 dimensional vector), we obtain $\mathbf{l_1}$, $\mathbf{l_2}$ and L_2 from:

$$\mathbf{l_1} = \frac{(x_0, x_1, x_2)}{||(x_0, x_1, x_2)||} \tag{9}$$

$$\mathbf{l_2} = \frac{(x_3, x_4, x_5)}{||(x_3, x_4, x_5)||} \tag{10}$$

$$L_2 = \frac{||(x_3, x_4, x_5)||}{||(x_0, x_1, x_2)||} \tag{11}$$

To handle color images we could treat each channel separately and solve (8) per channel. However, this typically yields three slightly different positions for the light source. We can obtain a more robust solution if we convert the image to luminance space and use the luminance values, instead. After we recover the direction of the light source, the relative scale L_2 for each channel c is obtained from the original images by averaging the following expression over all pixels:

$$L_{2,c}(x,y) = \frac{\max(\mathbf{n}(x,y) \cdot \mathbf{l}_1, 0)}{R(x,y) \max(\mathbf{n}(x,y) \cdot \mathbf{l}_2, 0)}. \tag{12}$$

We can only compute the relative intensities for each channel. Also, note that nothing is known about the absolute chromaticity of the light sources. By fixing the intensity L_1 to the same value for the all three channels, we assume that light to be white. This chromatic ambiguity can not be solved without further assumptions or resorting to a color calibration object.

3.2. Generalized illumination

We can extend the previous idea to a more general form of illumination. In this case we define irradiance as an expansion in terms of spherical harmonic bases. Ramamoorthi and Hanrahan [9] and Basri and Jacobs [1] have established that the image of a diffuse convex object under general illumination is well approximated by a low dimensional spherical harmonic expansion. Spherical harmonics are orthonormal basis defined over the sphere. Using this framework, we can approximate the incident irradiance as:

$$E(\mathbf{n}) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} A_l L_{lm} Y_{lm}(\mathbf{n}). \tag{13}$$

In (13) above, Y_{lm} are the spherical harmonic functions, L_{lm} are the spherical harmonic coefficients of the incident illumination, and A_l is a constant that represents the effects of multiplying the incident light with the half-cosine function. In other words, (13) is the frequency space equivalent of the integral (2) [9]. In this context, we want to solve for L_{lm}. Since A_l decays very rapidly, a very good approximation can be obtained by limiting $l \leq 2$. A first order approximation (up to $l = 1$) has a total of four terms and a second order approximation has a total nine. Before we write the expression of the irradiance ratio in spherical harmonics, we will make one more notation change for clarity purposes. We will replace the double-indexed Y_{lm} functions and L_{lm} coefficients by their single-index equivalent Y_s and L_s, where $s = l^2 + l + m$. Also, since we have to solve for two different illuminations L_s, we will denote these with L_{1s} and L_{2s}. Using this new notation, we can substitute (13) into our irradiance ratio expression to obtain:

$$R_i = \frac{\sum_{s=0}^{n} A_s L_{1s} Y_s(\mathbf{n}_i)}{\sum_{s=0}^{n} A_s L_{2s} Y_s(\mathbf{n}_i)}. \tag{14}$$

where $n = 4$ or $n = 9$ depending on the order of the desired approximation. We can now derive a system of linear equations similar to (8) on the unknown lighting coefficients L_{1s} and L_{2s}.

$$
\begin{bmatrix}
Y_0(\mathbf{n}_0)\ldots Y_n(\mathbf{n}_0) & -R_0 Y_0(\mathbf{n}_0)\ldots \\
\vdots & \vdots \\
Y_0(\mathbf{n}_k)\ldots Y_n(\mathbf{n}_k) & -R_k Y_0(\mathbf{n}_k)\ldots
\end{bmatrix}
\begin{bmatrix}
L_{10} \\ \vdots \\ L_{1n} \\ L_{20} \\ \vdots \\ L_{2n}
\end{bmatrix} = 0.
\tag{15}
$$

The solution to (13) is once more the null-space of $\boldsymbol{A}$ and the coefficients L_{1s} and L_{2s} will be defined up to a scale factor. This scaling factor can be fixed by setting $L_{10} = 1$, which fixes both the relative scale and chromaticity of the illumination.

