



ACADEMIC
PRESS

Available at
www.ComputerScienceWeb.com
POWERED BY SCIENCE @ DIRECT®

Graphical Models 65 (2003) 185–205

Graphical Models

www.elsevier.com/locate/gmod

Estimation of multiple directional light sources for synthesis of augmented reality images

Yang Wang* and Dimitris Samaras

*Computer Science Department, State University of New York at Stony Brook, Stony Brook,
NY 11794-4400, USA*

Received 10 February 2003; accepted 25 February 2003

Abstract

We present a new method for the detection and estimation of multiple directional illuminants, using a single image of any object with known geometry and Lambertian reflectance. We use the resulting highly accurate estimates to modify virtually the illumination and geometry of a real scene and produce correctly illuminated Augmented Reality images. Our method obviates the need to modify the imaged scene by inserting calibration objects of any particular geometry, relying instead on partial knowledge of the geometry of the scene. Thus, the recovered multiple illuminants can be used both for image-based rendering and for shape reconstruction. Our method combines information both from the shading of the object and from shadows cast on the scene by the object. Initially, we use a method based on shadows and a method based on shading independently. The shadow-based method utilizes brightness variation inside the shadows cast by the object, whereas the shading-based method utilizes brightness variation on the directly illuminated portions of the object. We demonstrate how the two sources of information complement each other in a number of occasions. We then describe an approach that integrates the two methods, with results superior to those obtained if the two methods are used separately. The resulting illumination information can be used (i) to render synthetic objects in a real photograph with correct illumination effects, and (ii) to virtually re-light the scene.

© 2003 Elsevier Science (USA). All rights reserved.

Keywords: Texture; Shading and color; Calibration; Surface geometry; Illumination estimation

* Corresponding author.

E-mail address: yangwang@cs.sunysb.edu (Y. Wang).

1. Introduction

In order to integrate seamlessly a virtual object in a real scene (i.e., synthesize an Augmented Reality image [1]), we need to simulate accurately the interactions of the virtual object with the illumination of the scene. Furthermore, to manipulate realistically existing images, 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. Lambertian reflectance is the most common type of reflectance. In this work we concentrate on directional light sources because they have the most pronounced effects in the appearance of a scene and any errors in their estimation will cause noticeable errors and inconsistencies in the resulting Augmented Reality images. Previous methods that estimate multiple light sources require images of a calibration object of given shape (typically a sphere) 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 in the scene for illumination calibration. We present a new method that integrates information from shadows and shading in the presence of strong directional sources of illumination. 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. The proposed integrated method is both more accurate and more general in its applicability, than the two methods applied separately.

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 images of the scene. The interest in computing illuminant directions first arose from shape from shading applications, and focused on recovering a single light source [7,12,20,31]. 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. An early attempt to recover a more general illumination description [8], modeled multiple light sources as a polynomial distribution. A discussion of the various types of light sources can be found in [11]. With the advent of Image Based Modeling and Rendering (IMBR) methods in Computer Graphics, it quickly became apparent that accuracy, photorealism, and generality of many IMBR applications depends on the knowledge of such properties and parameters. As a result a number of methods were proposed, which recovered illumination parameters or reflectance properties of the scene in the form of BRDFs (Bidirectional Reflectance Distribution Functions) [4,5,10,14–16,19,21,23,25,28,29]. Most of these methods are geared towards the production of high quality images, requiring extensive data collection with the use of specialized equipment [4,5,28,29] and off-line processing [10,16], or have particularly restrictive assumptions, e.g., a single light source [21]. Such methods would not be of use if only one or a few images of the scene are available. More relevant work is the system proposed by [14], aimed at interactive relighting of indoor scenes, requiring knowledge of complete scene geometry and using fast radiosity methods. Radiosity computa-

tions in a complex outdoor scene, even if accurate geometry was known, would be prohibitively slow.

A domain of computer graphics research that has recently attracted increased interest, Augmented Reality (AR) is a variation of Virtual Environments (VE) with the following three characteristics: (1) combines real and virtual, (2) interactive in real time, (3) real and virtual objects are registered in 3-D [1]. To achieve a high quality synthesized image in an Augmented Reality system, three aspects have to be taken into account: geometry, illumination, and real-time performance [22]. In particular, estimation of illumination parameters from images is necessary, in order to compensate for illumination artifacts, and also to allow super-imposition of synthetic images of new objects into real scenes. Most such methods need to use a calibration object of fixed shape, typically a sphere. In [17] a calibration object that comprises of diffuse and specular parts is proposed. In [4] a specular sphere is used as a light probe to measure the incident illumination at the location where synthetic objects will be placed in the scene. Such a sphere though might have strong inter-reflections with other objects in the scene, especially if they are close to it. Using the Lambertian shading model, Yang and Yuille [27] 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, Zhang and Yang [30] show 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. However, because the detection of critical points is sensitive to noise, the direction of extracted real lights is not very robust to noisy data. Recently, an illuminant direction detection method that minimizes the global error was proposed by Wang and Samaras [26]. In general, each point of a surface is illuminated by a subset of all the directional light sources in the scene. The method segments the surface robustly into regions (“virtual light patches”), with each region illuminated by a different set of sources. Then, real lights can be extracted, based on the segmented “virtual light patches” instead of critical points that are relatively sensitive to noise. Since there are more points in a region than on the boundary, the method’s accuracy does not depend on the exact extraction of the boundary and can tolerate noisy and missing data better. When the observed shape is not spherical, its normals are mapped to a sphere (for an example see Fig. 7), although a lot of normals will be missing. However the method works well even for incomplete spheres, as long as there are enough points inside each light patch for the least-squares method to work correctly.

Inserting calibration objects in the scene complicates the acquisition process, as they either need to be physically removed before re-capturing the image, which is not always possible, or they need to be electronically removed as a post processing step, which might introduce artifacts in the image. Our proposed method can be applied to objects of known arbitrary geometry, as long as that shape contains a fairly complete set of normals for a least-squares evaluation of the light sources. Thus, it would be possible to estimate the illuminants from the image of a scene, using

geometry that is part of the scene. Although estimating the object geometry from the image would be more convenient, current single image computer vision methods do not provide the necessary accuracy for precise illuminant estimation. Such accuracy can be achieved only by more complex multi-image methods [13] that require precise calibration. The idea of using arbitrary known shape can also be 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 [2,19], 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 [3]. Our method uses global information to overcome this problem, and in this sense, it is complementary to the methods of [2,15,19].

