Multiple Directional Illuminant Estimation from a Single Image

Yang Wang, Dimitris Samaras
Computer Science Department,
State University of New York at Stony Brook, NY 11794-4400, USA
{yangwang, samaras}@cs.sunysb.edu

Abstract

We present a new method for the detection and estimation of multiple directional illuminants, using only one single image of an object of arbitrary known geometry. The surface is not assumed to be pure Lambertian, instead, it can have both Lambertian and specular properties. We propose a novel methodology that integrates information from shadows, shading and specularities in the presence of strong directional sources of illumination, even when significant non-directional sources exist in the scene. Since the specular spots have much sharper intensity changes than the Lambertian part, we can locate them in the image of the sphere by first down-sampling the image and then applying a region growing algorithm. Once the specularities have been roughly segmented, the remaining regions of the image are mostly Lambertian, and can be segmented into regions, with each region illuminated by a different set of sources, in a robust way. The regions are separated by boundaries consisting of critical points (points where one illuminant is perpendicular to the normal). Our region-based recursive least-squares method is impervious to noise and missing data. The illuminant estimation can be further refined by making use of shadow information when available, in a novel, integrated single-pass estimation. The method is generalized to objects of arbitrary known geometry, by mapping their normals to a sphere. Furthermore, we introduce a hybrid approach that combines our method with spherical harmonic representations of non-directional light sources, when such sources are present in the scene. We demonstrate experimentally the accuracy of our method, both in detecting the number of light sources and in estimating their directions, by testing on synthetic and real images.

1. Introduction

In order to integrate seamlessly a virtual object in a real scene, we need to simulate accurately the interactions of the virtual object with the illumination of the scene. Furthermore, to manipulate existing images realistically, knowledge of illuminant directions is necessary both in image based computer graphics, and in computer vision for shape reconstruction. This problem is particularly hard for diffuse (Lambertian) surfaces and directional light sources and cannot be solved using local information only. In this work we concentrate on directional light sources because they have the most pronounced effects in the appearance of a scene. Previous methods that estimate multiple light sources require images of a calibration object of given shape (typically spheres) which needs to be removed from the scene and might cause artifacts. Instead, our method relies on partial knowledge of the geometry of the scene and can be used on objects of arbitrary shape. This allows us to possibly use any diffuse object of the scene for illumination calibration. In this paper we present a novel methodology that integrates information from shadows, shading and specularities in the presence of strong directional sources of illumination, even when significant non-directional sources exist in the scene. Specularities are detected based on sharp intensity changes. The shadow-based method utilizes brightness variations inside the shadows cast by the object, whereas the shading-based method utilizes brightness variations on the directly illuminated portions of the object. Here, we demonstrate that they can be cast in an one-pass estimation framework. The proposed integrated method is both more accurate and more general in its applicability, than any of the three methods applied separately. Furthermore, we introduce a hybrid approach that combines our method, with spherical harmonic representations of non-directional light sources (e.g. area sources).

In the last few years, there has been an increased interest in estimating the reflectance properties and the illumination conditions of a scene based on its images, initially recovering a single light source [6, 12, 34, 22]. However, illumination in most real scenes is more complex and it is very likely to have a number of co-existing light sources in a scene. Various types of light sources are discussed in [11]. Accuracy, photorealism and generality of many Image Based Modelling and Rendering (IMBR) applications depends on the knowledge of illumination. A number

of methods were proposed to recover illumination parameters or reflectance properties of the scene in the form of BRDFs (Bidirectional Reflectance Distribution Functions) [3, 32, 31, 23, 21, 16, 8, 25, 4, 17]. Most of these methods require extensive data collection [3, 4, 32, 31] and off-line processing [8, 17], or have particularly restrictive assumptions, e.g. a single light source [22]. Such methods would not be of use if only one or a few images are available.

Most illumination estimation methods need to use a calibration object of fixed shape, often a sphere with a specular component [18, 3]. However, such a specular sphere might have strong inter-reflections with other objects of the scene, especially if they are close to it. Using the Lambertian shading model, in [30] it was observed that multiple light sources can be deduced from boundary conditions, i.e., the image intensity along the occluding boundaries and at singular points. Based on this idea, in [33] it was shown that the illuminant directions have a close relationship to critical points on a Lambertian sphere and that, by identifying most of those critical points, illuminant directions may be recovered if certain conditions are satisfied. Conceptually, a critical point is a point on the surface such that all its neighbors are not illuminated by the same light sources. Recently, an illuminant direction detection method minimizing global error was proposed in [28]. It uses critical boundaries to segment the surface robustly into regions ("virtual light patches"), with each illuminated by a different set of sources. By using points inside a region instead of the boundary, the method's accuracy does not depend on the exact extraction of the boundary and can tolerate noise better. When the observed shape is not spherical, its normals are mapped to a sphere.

The idea of using arbitrary known shape, is also found in the approach of Sato et al. [23], which exploits information of a radiance distribution inside shadows cast by an object of known shape in the scene. Recently, under a signal processing approach [1, 21] a comprehensive mathematical framework for evaluation of illumination parameters through convolution is described. Unfortunately, this framework does not provide a method to estimate high-frequency illumination such as directional light sources when the BRDF is smooth as in the Lambertian case. Convolution is a local operation and the problem is ill-posed when only local information is considered [2]. Our method uses global information to overcome this problem, and in this sense, it is complementary to the methods of [1, 21], which are suitable for non-directional sources. When both types of illuminants are present, we propose a hybrid method that first estimates the directional sources, subtracts their effects from the image and then calculates the spherical harmonic coefficients approximating the non-directional illumination.