The stability of (15) is highly dependent on the distribution of surface normals over the region of overlap between the two images. For better results and higher robustness against noise, it is possible to re-cast the problem in terms of principal components. This means replacing the spherical harmonic basis by lower dimensional orthogonal basis obtained using principal component analysis (PCA). The rationale behind this change of basis is that the principal components are vectors in the direction of greater variability (in this case due to illumination changes). Ramamoorthi derived in [8] an analytic expression for the principal components of the image of an object under all possible point light sources and showed that these are related to the spherical harmonic basis Y_s. In particular, Ramamoorthi shows that the eigenvectors obtained from PCA of the image space of an object illuminated under all possible point light sources can be well approximated as a linear combination of spherical harmonic functions up to order two. If $\boldsymbol{V}$ is the matrix with the principal eigenvectors as columns, then there exists a matrix $\boldsymbol{U}$ such that $\boldsymbol{V} \approx \boldsymbol{Y}\boldsymbol{U}$, where $\boldsymbol{Y}$ is a matrix whose columns are the first nine spherical harmonics $Y_0...Y_8$. The matrix $\boldsymbol{U}$ can be computed analytically from the geometry of the object and details on how to do this are given in [8]. Using the eigenvectors $\boldsymbol{V}_i$ as the new basis, we can now write the incident irradiance as:

$$
E(\mathbf{n}) = \sum_{i=0}^{n} e_i \boldsymbol{V}_i(\mathbf{n}),
\tag{16}
$$

where e_i are the coefficients of the irradiance in principal components basis. The number of terms to employ in this new approximation will depend on the object geometry, but by looking at the eigenvalues associated to each vector it is possible to determine a good cut-off point. Now, we can write (15) as:

$$
\begin{bmatrix}
V_0(\mathbf{n}_0)\ldots V_n(\mathbf{n}_0) & -R_0 V_0(\mathbf{n}_0)\ldots \\
\vdots & \vdots \\
V_0(\mathbf{n}_k)\ldots V_n(\mathbf{n}_k) & -R_k V_0(\mathbf{n}_k)\ldots
\end{bmatrix}
\begin{bmatrix}
e_{10} \\ \vdots \\ e_{1n} \\ e_{20} \\ \vdots \\ e_{2n}
\end{bmatrix} = 0.
\tag{17}
$$

Once we find the coefficients e_{1i} and e_{2i} we can find the corresponding L_{1s} and L_{2s} coefficients by substitution into:

$$
L_{1s} = \frac{\sum_i^n \boldsymbol{U}_{si} e_{1i}}{A_s}, \qquad
L_{2s} = \frac{\sum_i^n \boldsymbol{U}_{si} e_{2i}}{A_s},
\tag{18}
$$

where $\boldsymbol{U}_s$ is the sth row of $\boldsymbol{U}$. To handle color images we treat each of the RGB channels separately and solve for three sets of coefficients L_{1i} and L_{2i}. Once again, as before, there is an inherent chromatic ambiguity that we can only solve for if we have an image of a color calibration object.

4 Extracting the albedo map

Once we have solved for the relative irradiance we can compute an albedo map. Note that the chromatic and scale ambiguity in the estimated irradiance will translate to the estimation of the albedo map, so this will also be defined up to scale. From the image pair I_1 and I_2 with estimated irradiance E_1 and E_2 we compute the albedos as:

$$
\rho_1(x,y) = \frac{I_1(x,y)}{E_1(\mathbf{n}(x,y))}
\tag{19}
$$

$$
\rho_2(x,y) = \frac{I_2(x,y)}{E_2(\mathbf{n}(x,y))}
\tag{20}
$$

$$
\rho(x,y) = \frac{I_1(x,y)\rho_1(x,y) + I_2(x,y)\rho(x,y)}{I_1(x,y) + I_2(x,y)}.
\tag{21}
$$

In other words, for each pixel (x,y) we set its albedo $\rho(x,y)$ to a weighted average of the albedos we obtain from I_1 and I_2. The weights are set to the pixel intensities, so that dark pixels, which are usually noise, are down-weighted.