In this paper, we propose a new method for multiple directional source estimation, that integrates illumination estimation from shading [26] and shadows [23]. Both methods rely on knowledge of the illuminated geometry but do not require a specific calibration object. However they have different strengths and weaknesses. Often the shadow of a light source that shades a large visible area of an object is occluded and vice versa. The Hough transform can introduce spurious lights in the shading-based method and the extended area source approximation of directional sources in the shadow-based method can introduce significant errors. We demonstrate how the two sources of information complement each other in a number of occasions. Even when both methods are applicable at the same time, combining them reduces the error and speeds up computation. Hence, we arrive at an approach that integrates the two methods, with results superior to those obtained if the two methods are used separately. The resulting illumination information can be used (i) to render synthetic objects in a real photograph with correct illumination effects, and (ii) to virtually re-light the scene.

The rest of this paper is structured as follows. Section 2 describes the notion of critical points and their properties as they pertain to our problem. Section 3 describes the basic shading-based algorithm and extensions that make it robust to noise and missing data. These properties of our algorithm allow its application to objects of arbitrary shape in Section 3.4. The shadow-based algorithm is in Section 4. The two methods are compared and integrated in Section 5. We apply our method to the synthesis of Augmented Reality images in Section 6 and conclude with future work in Section 7.

2. Critical points

Definition 1. Given an image, let \mathbf{L}_i , $i = 1, 2, \dots$, 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

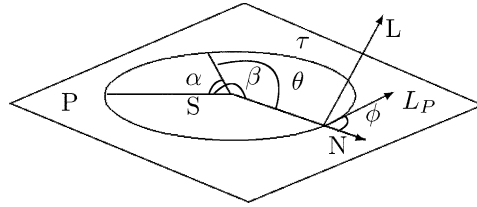


Fig. 1. L and its projection L_p onto plane P .

BRDF $f(\theta_i, \phi_i; \theta_e, \phi_e)$ is known to be a constant and each surface point appears equally bright from all viewing directions

$$E = \rho L_i \mathbf{L} \cdot \mathbf{n} = \rho L_i \cos \theta_i, \tag{1}$$

where E is the scene radiance of an ideal Lambertian surface, ρ is the albedo, L represents the direction, and L_i represents the amount of incident light, and \mathbf{n} is the unit normal to the surface.

Initially, the algorithm is developed using a sphere model and subsequently extended [26] to objects of arbitrary shape.

- We assume the observed object is a sphere with Lambertian reflectance properties whose physical size is already known.
- For light sources whose direction is co-linear with the lens axis of the camera, ¹ the best possible result is their equivalent frontal light source L_{frontal} .

It has been proven in [30] that it is not possible 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 kind of situation, i.e., an object illuminated by antipodal light sources, happens rarely, so for simplicity in the rest of this paper, we will make an additional assumption that there are no antipodal light sources.

2.1. Sphere cross-section with a plane P

Let P be an arbitrary plane such that S , the center of the sphere, lies on it (Fig. 1), L_i , $i = 1, 2, \dots$, be the light sources of the image and $(L_i)_p$ their projections on P . A point on the arc τ can be specified by its corresponding angle parameter in $[\alpha, \beta]$ using the following proposition [30]:

Proposition 1. Consider an angle interval $[\alpha, \beta]$ of a sphere cross-section (Fig. 1). We can always find a partition $\theta_0 = \alpha < \theta_1 < \dots < \theta_n = \beta$ 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 , $1 \leq i \leq n$ (Fig. 2), where $E(\theta)$ is the intensity function along the arc τ .

¹ We assume that they are co-linear when they form an angle less than a threshold ω depending on the resolution.

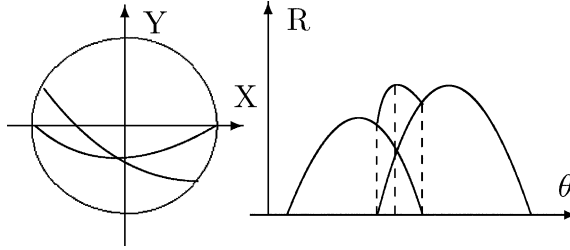


Fig. 2. The XY -space and the θR -space for the case with two directional light sources.

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. These two virtual light sources will be different, if each of these two parts is lit by a different illuminant configuration. More formally, Proposition 2 (from [30]) describes the difference between (b_{i-1}, c_{i-1}) and (b_i, c_i) .

Proposition 2. *In the configuration of Proposition 1, for any $1 \leq i \leq n$, we define A_i as the index set of real light sources contributed to the $[\theta_{i-1}, \theta_i]$ part of the arc τ . Then the Euclidean distance between two (b_i, c_i) pairs is*

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

where $\|(\mathbf{L}_j)_P\|$ is the Euclidean norm, $A' = A_{i-1} \setminus A_i$ (the index set of elements in A_{i-1} but not in A_i), $A'' = A_i \setminus A_{i-1}$, and $\sum_{j \in A_i} \mathbf{L}_j$ is the virtual light source corresponding to $[\theta_{i-1}, \theta_i]$.

Propositions 1 and 2 show that the difference between (b_{i-1}, c_{i-1}) (b_i, c_i) will be maximized at a critical point for these two virtual light sources. As we can see from Eq. (2), possible critical points can be detected by thresholding $\|(L_j)_P\|$.

2.2. Properties of critical points

Let Σ be the set of all critical points and Ω be the space of the sphere image. Then intuitively Σ will cut Ω into a decomposition, i.e.,

$$\Omega = \left(\bigcup_{i \in I} u_i \right) \cup \Sigma, \tag{3}$$

where each $u_i \subset \Omega$ is a subset of R^2 which does not contain any critical points and I is an index set.