Most objects are neither exactly diffuse nor perfectly specular. Instead, for many materials, the surface reflectance can be approximated by a parameterized BRDF with a diffuse and a specular component, such as [26, 29, 10, 5] etc. A number of authors have addressed estimation of illumination from specularities, usually in a multiple view context, [18, 14, 15, 35]. In this work, we are not interested in estimating the illuminant directions from specularities themselves, but only in removing their most pronounced effects from the image, so that we can utilize the diffuse part. This allows us to use an approximation for their segmentation, using only a single image.

In this paper, we propose a new method for multiple directional illuminant estimation, that integrates illumination information from shading [28], shadows [23] and specularities. These methods have different strengths and weaknesses. We demonstrate how the sources of information complement each other in a number of occasions. Combining the methods reduces error and speeds up computation. We demonstrate the effectiveness of our method with synthetic and real image experiments.

2. Shading-Based Illuminant Detection

2.1. Critical Points

Definition 1 *Given an image, let* $\mathbf{L}_i$, $i = 1, 2, \ldots$, *be the light sources of the image. A point in the image is called a critical point if the surface normal at the corresponding point on the surface of the object is perpendicular to some light source* $\mathbf{L}_i$.

We assume that images are formed by perspective or orthographic projection and the object in the image has a Lambertian surface with constant albedo, that is the BRDF is known to be a constant and each surface point appears equally bright from all viewing directions: $E = \rho L_i \mathbf{L} \cdot \mathbf{N} = \rho L_i cos\theta_i$, where E is the scene radiance of an ideal Lambertian surface, ρ is the albedo, $\mathbf{L}$ represents the direction and L_i the amount of incident light, and $\mathbf{N}$ is the unit normal to the surface. Initially, the algorithm is developed using a sphere model of known size and subsequently extended to objects of arbitrary shape.

As shown in [33], it is impossible to recover the exact value of the intensity of any individual light source among four (or more) pairs of antipodal light sources (i.e. opposite direction light sources). However, this happens rarely, so we assume that there are no antipodal light sources.

Let $\mathbf{P}$ be an arbitrary plane such that $\mathbf{S}$, the center of the sphere, lies on it (Fig. 1(a)), $\mathbf{L}_i$ be the light sources of the image and $(\mathbf{L}_i)_P$ their projections on $\mathbf{P}$. A point on the arc τ can be specified by its corresponding angle parameter in $[\alpha, \beta]$ using the following proposition [33]:

Proposition 1 *Consider an angle interval* $[\alpha, \beta]$ *of a sphere cross section (Fig. 1(a)). We can always find a partition*

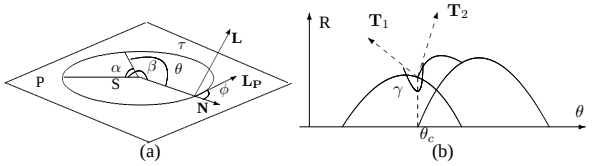


Figure 1. (a) $\mathbf{L}$ and its projection $\mathbf{L_P}$ onto plane $\mathbf{P}$. (b) inner angle γ between two tangent lines.

$\theta_0 = \alpha < \theta_1 < \ldots < \theta_n = \beta$ of the interval $[\alpha, \beta]$ such that in each $[\theta_{i-1}, \theta_i]$ we have $E(\theta) = b_i \sin\theta + c_i \cos\theta$ for some constants b_i and c_i, $1 \leq i \leq n$, where $E(\theta)$ is the intensity function along the arc τ.

Intuitively, (b_{i-1}, c_{i-1}) represents the virtual light source of the $[\theta_{i-2}, \theta_{i-1}]$ part, and (b_i, c_i) of the neighboring $[\theta_{i-1}, \theta_i]$ part. The virtual light source of an interval $[\theta_{i-2}, \theta_{i-1}]$ in the Lambertian case is the sum of all light sources illuminating $[\theta_{i-2}, \theta_{i-1}]$. As shown in [33] the distance between (b_{i-1}, c_{i-1}) and (b_i, c_i),

$$\sqrt{(b_i - b_{i-1})^2 + (c_i - c_{i-1})^2} = \sum_{j \in \Lambda' \cup \Lambda''} \|(\mathbf{L}_j)_P\| \quad (1)$$

will be maximized at a critical point for these two virtual light sources. Unfortunately, this criterion greatly depends on the intensities of virtual light sources. Together with the distance between (b_i, c_i) pairs, we can use the tangent angles defined on the intensity curve (Fig. 1(b)). By applying a standard recursive least-squares algorithm, we can use two consecutive windows to detect the local maximum points of inner angles γ and distance defined by Eqn.(1) [28].

2.2. Segmenting the Surface

Definition 2 *All critical points corresponding to one real light will be grouped into a cut-off curve which is called a critical boundary.*

Intuitively, each critical boundary of the sphere in our model is on a cross section plane through the center of the sphere. Therefore, critical points can be grouped into critical boundaries using the Hough transform. Although critical points provide information to determine the light source directions [33], they are relatively sensitive to noise. However, non-critical point areas are less sensitive to noise and provide important information to determine the light source directions.

