

Shedding Light on Reconstruction and Image-Based Rendering *

Peter N. Belhumeur
Center for Comp. Vision & Control
EE and CS
Yale University
New Haven, CT 06520-8267

David J. Kriegman
Beckman Institute
Computer Science
University of Illinois, Urbana-Champaign
Urbana, IL 61801

Abstract

We propose new methods for shape reconstruction and image-based rendering based on double covering a scene's light field. In particular, a method is presented for reconstructing an object's shape from a collection of images in which the viewpoint remains fixed and the object is illuminated by point light sources at different positions. The method can be applied to an object of any shape; the bidirectional reflectance density (BRDF) function may be arbitrary and changing over the surface; and shadows may be present. Unlike photometric stereo, nothing about the nature of the BRDF is assumed, and depth rather than the surface normal is estimated. The method for reconstruction leads directly to a novel method for rendering physically valid synthetic images of the object under arbitrary and previously unseen lighting conditions. Results for both surface reconstruction and image-based rendering are presented using images of a synthetic scene.

*D. J. Kriegman were supported by NSF under NYI IRI-9257990 and ARO DAAG55-98-1-0168. P. N. Belhumeur was supported by a Presidential Early Career Award, an NSF Career Award IRI-9703134, and ARO grant DAAH04-95-1-0494.

1 Introduction

We address an open problem in computer vision: how to reconstruct the shape of an object with an arbitrary, unknown bidirectional reflectance density function (BRDF). We also address an open problem in computer graphics: how to accurately render synthetic images of a real object – one for which the shape and BRDF are unknown – under arbitrary lighting conditions. The solutions to both problems are deeply interrelated and appear to require many images of the object illuminated by light sources at different positions. Yet, no assumptions are needed about the nature of the object’s shape, its surface reflectance, or the presence or absence of shadowing.

Our solution for reconstructing the object’s shape stands in contrast to existing methods which assume, either implicitly or explicitly, that the BRDF of points on the object’s surface are Lambertian, approximately Lambertian, or of some known parametric form. These methods include but are not limited to binocular stereopsis, structure from motion, shape from defocus, and photometric stereo. In contrast, we present a computationally simple, but data intensive method for estimating the shape of an object when the surface’s BRDF is unknown. The method correctly handles both attached and cast shadows.

The method requires only a single viewpoint of the object, but *many* images of the object illuminated by point light sources at different positions. In particular, we require two sets of images of the object: an inner and outer set. The inner set of images is created by moving a point light source over any known surface that is star-shaped (e.g., convex) with respect to all surface points. The outer set of images is similarly acquired, with the additional requirement that its star-shaped surface does not intersect the inner surface, see Fig. 1. Note that these surface need not be closed.

Using these two sets of images, a point for point reconstruction of the object’s visible surface is performed by estimating the depth of each point along the line of sight. The estimation of depth exploits a simple assumption: the radiance along a ray of light is constant. With this assumption in hand, we reconstruct the surface by double covering the light field (the space of light rays) emanating from the inner star-shaped surface of light source positions. In particular, we are able to equate the scene radiance of a point on the object’s surface produced by a point light source lying on the inner star-shaped surface with the scene radiance of the same point produced by some corresponding point light source lying on the outer star-shaped surface. The correct correspondence can then be set up as a one-dimensional optimization problem over the point’s depth along the light of sight.

In addition to recovering the object’s shape, we can use the recovered shape and samplings of the object’s light field to efficiently and accurately render images of real objects under novel illumination conditions. In [1, 4, 6, 7, 10, 13, 16], images of a real object taken from multiple viewpoints are used to render synthetic images of the object from arbitrary viewpoints. In [1, 4, 6, 10, 16], few real images are needed for the synthetic renderings, but the method first must determine a 3D model of the scene by establishing the correspondence of features pixels in the real images. A radical departure from reconstruction or correspondence-based approaches to image-based rendering was the 4-D lumigraph [7] or light field [13]. In these methods, renderings of the object from novel viewpoints can be synthesized without any 3D model of the scene; however, thousands of real images are needed for accurate renderings. See the discussion in Section 3.

In contrast to this work, we present a method for rendering synthetic images of an object or scene from a fixed viewpoint, but under arbitrary illumination conditions. The method uses many images of an object illuminated by point light sources, to recover the object’s shape and then render synthetic images of the object under arbitrary lighting conditions. The problem of synthesizing images for Lambertian surfaces with light sources at infinity without shadows is considered in [17] and with shadows in [3]. Methods for re-rendering images with diffuse linear combinations of images formed under diffuse light is considered in [14]. It is proposed in [20] to perform image-based rendering under variable illumination by estimating an apparent BRDF (for a fixed viewing direction) associated with each scene point by systematically moving light sources at infinity. However, to synthesize images for nearby light sources, 3-D scene geometry as well as the apparent BRDF at each point is needed, and it is assumed that geometry has been acquired by some other means (e.g., a range finder).