Proposition 3. *Given a decomposition of the image as described by (3), for any image region u_i , which corresponds to a 3D surface region s_i , there exists a light source \mathbf{L} such that when s_i is illuminated by \mathbf{L} , the resulting image of s_i is exactly the same as u_i .*

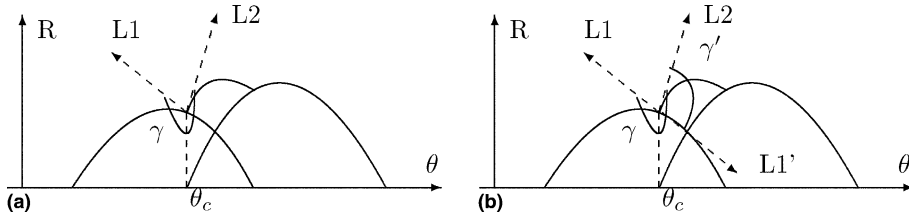


Fig. 3. (a) Inner angle γ . (b) Angles between two tangent lines.

Proposition 2 already provides us with a criterion to detect critical points on the sphere based on the distance between (b_i, c_i) pairs. Unfortunately, this criterion greatly depends on the intensities of virtual light sources, $\sum_{j \in A' \cup A''} \|(\mathbf{L}_j)_P\|$, which are projected on the plane with respect to each different cross-section. To locate the critical points more accurately, we provide another way to detect critical points on each cross-section. Instead of using the distance between (b_i, c_i) pairs, we can use the tangent angles defined on the intensity curve (Fig. 3a) [26].

Proposition 4. *Along a sine curve, at a critical point θ_c , the inner angle γ between two tangent lines $(\mathbf{T}_1, \mathbf{T}_2)$ of each side of θ_c will be larger than 180° .*

3. Shading-based illuminant detection

3.1. Critical point detection

From Proposition 1 we know that, for every cross-section of the sphere with a plane \mathbf{P} such that \mathbf{S} , the center of the sphere, lies on \mathbf{P} (illustrated in Figs. 2 and 4), there is a partition $\theta_0 = a < \theta_1 < \dots < \theta_n = \beta$ of the angle 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 $c_i, 1 \leq i \leq n$. By applying a standard recursive least-squares algorithm [6], we can use the following two consecutive windows (AW and WB) to detect the local maximum points of inner angle γ and distance defined by Eq. (2). Starting from an initial point A, any point B on the same arc can be determined uniquely by the angle θ

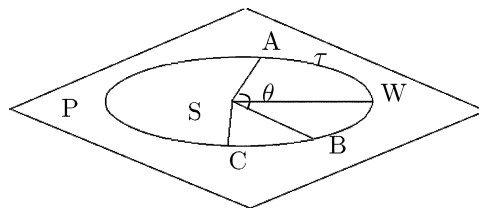


Fig. 4. A part of arc τ , AB, is covered by two consecutive windows AW and WB.

between SA and SB. Then we cover this part AB by the two consecutive windows AW and WB (Fig. 4).

With points B and W moving from the beginning point A of the visible part of the circle to its ending point C along the arc τ , we could estimate b_i and c_i from the data in the two consecutive windows AW and WB, respectively. Once a local maximum point of Eq. (2) is detected, it signifies that we have included at least one ‘critical point’ in the second window WB. Because the inner angle γ , defined in Proposition 4, is very sensitive to noise, we use two different criteria simultaneously to detect critical points. First we examine the distance defined in Proposition 2, then if the distance is above a threshold T_{distance} , we try to locate the critical point by searching for the maximum inner angle γ along the curve. In practice, for the distance criterion threshold T_{distance} , we use a ratio T_{ratio} instead of the direct Euclidean norm to normalize for the varying light intensities. T_{ratio} is calculated from Proposition 2:

$$T_{\text{ratio}} = \frac{\sqrt{(b_i - b_{i-1})^2 + (c_i - c_{i-1})^2}}{\max\{\sqrt{b_{i-1}^2 + c_{i-1}^2}, \sqrt{b_i^2 + c_i^2}\}}, \quad (4)$$

where (b_{i-1}, c_{i-1}) and (b_i, c_i) are calculated from the two consecutive windows AW and WB. Therefore, we can keep growing the first window AW to find critical point p_c using the recursive least-squares algorithm again. Then we fix the initial point A at p_c and keep searching for the next ‘critical point’ until we exhaust the whole arc τ .

3.2. Segmenting the surface

Definition 2. All critical points corresponding to one real light can 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 in a (ζ, θ) angle-pair parameter Hough space, i.e., we apply the cross-section plane equation in the following form:

$$\begin{cases} x \cdot n_x + y \cdot n_y + z \cdot n_z = 0, \\ n_x = r \cos \theta, \quad n_y = r \sin \theta \cos \zeta, \quad n_z = r \sin \theta \sin \zeta, \end{cases} \quad (5)$$

where (x, y, z) is the position of each critical point, (n_x, n_y, n_z) is the normal of the cross-section plane, r is the radius of the sphere, and $\zeta, \theta \in [0, 180]$. Typically, we use one-third of the highest vote count in the Hough transform as the threshold above which we detect a (ζ, θ) angle pair as a possible critical boundary.

Although critical points provide information to determine the light source directions [30], they are relatively sensitive to noisy data. Since most real images are not noise free, if we only use the Hough transform to extract critical boundaries, we will very likely find more boundaries than the real critical boundaries. Noise can either introduce many spurious critical points or move the detected critical points away from their true positions. 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.

Let **A**, **B**, **C**, and **D** be four points in a patch corresponding to one virtual light source and \mathbf{n}_A , \mathbf{n}_B , \mathbf{n}_C , and \mathbf{n}_D be their normals, respectively. From the Lambertian equation (1), augmented by an ambient light term α , we have

$$\begin{bmatrix} n_{Ax} & n_{Ay} & n_{Az} & 1 \\ n_{Bx} & n_{By} & n_{Bz} & 1 \\ n_{Cx} & n_{Cy} & n_{Cz} & 1 \\ n_{Dx} & n_{Dy} & n_{Dz} & 1 \end{bmatrix} \cdot \begin{bmatrix} L_x \\ L_y \\ L_z \\ \alpha \end{bmatrix} = \begin{bmatrix} I_A \\ I_B \\ I_C \\ I_D \end{bmatrix}, \tag{6}$$

where I_A , I_B , I_C , and I_D are brightness of each pixel in the source image corresponding to four points **A**, **B**, **C**, and **D**, respectively.