5 Some practical aspects

There are certain practical aspects that are worth mentioning:

Surface normal aggregation. Each pixel in the ratio image provides one constraint to the system of equations. If we use every pixel, the size of the equations matrix could become too large to handle. Instead of using all pixels we

Figure 1: Objects used for testing our algorithm. Starting from the left, first come the synthetic renderings: a sphere and the Armadillo; followed by three objects with their geometry acquired using photometric stereo: the buddha, the cat and the owl; and finally two scanned objects: the chicken and the girl. Each row shows the objects with a different illumination.

	Point source 1			Point source 2			Rel. intensity
	x	y	z	x	y	z	(R, G, B)
Actual position	-0.58	0.36	0.73	0.28	-0.28	0.92	(5.00, 10.00, 20.00)
Sphere	-0.58	0.36	0.73	0.27	-0.28	0.92	(5.03, 10.10, 20.48)
Armadillo	-0.58	0.35	0.73	0.27	-0.28	0.92	(5.11, 10.27, 20.88)

Table 1: Ground truth and recovered light directions and relative intensities for the synthetic images of the sphere and the Armadillo

can aggregate the ratios per surface normal. To do this, we can take any 2D parametrization of the sphere and compute the average normal and the average ratio for each (u, v) coordinates.

Weighted least squares. For robustness against outliers, we can solve the linear systems using weighted least squares. We weight the equations according to the angle that the surface normal makes with the camera axis, giving in this way less importance to foreshortened points. Also, because we aggregate the surface normals, we can ignore those orientations that received few contributing pixels during the aggregation.

Almost-convex objects and self-shadowing. Our analysis is based on the assumption that the objects of interest are convex. But, it is possible to still work with almost-convex objects. For an object illuminated by a point light source, it suffices to ignore those pixels that correspond to cast shadows. We implemented this technique by setting a threshold on the luminance value of the pixels. Pixels that do not meet the threshold condition are excluded in all stages of our algorithm. For the generalized illumination case, self-shadowing is more problematic. We have not completely addressed this case and a possible solution is proposed in the final discussion.

6 Results

We tested the point light source and generalized illumination algorithms on synthetic and real data (see Figure 1). The synthetic scenes were generated from a model of a sphere and a model of the Armadillo[1] textured with a synthetic wooden pattern and rendered using a ray-tracer . For the real data sets, we used three models acquired using photometric stereo: - a buddha, a cat and an owl; and two models that were scanned using a Polhemus hand-held scanner: a chicken and a figure of a girl. The synthetic and photometric stereo models are good for ground-truth comparisons since we know the position of the light sources and do not require image registration. For the chicken and girl models, we captured several images varying the position of a point light source but leaving the viewpoint fixed. We then manually registered these images with the 3D model using software that we have developed for that purpose.

6.1. Point light source model

We ran our point-light source estimation model of section 3.1 on all of the image pairs shown in Figure 1. Tables 1 and 2 show the ground truth and recovered light source positions and relative scales for the synthetic and photometric stereo models. For the synthetic scenes, the recovered light source

---

[1]The Armadillo model was downloaded from the Stanford scanning repository.

	Point source 1			Point source 2			Rel. intensity
	x	y	z	x	y	z	(R, G, B)
Actual position	0.40	0.48	0.78	-0.32	0.49	0.86	(1.00, 1.00, 1.00)
Buddha	0.39	0.57	0.71	-0.34	0.54	0.76	(1.03, 1.03, 1.04)
Cat	0.39	0.49	0.78	-0.33	0.47	0.82	(1.09, 1.09, 1.05)
Owl	0.39	0.48	0.78	-0.31	0.44	0.84	(1.02. 1.01, 1.00)

Table 2: Ground truth and recovered light directions and intensities for the buddha, cat and owl models.