Definition 3 *Critical boundaries will segment the whole sphere image into several regions, and intuitively, each segmented region is corresponding to one virtual light. Each region is called a virtual light patch.*

Once we get the patches corresponding to each virtual light, the directions of virtual light sources can be calculated, via a least-squares method, using the internal non-critical points of each patch.

2.3. Recovering the True Lights

Proposition 2 *If a critical boundary separates a region into two virtual light patches with one virtual light each, e.g. $\mathbf{L}_1$, $\mathbf{L}_2$, then the difference vector between $\mathbf{L}_1$ and $\mathbf{L}_2$, $\mathbf{L}_{pre} = \mathbf{L}_1 - \mathbf{L}_2$, is called the real light pre-direction with respect to this critical boundary. Since we have already assumed that there are no antipodal light sources (i.e. opposite direction light sources), the real light direction will be either the pre-direction $\mathbf{L}_1 - \mathbf{L}_2$, or its opposite $\mathbf{L}_2 - \mathbf{L}_1$.*

To find out the true directions, we pick a number of points on the surface, e.g. $P_1, ..., P_k$ and their normals, e.g. $\mathbf{N}_1, ..., \mathbf{N}_k$, then the true directions will be the solution of: $E(P_j) = \sum_{i \in \Lambda} max(e_i \mathbf{L}_i \cdot \mathbf{N}_j, 0) + \mathbf{L}_v \cdot \mathbf{N}_j, 1 \leq j \leq k$, where $\mathbf{L}_v$ is the virtual light source of a possible frontal illuminant whose critical boundary can not be detected and will be checked as a special case. Selecting points in the area inside the critical boundaries is a robust way to detect real lights. This can be done using standard least-squares methods [19]. An angle threshold is introduced to cluster the light difference vectors into real light groups, each is approximated by one vector. By minimizing the least-squares errors of virtual light patches, we are able to merge the spurious critical boundaries detected by the Hough transform.

2.4. Arbitrary Shape

In this section we extend our method to work with any object of known shape. Obviously, there should exist enough non-coplanar points on the object illuminated by each light to allow for a robust least-squares solution. We assume no inter-reflections. We map the image intensity of each point P_i of the arbitrary shape to a point S_i of a sphere, so that the normal at P_i is the same as the normal at S_i. We detect all potential critical points based on the points mapped on the sphere. As expected, not every point on the surface of the sphere will be corresponding to a normal on the surface of the arbitrary shape, so there will be many holes on the mapped sphere, e.g. the black area in Fig. 2. Thus, many critical points' locations will be erroneously calculated even for noise-free data. Consequently, the critical boundaries calculated by the Hough transform based on these critical points might not be correct. Since we can not recover these missing data from the original image, it is impossible to adjust the critical boundaries detected by the Hough transform itself. On the other hand, as long as the critical boundaries are not too far from the truth, the majority of the points in a virtual patch will still correspond to the correct virtual light (especially after the adjustments steps described in Sec. 2.3). Thus it is still possible, using sparse points on the sphere, to calculate the true light for each virtual light patch based on Proposition 2. If two points have the same normal but different intensities, we use the brighter one (assuming that the other is in shadow).

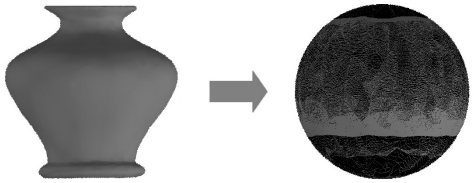


Figure 2. Vase and its sphere mapping. Both image sizes are 400 by 400. Black points on the sphere represent normals that do not exist on the vase's surface.

3. Generalized Shading-Based Method

The original algorithm of our method assumes directional light sources and perfect Lambertian surfaces, however, in most real world situations there are significant deviations from these assumptions. Here we extend our method to allow for specular surfaces and combinations of directional and area light sources.

3.1. Non-Lambertian Surfaces

Here we extend the applicability of our method to surfaces which have both Lambertian and specular properties. Instead of the Dichromatic Refection Model [9], we use Ward's parameterization in which the BRDF is modeled as the sum of a diffuse term $\frac{\rho_d}{\pi}$ and a specular term $\rho_s K(\mathbf{L},\mathbf{V},\mathbf{\Sigma})$, where ρ_d and ρ_s are the diffuse and specular reflectance of the surface respectively, and $K(\mathbf{L},\mathbf{V},\mathbf{\Sigma})$ is a function of light and viewing directions, parameterized by $\mathbf{\Sigma}$, the surface roughness vector. Assuming the surface is isotropic, the detailed Ward model can be simplified as,

$$\begin{aligned} L_i &= (\tfrac{\rho_d}{\pi} + \rho_s K(\mathbf{L},\mathbf{V},\mathbf{\Sigma}))I_i \\ &= \tfrac{\rho_d}{\pi} I_i + \rho_s \tfrac{1}{\sqrt{(\mathbf{L}\cdot\mathbf{N})(\mathbf{V}\cdot\mathbf{N})}} \cdot \tfrac{e^{\frac{-\tan^2 \angle(\mathbf{H},\mathbf{N})}{\sigma^2}}}{4\pi\sigma^2} I_i \end{aligned} \quad (2)$$

where L_i and I_i are the radiance and irradiance on the point P_i respectively, $\mathbf{H}$ is the bisector of the light and viewing directions and σ is the surface roughness measured as the standard deviation of the surface slope. The treatment for the Torrance-Sparrow model is similar.