If \mathbf{n}_A , \mathbf{n}_B , \mathbf{n}_C , and \mathbf{n}_D are non-coplanar, we can obtain the direction of the corresponding virtual light source \mathbf{L} , $[L_x, L_y, L_z]^T$, and the ambient light intensity α by solving the system of equations in (6). Ideally, we would solve for the directions of virtual light sources by using four non-coplanar points from corresponding patches. Due to computation and rounding errors, four non-coplanar points are not always enough for us to get a numerically robust estimate of the direction of a virtual light source. Furthermore, it is not possible that we can always find several non-coplanar points in an interval of an arc in some plane **P** as described above. These problems are avoided by scanning the image both horizontally and vertically instead of one direction only and recovering the two-dimensional patches that are separated by critical boundaries. Then from each two-dimensional patch, we use the internal non-critical points of each virtual light patch to solve for the direction of the virtual light source.²

3.3. Recovering the true lights

Proposition 5. *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 $\mathbf{L}_2 - \mathbf{L}_1$ (Fig. 5).*

² We only use points that are at least 2 pixels away from the critical boundary for increased robustness to noise.

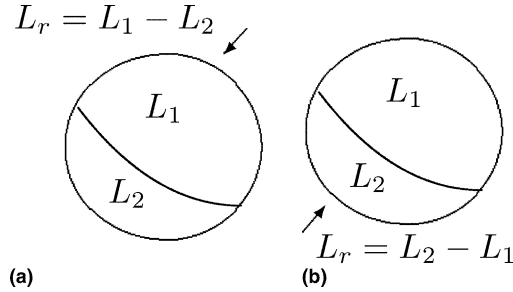


Fig. 5. Illustration of real light pre-direction. \mathbf{L}_r is the real light direction.

To find out the true directions, we pick a number of points on the surface, e.g., P_1, P_2, \dots, P_k and their normals, e.g., $\mathbf{N}_1, \mathbf{N}_2, \dots, \mathbf{N}_k$, then the true directions will be the solution of

$$E(P_j) = \sum_{i \in A} \max(e_i \mathbf{L}_i \cdot \mathbf{N}_j, 0) + \mathbf{L}_v \cdot \mathbf{N}_j, \quad 1 \leq j \leq k, \quad (7)$$

where \mathbf{L}_v is the virtual light source of a possible frontal illuminant whose critical boundaries could 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 [6,18].

After we find all the potential critical boundaries, Proposition 5 provides a way to extract real lights by calculating the light difference vector of two virtual light patches of two sides along the critical boundary. However, one real light might be calculated many times by different virtual light patch pairs, and since our data will not be perfect, they will not be necessary exactly the same vector. We introduce an angle threshold to cluster the resulting light difference vectors into real light groups, that can be 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, by the following steps (for an example see Fig. 6):

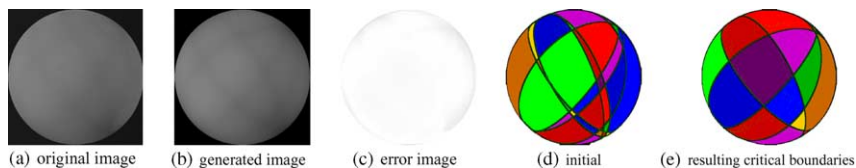


Fig. 6. Real sphere image: an almost Lambertian rubber ball with five light sources. Image size: 456×456 . (a) The original image. (b) The generated image of a Lambertian ball with the five light sources extracted from (a). (c) The error image: darker color means higher error. (d) The initial eight boundaries and virtual light patches extracted by the Hough transform. (e) The resulting critical boundaries and virtual light patches calculated by our algorithm, three out of the initial eight boundaries were automatically merged and the locations of the other five boundaries were automatically adjusted.

- (1) Detect critical points using the method described in Section 3.1.
- (2) Find initial critical boundaries by the Hough transform based on all detected critical points.
- (3) Adjust critical boundaries. We adjust every critical boundary by moving it by a small step, with a reduction in the least-squares error indicating a better solution. We keep updating boundaries using a “greedy” algorithm in order to minimize the total error.
- (4) Merge spurious critical boundaries. If two critical boundaries are closer than a threshold angle $T_{\text{mergeangle}}$ (e.g., 5°), they can be replaced by their average, resulting into one critical boundary instead of two.
- (5) Remove spurious critical boundaries. We 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. We test boundaries in increasing order of Hough transform votes (intuitively we test first the boundaries that are not as trustworthy).
- (6) Calculate the real lights along a boundary by subtracting neighboring virtual lights as described in Proposition 5.

3.4. 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. 7. Thus, many critical points’ locations will be calculated erroneously even for noise-free data. Consequently, the critical boundaries calculated by the Hough transform based on these critical points might not be correct or might be far away from their correct positions in some cases. Since we cannot recover these missing data from the original image, it is impossible to

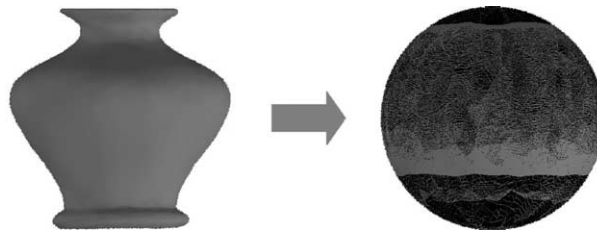


Fig. 7. Vase and its sphere mapping. Both image sizes are 400×400 . Black points on the sphere represent normals that do not exist on the vase’s surface.

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 adjustment steps described in Section 3.3). Thus it is still possible, using sparse points on the sphere, to calculate the true light for each virtual light patch based on Proposition 5. If two points have the same normal but different intensities, we use the brighter one (assuming that the other is in shadow).

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. 8. The total irradiance E at the shadow surface received from the entire illumination distribution is computed by

$$E = \sum_{i=1}^n L_i S_i \cos \theta_i, \tag{8}$$

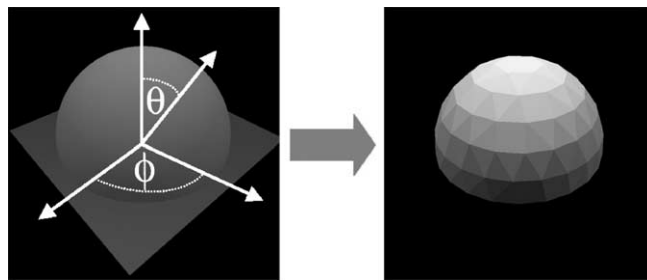


Fig. 8. The illumination distribution of a scene is approximated by discrete sampling over the entire surface of the extended light source.