Figure 2: Results obtained using the point light source model for the images in Figure 1. The top row shows the recovered albedo, the middle row shows the factored irradiance for the first illumination, and the last row the factored irradiance for the second illumination. Notice how the factorization de-couples texture from shading.

directions and scaling factors shown in Table 1 are almost identical to the actual directions. Similarly, the computed light source directions for the buddha, cat and owl models listed in Table 2 are very close the ground truth data. For the chicken and girl models we did not have ground truth data, but we ran our algorithm and computed the light source directions.

As a second step, we used the computed light direction to factor the input images into their corresponding texture (albedo maps) and shading (irradiance) components. The results are shown in Figure 2 - the top row shows the albedo map, and the second and third rows the irradiance maps for each of the input images. Note that, with the exception of a few minor artifacts, the albedo maps do not contain the effects of the illumination. This is certainly true for the synthetic models: both the sphere and the armadillo look completely flat. For the real data sets, some artifacts can be seen where the surfaces are not purely Lambertian. For example, the owl shows some specular components in the albedo map. Other artifacts are brighter spots in non-convex regions, in particular at the junction of the head and

body of the cat and owl models, the junction of the arm and body in the chicken model, and the junction of the hair and face in the girl model. The convexity assumption fails here and inter-reflections influence the final result. The result can be seen as a brightening of the albedo map, since the pure Lambertian model can not explain the increase in irradiance due to inter-reflections. As a final comment, the pants in the chicken model are not in the albedo map since the luminance value of those pixels falls below the shadow threshold, and hence ignored by the algorithm.

To obtain a quantitative measurement of the quality of the factorization, we compared the obtained irradiance images with the irradiance images generated using the ground truth light source position. Since there is a scale ambiguity that is inherent to our method, we normalized all images before computing the error metric. This normalization was achieved by setting the norm $\parallel I \parallel^2 = \sum_{x,y} I(x,y)^2$ equal to one. Then, for a given pair of normalized ground truth image I^0 and reconstructed irradiance image I^1, we com-

6

Figure 3: Results obtained using the generalized light source model the images in Figure 1. The top row shows the recovered albedo, the middle row shows the factored irradiance for the first illumination, and the last row the factored irradiance for the second illumination.

Model	Method	Error 1	Error 2
Sphere	PL	< 0.1%	< 0.1%
	PCA 5	0.40%	0.20%
Armadillo	PL	< 0.1%	< 0.1%
	PCA 3	3.50%	4.30%
Buddha	PL	0.40%	0.50%
	PCA 3	0.10%	1.60%
Cat	PL	< 0.1%	< 0.1%
	PCA 3	4.40%	4.50%
Owl	PL	< 0.1%	< 0.1%
	PCA 3	3.80%	3.50%

Table 3: Normalized reconstruction error for the irradiance images. The first column indicates the object, the second one the method used (PL = point light, PCA n = generalized illumination using PCA of size n), and the last two columns show the normalized reconstruction error for the two irradiance images.

puted the relative squared error of the reconstruction[2]:

$$err(I^1, I^0) = \frac{\| I^1 - I^0 \|^2}{\| I^0 \|^2}. \qquad (22)$$

The resulting reconstruction errors are reported as percentages in Table 3. It can be observed that the reconstructed irradiance images using the point light source algorithm are very accurate.

---

[2]The relative squared error is frequently used in the literature (e.g. [1, 4]) as a metric for evaluating the goodness of image reconstructions.

6.2. Generalized light model

We tested our generalized light algorithm with the images in Figure 1. In all cases, we first analytically computed the PCA basis and we used a dimension 3 basis, except for the sphere for which we used a basis set of dimension 5. We found these dimensions to be optimal empirically, by testing with different basis size. The resulting factorizations into irradiance images and albedo maps are shown in Figure 3 and the reconstruction errors for the irradiance images are tabulated in Table 3. It can be seen that the model approximates well the irradiance and produces a good factorization. The reported reconstruction errors are less than 4.5%. For the sphere, the quality of the approximation is as good as for the point-light source model. For the remaining models, the irradiance reconstruction error varies between 0.5% for the buddha to 4.50% for the cat.