2 Reconstruction

Consider the general problem of reconstructing from one or more images the 3-D geometry of a surface whose bidirectional reflectance density function (BRDF) is arbitrary (i.e., not well approximated by a parametric model such as Phong) and varying over the surface. A general BRDF at a point on the surface can be represented as a 4-D function $\rho(\hat{\mathbf{d}}, \hat{\mathbf{r}})$ where $\hat{\mathbf{d}} \in S^2$ (unit sphere) is the direction of an incoming light ray, and $\hat{\mathbf{r}} \in S^2$ is the direction of the outgoing ray.

For a surface whose BRDF is not a function $\hat{\mathbf{r}}$ (e.g., Lambertian), the image intensity of a surface point will be the same irrespective of the observer’s viewpoint. This “constant brightness assumption” is the basis for establishing correspondence in all dense stereo and motion methods. Yet for objects with a general BRDF, this constant brightness assumption is violated, and so establishing correspondences between images gathered from different viewpoints is difficult if not impossible; hence methods like stereo, motion, or wide baseline structure from motion are severely hampered by the invalidity of this assumption.

Similarly, nearly all photometric stereo methods assume that the BRDF is Lambertian [12, 15, 18, 21] or can be specified using a small number of parameters usually derived from limited physical models [9, 8, 19]. In these methods, multiple images under varying lighting directions (but from fixed viewpoint) are used to estimate the surface normal for a surface patch projecting onto a single pixel. This procedure is done for each pixel in the image, and the resulting normal field is integrated to produce a surface.

In this section, we present a method for surface reconstruction that resembles photometric stereo in that a single viewpoint and multiple lighting directions are used. Yet, our method differs significantly in that depth is directly estimated, and no assumptions are made about the surface BRDF.

Let us first consider a fixed calibrated pinhole camera observing a static scene; see Figure 1. Let the coordinates of a point on the image plane be given by $\mathbf{q} \in \mathbb{R}^2$. For every \mathbf{q} , there is a line passing through the optical center \mathbf{o} in the direction $\hat{\mathbf{r}}(\mathbf{q})$ which we call the line of sight of pixel \mathbf{q} . We obtain the function $\hat{\mathbf{r}}(\mathbf{q})$ during camera calibration, and \mathbf{o} will be taken as the origin. The image point \mathbf{q} is the projection of a scene point \mathbf{p} lying on the line defined by \mathbf{o} and $\hat{\mathbf{r}}(\mathbf{q})$. The depth $\lambda(\mathbf{q})$ of \mathbf{p} from \mathbf{o} is unknown, and the relation can be expressed as

$$\mathbf{p}(\mathbf{q}, \lambda) = \lambda(\mathbf{q})\hat{\mathbf{r}}(\mathbf{q}) + \mathbf{o}. \tag{1}$$