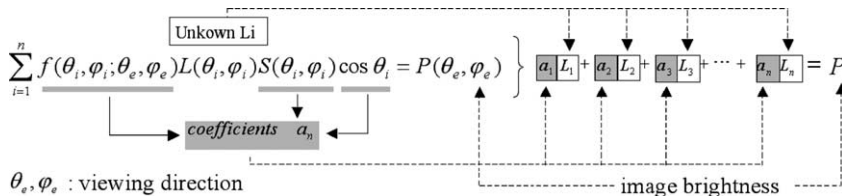


Fig. 9. Each shadow pixel provides a linear equation for estimating illumination distribution by shadows.

Fig. 11. Synthetic arbitrary shape image experiment: a top view of a vase illuminated by two light sources. Image size: 800×800 (scene), 400×400 (mapping sphere). (a) The original synthetic image of a top lit synthetic vase. The generated image of the vase with the illumination distribution estimated by (b) the shading-based method, (c) the shadow-based method and (d) the integrated method, (e) The critical boundary extracted using shading information only, where blue color represents points which are not mapped by the normals on the sphere's surface. One critical boundary (in red color) was missing when using shading information only. (f) The resulting estimation using shadow information. (g) The resulting estimation using the integrated method. Yellow areas are illuminant estimates using shadow information. Red points represent the true light directions, green points the estimates of the integrated method. The error image for (h) the shading-based method, (i) the shadow-based method and (j) the integrated method: darker color means higher error.

Fig. 12. Synthetic arbitrary shape image experiment: a vase illuminated by two light sources with partial shadows. Image Size: 800×800 (scene), 400×400 (mapping sphere). (a) The original synthetic image of a synthetic vase with partial shadows. The generated image of the vase with the illumination distribution estimated by (b) the shading-based method, (c) the shadow-based method and (d) the integrated method. (e) The critical boundary extracted using shading information only, where blue color represents points which are not mapped by the normals on the sphere's surface. (f) The resulting estimation using shadow information. Yellow areas are illuminant estimates using shadow information with 4° average angle error and high error of illuminant intensity. (g) The resulting estimation using the integrated method. Yellow areas are illuminant estimates using shadow information. Red points represent the true light directions, green points the estimates of the integrated method. The error image for (h) the shading-based method, (i) the shadow-based method and (j) the integrated method: darker color means higher error.

where L_i ($i = 1, 2, \dots, n$) 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 this approximation leads each image pixel inside shadows to provide a linear equation with unknown radiance of those sources as shown in Fig. 9 [24].

Finally, a set of linear equations (9) is derived from the brightness changes observed in the shadow image and solved for unknown L_i 's.

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{bmatrix} \cdot \begin{bmatrix} L_1 \\ L_2 \\ L_3 \\ \vdots \\ L_n \end{bmatrix} = \begin{bmatrix} P_1 \\ P_2 \\ P_3 \\ \vdots \\ P_m \end{bmatrix}. \tag{9}$$

Under the assumption of the Lambertian model, the BRDF $f(\theta_i, \phi_i; \theta_e, \phi_e)$ for a Lambertian surface is known to be a constant. Then, in Eq. (9) the coefficients a_i ($i = 1, 2, \dots, n$) represent $K_d \cos \theta_i S_i$ where K_d is a diffuse reflectance 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.

To estimate the illumination distribution of a real scene, we need to assume that the number of image pixels in shadows is far larger than the number of illumination radiance values to be calculated.

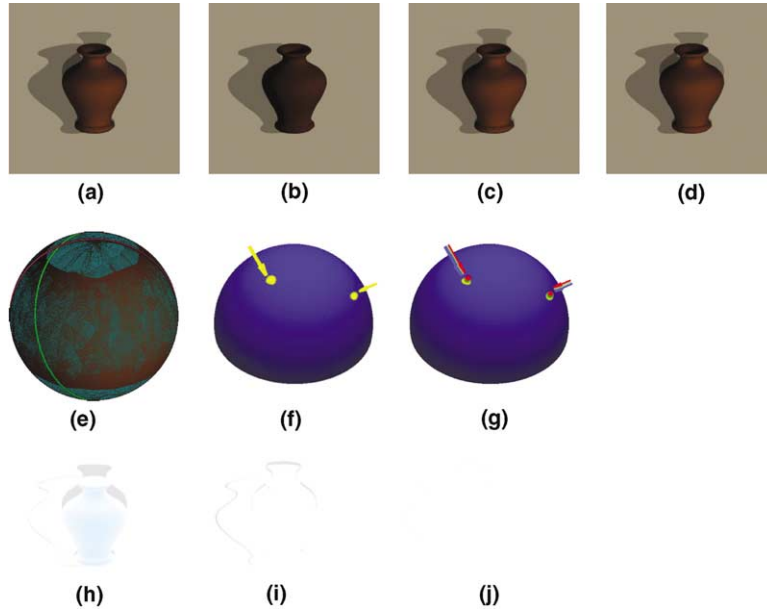


Fig. 11.

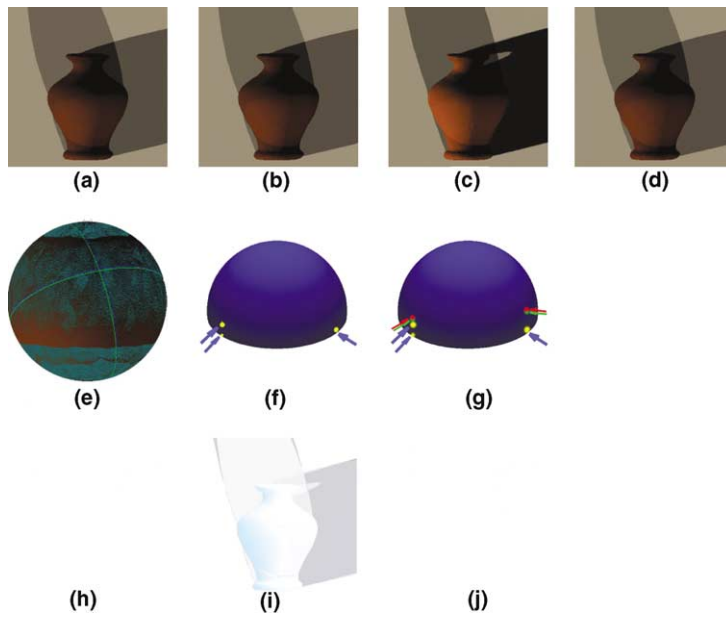


Fig. 12.