7. Discussion and future work.

We have presented two different methods of factoring texture and illumination from a pair of images of a diffuse convex object based on the ratio of two images. We ran tests on real and synthetic data and have shown good results. In running the experiments we have noticed a number factors that can affect these results: accuracy of the normals, non-convex regions and shadows. We found that for the real data, our algorithm was sensitive to the shadow threshold. Varying the shadow threshold, which determines which pixels are in shadow and should not be considered in the computation of irradiance, produced different results, some of

which in the extreme case of very low or very high thresholds deviated significantly from the ground truth. We also found, as expected, that good surface normals and good image registrations are required to obtain good results. This was quite evident when working with the chicken and girl models. In particular, for the girl, we had to mask out the region around the nose and the flowers on the umbrella, since the surface normals in those areas where not that accurate. In addition, the scanned geometry was smoothed out significantly to reduce noise. Finally, as we have seen in the results section, inter-reflections in non-convex regions can cause a bias in the albedo map estimation.

The applicability of the presented algorithms to different objects and illumination environments is conditioned by the geometry of the object and its illumination, but not by the object's albedo map, since the albedos are factored out when computing the ratio image. In particular, the generalized illumination algorithm using PCA basis has proved to work well for point-light source illumination and some area light sources, but further research is being conducted to evaluate its applicability to more complex illumination environments.

As for future work, we still have to address self-shadowing problem when working with the generalized light source illumination model. We can address this issue by using the geometry information to find those regions in the object that are occlusion free. If we consider only the pixels associated with those regions, we fall back into the convex case and we can estimate the irradiance.

Finally, we want to apply our method to outdoor scenes. This was our main motivation in developing the ideas explored in this paper, since illumination in outdoor scenes can not be controlled during day-time. Our current line of work is in the area of 3D modeling of historic sites, most of which are outdoor structures.

Acknowledgments

References

[1] R. Basri and D. W. Jacobs. Lambertian reflectance and linear subspaces. *IEEE Trans. Pattern Anal. Mach. Intell.*, 25(2):218–233, 2003.

[2] E. Beauchesne and S. Roy. Automatic relighting of overlapping textures of a 3D model. In *Proceedings of Computer Vision and Pattern Recognition*, 2003.

[3] P. Debevec, T. Hawkins, C. Tchou, H.-P. Duiker, W. Sarokin, and M. Sagar. Acquiring the reflectance field of a human face. In *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, pages 145–156. ACM Press/Addison-Wesley Publishing Co., 2000.

[4] D. Frolova, D. Simakov, and R. Basri. Accuracy of spherical harmonic approximations for images of lambertian objects under far and near lighting. In *ECCV (1)*, pages 574–587, 2004.

[5] K. Ikeuchi and K. Sato. Determining reflectance properties of an object using range and brightness images. *IEEE Trans. Pattern Anal. Mach. Intell.*, 13(11):1139–1153, 1991.

[6] H. P. A. Lensch, J. Kautz, M. Goesele, W. Heidrich, and H.-P. Seidel. Image-based reconstruction of spatial appearance and geometric detail. *ACM Trans. Graph.*, 22(2):234–257, 2003.

[7] S. R. Marschner and D. P. Greenberg. Inverse lighting for photography. In *Proc. 5th Color Imaging Conference*, 1997.

[8] R. Ramamoorthi. Analytic PCA construction for theoretical analysis of lighting variability in images of a Lambertian object. *IEEE Trans. Pattern Anal. Mach. Intell.*, 24(10):1322–1333, 2002.

[9] R. Ramamoorthi and P. Hanrahan. On the relationship between radiance and irradiance: Determining the illumination from images of a convex Lambertian object. *Journal of the Optical Society of America A*, 18(10):2448–2459, October 2001.

[10] R. Ramamoorthi and P. Hanrahan. A signal-processing framework for inverse rendering. In *SIGGRAPH '01: Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, pages 117–128, New York, NY, USA, 2001. ACM Press.

[11] A. Shashua and T. Riklin-Raviv. The quotient image: Class-based re-rendering and recognition with varying illuminations. *IEEE Trans. Pattern Anal. Mach. Intell.*, 23(2):129–139, 2001.

[12] H. Wang, S. Z. Li, and Y. Wang. Generalized quotient image. In *CVPR (2)*, pages 498–505, 2004.