From Eqn.(2), we notice that $e^{\frac{-\tan^2 \angle(\mathbf{H},\mathbf{N})}{\sigma^2}}$ plays a dominant role in the specular component, which is plotted in the 0-255 gray scale range for different values of σ (Fig. 3), and depends heavily on the angle α between $\mathbf{H}$ and $\mathbf{N}$.

As can be seen in Fig. 3, the specular component has a much sharper intensity change than the diffuse part, and for values of the angle α greater than a threshold around $10^\circ - 15^\circ$, it becomes less than 1 gray scale level. We refer to this threshold as $\alpha_{specular}$, pre-calculated based on σ.

Proposition 3 *Let α be the angle between $\mathbf{H}$ and $\mathbf{N}$, $\angle(\mathbf{H},\mathbf{N})$, and k_g be the gray scale coefficient, i.e. $k_g = \rho_s \frac{1}{\sqrt{(\mathbf{L}\cdot\mathbf{N})(\mathbf{V}\cdot\mathbf{N})}} \cdot \frac{1}{4\pi\sigma^2} \cdot I_i$, $L_i = \frac{\rho_d}{\pi} I_i + k_g e^{\frac{-\tan^2 \angle(\mathbf{H},\mathbf{N})}{\sigma^2}}$. Since intensities in the image less than 1 are not observable, we have $k_g e^{\frac{-\tan^2 \alpha}{\sigma^2}} \geq 1$, that is, $\tan\alpha \leq \sigma\sqrt{\ln k_g}$. Thus, $\alpha_{specular} = \arctan(\sigma\sqrt{\ln k_g})$.*

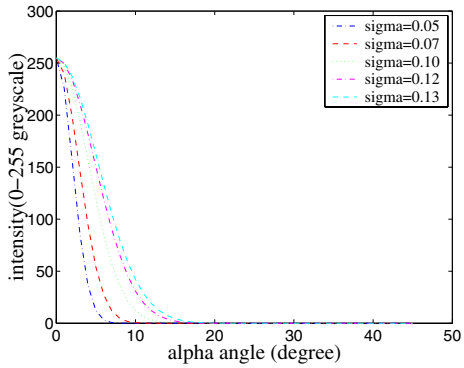


Figure 3. The exponential part of the specular component in the Ward BRDF (Eqn. 2).

Based on this observation, we can locate specularities in the image of a sphere by first down-sampling the image and then applying a region growing algorithm. Since there are many more peaks on an image than specular spots, down-sampling is used to locate the specular spots more efficiently by taking advantage of the approximately known BRDF parameters. In the first stage, we down-sample the original image to such a level that a region of the size of $\alpha_{specular}$ (typically $\arctan(140\sigma)$ degrees), is within 1-2 pixel resolution size. Then, we use a 3 by 3 convolution kernel to detect the sharp intensity change above $threshold_{sharp}$ and then locate the local maximum within each detected region as a potential specular spot. In the second stage, we grow the potential specular spots in the original image until they reach the size defined by $\alpha_{specular}$. These regions are treated as potential specular segments and the remaining regions of the image are mostly Lambertian. Because our region-based recursive least-squares method is impervious to noise and missing data, we can apply it to the segmented image and estimate the illuminants. We are interested in a rough segmentation of the specular area, so our method works even if the reflectance model or the specular peak are not extremely accurate.

3.2. Non-Directional Illumination Sources

Unlike directional light sources, broad area illumination sources produce low-frequency shading which does not generate clear critical boundaries on diffuse surfaces. Recently, under a signal processing approach [1, 21], a comprehensive mathematical framework for evaluation of illumination parameters through convolution is described based on the spherical harmonics Y_n^m

$$\begin{aligned} P_n^m(z) &= \tfrac{(1-z^2)^{\frac{m}{2}}}{2^n n!} \tfrac{d^{n+m}}{dz^{n+m}} (z^2-1)^n \\ Y_n^m(\theta,\phi) &= \sqrt{\tfrac{2n+1}{4\pi} \tfrac{(n-m)!}{(n+m)!}} P_n^m(\cos\theta) e^{Im\phi} \end{aligned} \quad (3)$$

where P_n^m are the associated Legendre functions.

As shown in [1, 20], the set of Lambertian reflectance functions can be well approximated by a low dimensional linear space, based on the fact that the Lambertian kernel has most of its energy (more than 99%) concentrated