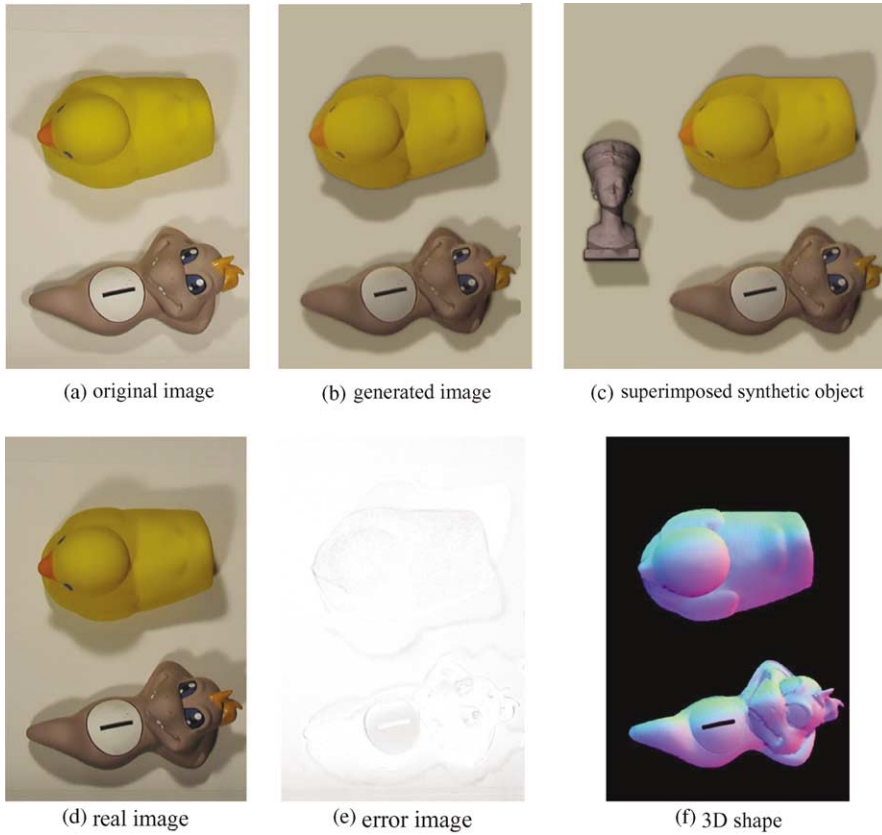


Fig. 15.

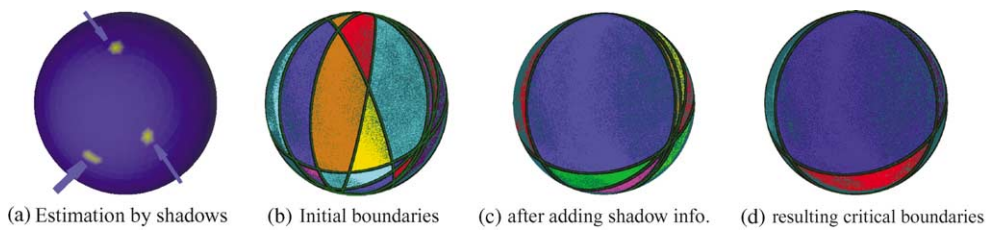


Fig. 16.

5. Integration of shading and shadows

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.



Fig. 15. Real arbitrary shape image experiment: a scene illuminated by three light sources. Image Size: 1534×1024 (scene), 400×400 (mapping sphere). (a) The original image. (b) The generated image of a scene with the two light sources extracted from (a). (c) A synthetic object is superimposed into the generated image (b). (d) The real image of the scene illuminated by the two real lights. (e) 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. (f) The 3D shape of the two objects' frontal surfaces, R,G,B color values represent the x, y, z components of the normal.

Fig. 16. The results of the various steps of the algorithm in Section 5.3 for the images of Fig. 15. (a) The distribution of illuminants estimated by shadow information. Notice that for each direction light source there are more than one non-zero intensity solid angles corresponding to it. (b) The initial eight boundaries extracted by the Hough transform. (c) The remaining five boundaries after adding shadow information. (d) The resulting critical boundaries calculated by our algorithm, two of the five boundaries in (c) were automatically removed and the locations of the other three boundaries were automatically adjusted.

5.1. Advantages of shadows over shading

As can be seen from Fig. 7, 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. Fig. 11 is an example of a synthetic vase image whose top part is lit by two directional light sources. From the sphere mapping (Fig. 11e), it is clear that very few of the object's normals map to the top part of the sphere and so not enough critical points can be detected. However, shadow information can be used to estimate the intensity and direction of each light source. Fig. 13a compares the the experiment results from shading-based, shadow-based, and integrated method.

5.2. Advantages of shading over shadows

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. Obviously, in this case shadows can not provide enough information to estimate the real illuminants. In particular the azimuth of the light source can still be estimated reliably, but not the elevation. An experiment showed that in this situation, big errors will be introduced to the illumination distribution estimated by shadow information only (Fig. 12d–e). Furthermore, in the method proposed in [23], a large number of samples are needed to capture the rapid change of radiance distribution around a direct light source. The radiance distribution inside a direct light source has to be sampled densely and the estimation becomes more stable if we observe the difference between the radiances of two shadow regions for each light source: one illuminated and the other not illuminated. Therefore, due to the discrete sampling of the geodesic dome, it is very likely that one directional light will be represented by several adjacent sampling solid angles so the precision of estimation will