in the first 9 terms and are given numerically in the cartesian coordinates $(x, y, z) = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$, $Y_0^0 = \sqrt{\frac{1}{4\pi}}$, $(Y_1^1; Y_1^0; Y_1^{-1}) = \sqrt{\frac{3}{4\pi}}(x; z; y)$, $Y_2^0 = \frac{1}{2}\sqrt{\frac{5}{4\pi}}(3z^2 - 1)$, $(Y_2^1; Y_2^{-1}; Y_2^{-2}) = 3\sqrt{\frac{5}{12\pi}}(xz; yz; xy)$, $Y_2^2 = \frac{3}{2}\sqrt{\frac{5}{12\pi}}(x^2 - y^2)$. Then the irradiance E and the distant lighting distribution L will be approximated by a linear combination of the spherical harmonics, $E(\theta, \phi) = \sum_{n,m} E_n^m Y_n^m(\theta, \phi)$, $L(\theta, \phi) = \sum_{n,m} L_n^m Y_n^m(\theta, \phi)$, and the relationship between E_n^m and L_n^m is given by $E_n^m = \sqrt{\frac{4\pi}{2n+1}} A_n L_n^m$, where $A_1 = \sqrt{\frac{\pi}{3}}$, $A_{n>1,odd} = 0$ and $A_{n>1,even} = 2\pi\sqrt{\frac{2n+1}{4\pi}}\frac{(-1)^{\frac{n}{2}-1}}{(n+2)(n-1)}\left[\frac{n!}{2^n(n!/2)^2}\right]$.

Unfortunately, this framework does not provide a method to estimate high-frequency illumination such as directional light sources when the BRDF is smooth as in the Lambertian case. Convolution is a local operation and the problem is ill-posed when only local information is considered [2]. Here we propose a hybrid approach that combines the signal processing approach for broad area illumination sources with the method of Sec. 2. Assuming the directional light sources play a non-negligible role in the lighting distribution, we use the following algorithm to estimate the lighting distribution:

1. Estimate directional illuminants as described in Sec.2.

2. Subtract the effects of directional illuminants from the image which will then contain mostly area source illumination information.

3. Calculate the spherical harmonic coefficients as a representation of the area sources.

4. The unaccounted information in the residual image is error from two sources, the spherical harmonic approximation and errors in the initial estimation of the directional illuminants. We minimize this error by small adjustments in the estimates of the directional illuminants.

4. Shadow-Based Illuminant Detection

Besides the shading information we explored above, a picture of a real scene is very likely to contain some shadow information. Hence the illumination distribution of the scene might also be recovered from a radiance distribution inside shadows cast by an object of known shape onto another object surface of known shape and reflectance. In [23], the illumination distribution of a scene is approximated by discrete sampling of an extended light source and the whole distribution is represented as a set of point sources equally distributed in the scene as shown in Fig. 4. The total irradiance E at the shadow surface received from the entire illumination distribution is computed by $E = \sum_{i=1}^{n} L_i S_i \cos\theta_i$, where $L_i (i = 1, 2, \ldots, n)$ is the illumination radiance per solid angle $\delta = 2\pi/n$ coming from the direction (θ_i, ϕ_i), and S_i are occlusion coefficients. $S_i = 0$ if L_i is occluded by objects, and $S_i = 1$ otherwise. Then

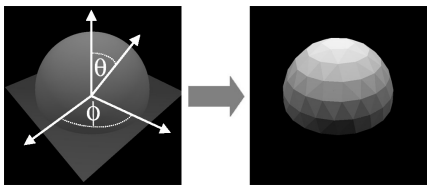


Figure 4. Illumination distribution is approximated by discrete sampling over the entire surface of the extended light source.

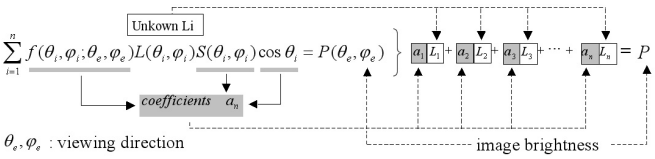


Figure 5. Each shadow pixel provides a linear equation for estimating illumination distribution by shadows.

this approximation leads each image pixel inside shadow regions to provide a linear equation where the radiance of the point sources is unknown, as shown in Fig. 5 [24]. Finally, a set of linear equations is derived from the brightness changes observed in the shadow image and solved for unknown L_i's.

The BRDF $f(\theta_i, \phi_i; \theta_e, \phi_e)$ for a Lambertian surface is known to be a constant. Then, in Fig. 5 the coefficients $a_i (i = 1, 2, \ldots, n)$ represent $K_d \cos\theta_i S_i$ where K_d is a diffuse reflection parameter of the surface. Therefore, by selecting a sufficiently large number of image pixels, it is possible to solve for a solution set of unknown L_i's.

5. Integration of Information Sources

In this section, we are going to propose a framework that combines the respective advantages of shading and shadow information, allowing us to obtain improved results compared to using each of them independently.

In the case of non-Lambertian reflectance, specularities need to be removed first. As can be seen from Fig. 2, arbitrary shapes do not always provide enough normals on the surface to make a complete sphere mapping, so many data points on the sphere will be missing non-uniformly. Consequently, there is a possibility that some critical boundaries will be lost and the corresponding real lights will not be estimated. However, shadow information can be used to estimate the intensity and direction of each light source.

While recovering the illumination distribution of the scene from a radiance distribution inside shadows, complete shadows cast by an object of known shape onto another object surface of known shape and reflectance are required. However, this might not be possible in situations where the light direction is nearly parallel to the surface. Then the azimuth of the light source can still be estimated reliably but not the elevation. Furthermore, in [23] a large number of samples are needed to capture the rapid change of radiance distribution around a direct light source. Radiance distribution inside a direct light source has to be sam-

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

Figure 6. (a) A synthetic vase illuminated by three directional light sources. (b) Estimated illumination distribution using the shadow information only. (c) Error image generated by illumination distribution estimated in (b). (d) Detected critical boundaries using the shading information only. (e) Error image generated by illumination distribution estimated in (d).

pled densely. Therefore, due to the discrete sampling of the geodesic dome, it is very likely that the precision of estimation will be limited. In Fig. 6, we can see that the illumination distribution estimated by shading information provides higher accuracy than the one estimated by shadows.

A shadow is called a complete shadow when all the parts of the scene the shadow falls on are visible. The outmost edge of a complete shadow corresponding to a directional light source is generated by the occluding boundary of the object surface. As we can see, the occluding boundary of a smooth surface will be a critical boundary in the context of shading. Consequently, when there is information both from shading and from shadows, we can use the shadow information to give us an initial estimation of the illumination distribution, compute the corresponding critical points and then use the shading information to refine this estimation to compute the directions and intensities of the real light sources.

Incorporating specularity and shadow information, the complete proposed algorithm is:

1. Estimate specularities and remove them from the image, as described in Sec. 3.1.

2. Detect critical points using the method in Sec. 2.1.

3. Calculate an initial illumination distribution using estimation from shadows [23].

4. For each light direction detected in Step 3, add the corresponding critical points in the results from step 2.

5. Find initial critical boundaries by the Hough transform based on all detected critical points.

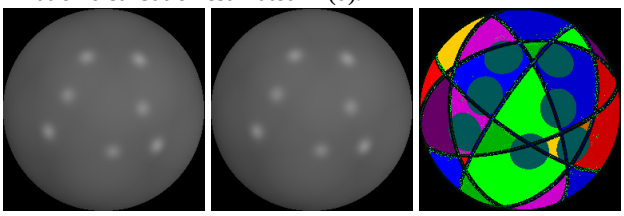
6. Adjust critical boundaries. Adjust every critical boundary by moving it by a small step, with a reduction in the least-squares error indicating an improved solution. Keep updating boundaries using a "greedy" algorithm in order to minimize the total error.

7. Merge spurious critical boundaries. If two critical boundaries are closer than a threshold angle $T_{mergeangle}$ (e.g. $5°$), they can be replaced by their average, resulting into one critical boundary instead of two.

8. Remove spurious critical boundaries. Test every critical boundary, by removing it temporarily and if the least-squares error does not increase, we can consider it a spurious boundary and remove it completely. Test boundaries in increasing order of Hough transform votes (intuitively we test first boundaries that are not as trustworthy).

9. Calculate the real lights along a boundary by subtracting neighboring virtual lights as described in Proposition 2.

Step 4 adds more votes close to the true direction for the Hough transform, which will reduce the adjustment time significantly, comparing to a previous implementation where shading and shadows were estimated separately in a two-pass approach [27].



(a)original image (b)generated image (c)critical boundaries

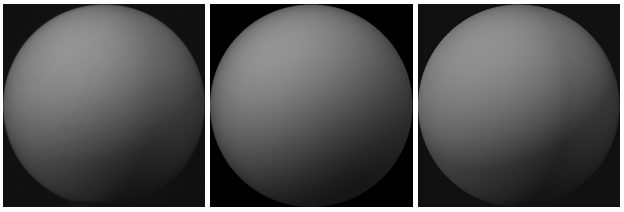
Figure 7. Synthetic sphere image: (a) a synthetic ball illuminated by seven light sources using the Ward BRDF model. Image size: 256x256. (b) the rerendered image with the seven light sources extracted from (a). Green dots on (c) represent the detected critical points.

6. Experiments

6.1. Specularity Detection

We performed a number of experiments with synthetic sphere images to better understand our method. We did experiments on images of seven random light sources. The following parameter values were chosen for the algorithm: image size is 256 by 256, sliding window width $w = 12$ pixels (approximately $8°$), distance ratio $T_{ratio} = 0.5$ and angle threshold for boundary merging (described in Sec. 2.3) $T_{mergeangle} = 5°$. Results are shown in Fig. 7. In all the segmented light patch images in this report, detected critical points are denoted with green dots.

As an example of a real object with specularities, we used a white clay vase illuminated by three light sources. The original image and 3D geometry were captured by the range scanner system described in [7]. The following parameter values were used: sliding window width $w = 32$ pixels (approximately $14°$), distance ratio $T_{ratio} = 0.6$ and angle threshold for boundary merging $T_{mergeangle} = 5°$. Although the surface did not conform very closely to the Ward model, the correct number of light sources was de-



(a)original image generated by (b) 9 spherical harmonics (c)our hybrid method

Figure 8. Real sphere image: a white plastic ball with three light sources. Image size: 460x460. (a) the original image, the generated image of a sphere using (b) only the first 9 spherical harmonics and (c) our hybrid method to estimate the lighting distribution extracted from (a). Although hard to discern critical boundary information is present in image (a) and (c) but gets lost in image (b) due to its high frequency.

tected and the estimated lights appear to be very close to the truth (no ground data was available). Results are shown in Fig. 9(e-g).