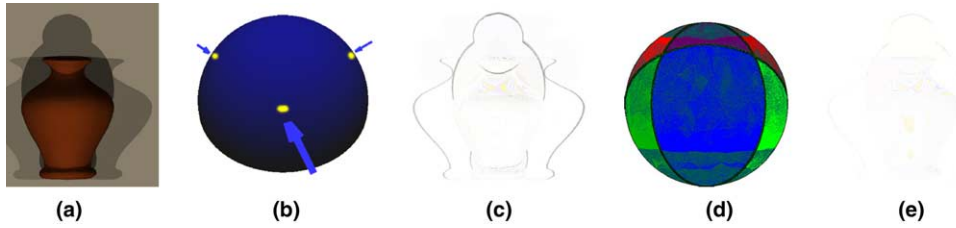


Fig. 10. (a) A synthetic vase illuminated by three directional light sources. (b) Estimated illumination distribution using shadow information only. (c) Error image generated by the illumination distribution estimated in (b). (d) Detected critical boundaries using shading information only. (e) Error image generated by the illumination distribution estimated in (d).

also be limited. In the following sections, a region on the geodesic dome in [23] composed by adjacent sampling solid angles, whose estimated illuminant intensity is not close to zero, will be referred to as an *illumination region*. In Fig. 10, we can see that the illumination distribution estimated by shading information provides higher accuracy than the one estimated by shadows. Fig. 13b compares the the experiment results from shading-based, shadow-based, and integrated method.

5.3. Shading and shadows

Definition 4. 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 corre-

Light	Average Angle Error (degree)			Average Image Intensity Error (gray scale level)			Average Image Intensity (gray scale level)
	Shading	Shadows	Integrated	Shading	Shadows	Integrated	original image
1	0.64	0.50	0.13	0.51	0.23	0.17	102.29
2	—	1.51	0.23	—	0.80	0.09	107.55

(a)

Light	Average Angle Error (degree)			Average Image Intensity Error (gray scale level)			Average Image Intensity (gray scale level)
	Shading	Shadows	Integrated	Shading	Shadows	Integrated	original image
1	0.03	3.96	0.03	0.03	4.08	0.03	98.96
2	0.03	3.42	0.03	0.28	19.76	0.26	93.62

(b)

Fig. 13. Error comparison of the shading-based, shadow-based and integrated methods. The experiments are based on (a) a top view of a vase illuminated by two light sources as described in Section 5.1 and (b) a vase illuminated by two light sources with partial shadows described in Section 5.2. In (a) the second light cannot be estimated by shading information only, due to the missing critical boundary.

sponding to a directional light source is generated by the occluding boundary of the object surface.

Fig. 14 shows that 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 shadow information to give us an initial estimate of the directions of the light sources, and then we can use the shading information to refine it to compute the directions and intensities of the real light sources.

In order to incorporate shadow information, we augment the algorithm of Section 3.3 by steps 3, 4, and 5:

- (1) Detect critical points using the method described in Section 3.1.
- (2) Find initial critical boundaries by the Hough transform based on all detected critical points.
- (3) Calculate an initial illumination distribution using the estimation from the shadow-based method [23]. Mark directions on the geodesic dome, for which possible shadows are not complete or observable due to occlusions, as ‘excluded’.
- (4) For each critical boundary, if its pre-direction is in a ‘non-excluded’ solid angle whose illuminant intensity is close to zero, consider it a spurious critical boundary and reject it. Otherwise mark the illumination region on the geodesic dome, containing this solid angle, as ‘registered.’
- (5) For each ‘non-registered’ illumination region, add a critical boundary whose pre-direction is close to the direction determined by the peak center of this region as an initial critical boundary.
- (6) Adjust critical boundaries. We adjust every critical boundary by moving it by a small step, with a reduction in the least-squares error indicating a better solution. We 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 $T_{\text{mergeangle}}$ (e.g., 5°), they can be replaced by their average, resulting into one critical boundary instead of two.
- (8) Remove spurious critical boundaries. We 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. We test boundaries in increasing order of Hough transform votes (intuitively we test first the boundaries that are not as trustworthy).

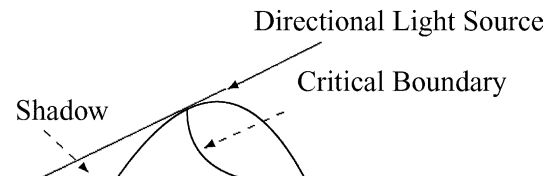


Fig. 14. Outline of estimating illumination distribution by shadows.

(9) Calculate the real lights along a boundary by subtracting neighboring virtual lights as described in Proposition 5.

Step 3 reduces significantly the spurious critical boundaries to be processed in steps 6 and 7, which are the most time consuming steps of the method. In our experiments this amounts to a 30% speed up (or more in the case of noisy data with a lot of spurious boundaries).

6. Augmented reality image synthesis

The combination of shading and shadow information can provide better estimation of illumination distribution. These estimates can be used to synthesize Augmented Reality images, i.e., real images with superimposed virtual objects rendered under correct illumination. 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. These abilities are demonstrated by the following real image experiments: 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. 15b, which can be compared with a real image of the scene with the same light truly switched off. Re-lighting is using the ratio images as described by Marschner et al. [15]. In the generated image we superimpose a synthetic object with correct shading and cast shadows in Fig. 15c. The original image and 3D geometry were captured by the range scanner system described in [9]. In Fig. 15f we can see that there are some inaccuracies and noise on the recovered 3D shape. The original image is 1534×1024 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 $T_{\text{ratio}} = 0.5$, and angle threshold for boundary merging (described in Section 3.3) $T_{\text{mergeangle}} = 5^\circ$.

Notice that the error of the generated image is located mostly along the edges of surfaces and shadows and this is because (1) the range scanned 3D shape in Fig. 15a has higher levels of estimation noise near the edges and the inter-reflections between the object and the table were not modeled, and (2) the simple rendering program used, does not simulate perfectly the shadowing effects of the real lights. The remaining noise in the generated image in Fig. 15b is due to inaccuracies in shape estimation and violations of the Lambertian assumption. Nonetheless illuminant estimation is still possible. In Fig. 16 we display the results of the various steps of the algorithm in Section 5.3.

7. Conclusions and future work

In this paper we presented a method for the estimation of multiple illuminant directions from a single image, incorporating shadow and shading information. We

demonstrate how information from each source enhances the information from the other source. 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 consists of directional light sources. Experiments on synthetic and real data show that the method is robust to noise, even when the surface is not completely Lambertian. We apply the results of our method to generate Augmented Reality images, by successfully modifying scene illumination and seamlessly re-rendering the scene, including superimposed synthetic objects. Future work includes study of the properties of arbitrary surfaces (so that we can avoid the intermediate sphere mapping), speeding up of the least-squares method and extending the method to non-Lambertian diffuse reflectance for rough surfaces.