6.2. Non-Directional Illumination Estimation

We used a rubber ball illuminated with two directional light sources and one broad area light source at the same time. The original photo, taken by an SONY DFW-X700 digital camera, was cropped to 460 by 460 pixels with the ball at the center of the image. The color image is converted to gray-scale with 256 levels. The following parameter values were used: sliding window width $w = 12$ pixels (approximately 5°), distance ratio $T_{ratio} = 0.5$ and angle threshold for boundary merging $T_{mergeangle} = 5^\circ$. Results are shown in Fig. 8. The recovered directional light sources are $\mathbf{L}_1 = 64.94 \times (-0.87, 0.04, 0.50)$, $\mathbf{L}_2 = 57.73 \times (-0.34, 0.77, 0.54)$. Comparing to estimation from single light images of the same illuminants which were captured only for evaluation purposes, $\mathbf{L}_1 = 64.71 \times (-0.84, 0.07, 0.54)$, $\mathbf{L}_2 = 60.75 \times (-0.29, 0.77, 0.57)$, angle errors are 3.2° and 3.7°, and the intensity errors are 0.2 and 3.0 gray-scale levels. Although the surface did not conform very closely to the Lambertian model, the correct number of light sources was detected and the estimated lights appear to be reasonably close to the truth. The average intensity error is 1.14 gray-scale levels between the original image and the image generated by the first 9 spherical harmonics (Fig. 8(b)), and 0.87 by our hybrid method (Fig. 8(c)).

6.3. Mixed Reality Image Synthesis

The combination of shading and shadow information can provide better estimation of illumination distribution. These estimates can be used to synthesize Mixed Reality images, i.e. real images with superimposed virtual objects correctly lit. Furthermore, under the assumptions of Lambertian BRDF and known geometry, we can re-render the real images to generate new images by modifying the estimated illumination configuration. Based on a scene containing two rubber toys illuminated by three light sources, we generated a new image where one light has been switched off in Fig. 9(b), which can be compared with a real image of the scene with the same light truly switched off in Fig. 9(c). The original image is 1534 x 1024 pixels with the two toys at the center of the image. To demonstrate the ability of our algorithm to use only partial scene information for accurate estimation, only the duck toy was used to estimate the illuminant directions. The second toy is used for visual evaluation of the results. Based on the size of the duck, the diameter of the mapping sphere is 400. The following parameter values were chosen for the algorithm: sliding window width $w = 30$ pixels (approximately 13.5°), distance ratio $T_{ratio} = 0.5$ and angle threshold for boundary merging $T_{mergeangle} = 5^\circ$.

7. Conclusions and Future Work

In this paper we presented a method for the estimation of multiple illuminant directions from a single image, incorporating specularity, shadow and shading information. We demonstrated how information from each source enhances the information from the others. We do not require the imaged scene to be of any particular geometry (e.g. a sphere). This allows our method to be used with the existing scene geometry, without the need for special light probes when the illumination of the scene contains directional light sources. The original algorithm of our method is extended to allow for specular surfaces and combinations of directional and area light sources, assuming that the directional light sources play a non-negligible role in the lighting distribution. Experiments on synthetic and real data show that the method is robust to noise. We apply the results of our method to generate Mixed Reality images, by successfully modifying scene illumination and seamlessly re-rendering, including superimposed synthetic objects. Future work includes study of the properties of arbitrary surfaces (so that we can avoid the intermediate sphere mapping) and modelling of inter-reflections.

References

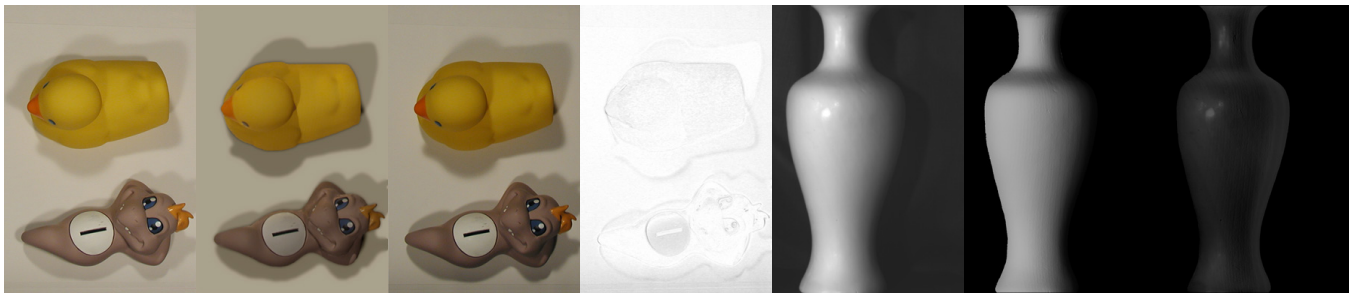
[1] R. Basri and D. Jacobs. Lambertian reflectance and linear subspaces. *ICCV*, pages 383–390, 2001.

[2] W. Chojnacki, M. J. Brooks, and D. Gibbins. Can the sun's direction be estimated prior to the determination of shapes? *Australian Jnt. Conf. on A.I.*, pages 530–535, 1994.

[3] P. Debevec. Rendering synthetic objects into real scenes. *SIGGRAPH*, pages 189–198, 1998.

[4] P. Debevec, T. Hawkins, C. Tchou, H. Duiker, W. Sarokin, and M. Sagar. Acquiring the reflectance field of a human face. *SIGGRAPH*, pages 145–156, 2000.

[5] T. K.-S. F. He, X.D. and D. Greenberg. A comprehensive physical model for light reflection. *SIGGRAPH*, pages 175–186, 1991.



(a)original image (b)generated image (c)real image (d)error image (e)original image (f)generated image (g)error image

Figure 9. Real arbitrary shape image experiment: a scene illuminated by three light sources. Image Size: 1534x1024. (a) the original image, (b) the generated image of a scene with the two light sources extracted from (a), (c) the real image of the scene illuminated by the two real lights, (d) the error image: darker color means higher error. The noise in the generated image is mainly due to the inaccuracies in the estimation of shape and the edges of each shadow. Nonetheless illuminant estimation is still possible. A clay vase illuminated by three light sources. Image Size: 1534x1024. (e) the original image, (f) the generated image of a scene with the three light sources extracted from (e) after specularity removal, (g) the error image: lighter color means higher error.

[6] B. Horn and M. Brooks. Shape and source from shading. *IJCAI*, pages 932–936, 1985.

[7] Q. Hu. *3-D Shape Measurement Based on Digital Fringe Projection and Phase-Shifting Techniques*. PhD thesis, M.E. Dept., SUNY at Stony Brook, 2001.

[8] T. Kim, Y. Seo, and K. Hong. Improving ar using shadows arising from natural illumination distribution in video sequences. *ICCV*, pages II: 329–334, 2001.

[9] G. Klinker, S. Shafer, and T. Kanade. The measurement of highlights in color images. *IJCV*, 2(1):7–32, 1988.

[10] F. S.-C. T. K. Lafortune, E.P.F. and D. Greenberg. Non-linear approximation of reflectance functions. *SIGGRAPH*, pages 117–126, 1997.

[11] M. Langer and S. Zucker. What is a light source? *CVPR*, pages 172–178, 1997.

[12] C. Lee and A. Rosenfeld. Improved methods of estimating shape from shading using the light source coordinate system. *AI*, 26(2):125–143, 1985.

[13] H. Lin and M. Subbarao. A vision system for fast 3d model reconstruction. *CVPR*, pages 663–668, 2001.

[14] S. Lin and S. Lee. Estimation of diffuse and specular appearance. *ICCV*, pages 855–860, 1999.

[15] S. Lin, Y. Li, S. Kang, X. Tong, and H.-Y. Shum. Diffuse-specular separation and depth recovery from image sequences. pages III:210–224, 2002.

[16] S. Marschner and D. Greenberg. Inverse lighting for photography. *Fifth Color Imaging Conference*, pages 262–265, 1997.

[17] S. Marschner, S. Westin, E. Lafortune, and K. Torrance. Image-based brdf measurement. *Applied Optics*, pages 2592–2600, 2000.

[18] M. Powell, S. Sarkar, and D. Goldgof. A simple strategy for calibrating the geometry of light sources. *PAMI*, 23(9):1022–1027, 2001.

[19] W. Press, S. Teukolsky, W. Vettering, and B. Flannery. *Numerical Recipes in C*. Cambridge Univ. Press, 1992.

[20] R. Ramamoorthi and P. Hanrahan. An efficient representation for irradiance environment maps. *SIGGRAPH*, pages 497–500, 2001.

[21] R. Ramamoorthi and P. Hanrahan. A signal-processing framework for inverse rendering. *SIGGRAPH*, pages 117–128, 2001.

[22] D. Samaras and D. Metaxas. Coupled lighting direction and shape estimation from single images. In *ICCV*, pages 868–874, 1999.

[23] I. Sato, Y. Sato, and K. Ikeuchi. Illumination distribution from brightness in shadows. *ICCV*, pages 875–883, 1999.

[24] I. Sato, Y. Sato, and K. Ikeuchi. Stability issues in recovering illumination distribution from brightness in shadows. *CVPR*, pages 400–407, 2001.

[25] Y. Sato, M. Wheeler, and K. Ikeuchi. Object shape and reflectance modeling from observation. *Computer Graphics*, pages 379–388, 1997.

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

[27] Y. Wang and D. Samaras. Estimation of multiple directional light sources for synthesis of mixed reality images. In *Pacific Graphics*, pages 38–47, 2002.

[28] Y. Wang and D. Samaras. Estimation of multiple illuminants from a single image of arbitrary known geometry. In *ECCV*, pages III: 272–288, 2002.

[29] G. Ward. Measuing and modeling anisotropic reflection. *SIGGRAPH*, pages 265–272, 1992.

[30] Y. Yang and A. Yuille. Sources from shading. *CVPR*, pages 534–439, 1991.

[31] Y. Yu, P. Debevec, J. Malik, and T. Hawkins. Inverse global illumination: Recovering reflectance models of real scenes from photographs. *SIGGRAPH*, pages 215–224, 1999.

[32] Y. Yu and J. Malik. Recovering photometric properties of architectural scenes from photographs. *SIGGRAPH*, pages 207–217, 1998.

[33] Y. Zhang and Y. Yang. Illuminant direction determination for multiple light sources. *CVPR*, pages I:269–276, 2000.

[34] Q. Zheng and R. Chellappa. Estimation of illuminant direction, albedo, and shape from shading. *PAMI*, 13(7):680–702, 1991.

[35] W. Zhou and C. Kambhamettu. Estimation of illuminant direction and intensity of multiple light sources. *ECCV*, page IV:206 ff., 2002.