Acknowledgments

This work was supported by DOE Grant MO-068. 3D shape of toys was captured by the 3D measurement system in the Optical Metrology Lab in SUNY Stony Brook, with the help of Song Zhang and Jiahui Pan.

References

- [1] R.T. Azuma, A survey of augmented reality, ACM SIGGraph, Computer Graphics, Course Notes for Developing Advanced Virtual Reality Applications 9 (1995) 1–38.
- [2] R. Basri, D. Jacobs, Lambertian reflectance and linear subspaces, in: *Int. Conf. on Computer Vision*, 2001, pp. 383–390.
- [3] W. Chojnacki, M.J. Brooks, D. Gibbins, Can the Sun's direction be estimated prior to the determination of shapes?, in: *Seventh Austral. Joint Conf. on Artificial Intelligence*, 1994, pp. 530–535.
- [4] P. Debevec, Rendering synthetic objects into real scenes, ACM SIGGraph, Computer Graphics (1998) 189–198.
- [5] P. Debevec, T. Hawkins, C. Lhou, H.P. Duiker, W. Sarokin, M. Sagar, Acquiring the reflectance field of a human face, ACM SIGGraph, Computer Graphics (2000) 145–156.
- [6] S. Haykin, *Adaptive Filter Theory*, Prentice-Hall, Englewood Cliffs, NJ, 1986.
- [7] B.K.P. Horn, M.J. Brooks, Shape and source from shading, in: *Int. Joint Conf. on Artificial Intelligence*, 1985, pp. 932–936.
- [8] D.R. Hougen, N. Ahuja, Estimation of the light source distribution and its use in integrated shape recovery from stereo and shading, in: *Int. Conf. on Computer Vision*, 1993, pp. 29–34.
- [9] Q.Y. Hu, 3-D Shape Measurement Based on Digital Fringe Projection and Phase-Shifting Techniques, M.E. Dept., SUNY at Stony Brook, 2001.
- [10] T. Kim, Y.D. Seo, K.S. Hong, Improving AR using shadows arising from natural illumination distribution in video sequences, in: *Int. Conf. on Computer vision*, vol. II, 2001, pp. 329–334.
- [11] M.J. Langer, S.W. Zucker, What is a light source?, *IEEE Comput. Vision Pattern Recognition* (1997) 172–178.
- [12] C.H. Lee, A. Rosenfeld, Improved methods of estimating shape from shading using the light source coordinate system, *Artif. Intell.* 26 (2) (1985) 125–143.
- [13] H.Y. Lin, M. Subbarao, A vision system for fast 3D model reconstruction, *IEEE Comput. Vision Pattern Recognition* (2001) 663–668.

- [14] C. Loscos, M.-C. Frasson, G. Drettakis, B. Walter, X. Granier, P. Poulin, Interactive virtual relighting and remodeling of real scenes, *ACM Eurographics Workshop on Rendering* (1999) 235–246.
- [15] S.R. Marschner, D.P. Greenberg, Inverse lighting for photography, in: *Fifth Color Imaging Conf.*, 1997, pp. 262–265.
- [16] S.R. Marschner, S.H. Westin, E.P.F. Lafortune, K.E. Torrance, Image-based BRDF measurement, *Appl. Opt.* 39 (16) (2000) 2592–2600.
- [17] M.W. Powell, S. Sarkar, D. Goldgof, A simple strategy for calibrating the geometry of light sources, *IEEE Trans. Pattern Anal. Machine Intell.* 23 (9) (2001) 1022–1027.
- [18] W.H. Press, S.A. Teukolsky, W.T. Vetterling, B.P. Flannery, *Numerical Recipes in C*, Cambridge University Press, Cambridge, 1992.
- [19] R. Ramamoorthi, P. Hanrahan, A signal-processing framework for inverse rendering, *ACM SIGGraph, Computer Graphics* (2001) 117–128.
- [20] D. Samaras, D. Metaxas, Coupled lighting direction and shape estimation from single images, in: *Int. Conf. on Computer Vision*, 1999, pp. 868–874.
- [21] D. Samaras, D. Metaxas, P. Fua, Y. Leclerc, Variable albedo surface reconstruction from stereo and shape from shading, *IEEE Comput. Vision Pattern Recognition* 1 (2000) 480–487.
- [22] I. Sato, Y. Sato, K. Ikeuchi, Acquiring a radiance distribution to superimpose virtual objects onto a real scene, *IEEE Trans. Visualization Comput. Graphics* 5 (1) (1999) 1–12.
- [23] I. Sato, Y. Sato, K. Ikeuchi, Illumination distribution from brightness in shadows, in: *Int. Conf. on Computer Vision*, 1999, pp. 875–883.
- [24] I. Sato, Y. Sato, K. Ikeuchi, Stability issues in recovering illumination distribution from brightness in shadows, *IEEE Comput. Vision Pattern Recognition* (2001) 400–407.
- [25] Y. Sato, M.D. Wheeler, K. Ikeuchi, Object shape and reflectance modeling from observation, *ACM SIGGraph, Computer Graphics* (1997) 379–388.
- [26] Y. Wang, D. Samaras, Estimation of multiple illuminants from a single image of arbitrary known geometry, in: *Eur. Conf. on Computer Vision*, vol. III, 2002, pp. 272–288.
- [27] Y. Yang, A. Yuille, Sources from shading, *IEEE Comput. Vision Pattern Recognition* (1991) 534–539.
- [28] Y. Yu, J. Malik, Recovering photometric properties of architectural scenes from photographs, *ACM SIGGraph, Computer Graphics* (1998) 207–217.
- [29] Y. Yu, P. Debevec, J. Malik, T. Hawkins, Inverse global illumination: recovering reflectance models of real scenes from photographs, *ACM SIGGraph, Computer Graphics* (1999) 215–224.
- [30] Y. Zhang, Y.H. Yang, Illuminant direction determination for multiple light sources, *IEEE Comput. Vision Pattern Recognition* vol. I (2000) 269–276.
- [31] Q. Zheng, R. Chellappa, Estimation of illuminant direction, Albedo, and shape from shading, *IEEE Trans. Pattern Anal. Machine Intell.* 13 (7) (1991) 680–702.