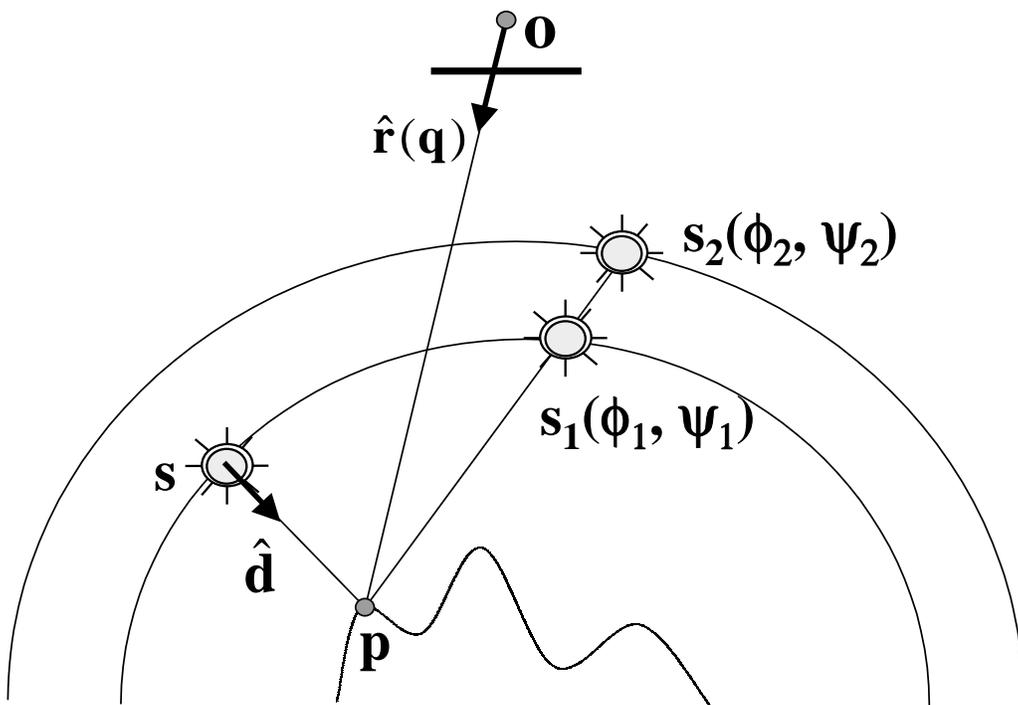


Figure 1: A two-dimensional schematic of the reconstruction setup. A camera whose origin is at \mathbf{o} observes a scene point \mathbf{p} which is illuminated by light sources covering two surfaces, parameterized as $\mathbf{s}_1(\phi_1, \psi_1)$ and $\mathbf{s}_2(\phi_2, \psi_2)$.

The process of reconstruction is to estimate the depth map $\lambda(\mathbf{q})$, in this case from images gathered under different lighting conditions. Since we will be able to independently estimate λ for each \mathbf{q} , we will drop \mathbf{q} from our notation and write \mathbf{p} as a function of the unknown depth.

Consider the scene to be illuminated by an isotropic point light source (not at infinity) whose location $\mathbf{s} \in \mathbb{R}^3$ is known. The direction of the light ray from \mathbf{s} to \mathbf{p} is $\hat{\mathbf{d}}(\mathbf{s}, \lambda) = \frac{1}{\|\mathbf{p}(\lambda) - \mathbf{s}\|}[\mathbf{p}(\lambda) - \mathbf{s}]$, while the distance between \mathbf{s} and \mathbf{p} is $d(\mathbf{s}, \lambda) = \|\mathbf{p}(\lambda) - \mathbf{s}\|$. While the BRDF is typically defined with respect to a coordinate system attached to the surface and with one of the axis in the direction of the surface normal $\hat{\mathbf{n}}$, we will specify it in a global coordinate system as a function of the incoming light ray $\hat{\mathbf{d}}$ and the outgoing direction $-\hat{\mathbf{r}}$; i.e., we write the apparent BRDF as $\rho(\hat{\mathbf{d}}, \hat{\mathbf{r}})$. (Note also that this apparent BRDF will include global properties of the scene like cast shadows.) While the relation between the incoming irradiance to the outgoing radiance is proportional to the true BRDF and the cosine between the incoming light and surface normal, we “fold” the cosine term into the apparent BRDF $\rho(\hat{\mathbf{d}}, \hat{\mathbf{r}})$.

The image intensity measured at \mathbf{q} is a function of the light source intensity, $d^2(\mathbf{s}, \lambda)$ and $\rho(\hat{\mathbf{d}}, \hat{\mathbf{r}})$. Without loss of generality, let all images be acquired with the same light source which can be taken to have unit intensity. The measured image intensity (irradiance) for image point \mathbf{q} corresponding to a surface point at depth λ illuminated by light source \mathbf{s} can be expressed as:

$$i(\mathbf{s}) = \frac{1}{d^2(\mathbf{s}, \lambda)} \rho(\hat{\mathbf{d}}(\mathbf{s}, \lambda), \hat{\mathbf{r}}). \quad (2)$$

As shown in Figure 1, let the light source be moved over any known surface which is star-shaped with respect to all points on the object – any convex surface is sufficient. Parameterizing the surface by (ϕ_1, ψ_1) , it can be expressed as $\mathbf{s}_1(\phi_1, \psi_1)$. For every light source position $\mathbf{s}_1(\phi_1, \psi_1)$, an image $i_1(\phi_1, \psi_1)$ is measured. If the depth λ were known, then from the image data, a two-dimensional slice for fixed $\hat{\mathbf{r}}$ of the apparent BRDF at \mathbf{p} could be determined as $\rho(\hat{\mathbf{d}}(\mathbf{s}, \lambda), \hat{\mathbf{r}}) = d^2(\mathbf{s}_1(\phi_1, \psi_1), \lambda) i_1(\phi_1, \psi_1)$. During modeling, the viewing direction $\hat{\mathbf{r}}$ is fixed, but all light source directions are covered. Alternatively, if the BRDF were known, then one can solve for the depth λ given the measured intensities $i_1(\phi_1, \psi_1)$ and the known light source positions $\mathbf{s}_1(\phi_1, \psi_1)$.

We now consider a method for simultaneously estimating the BRDF and the depth. Let a second light source be moved on second star-shaped surface which does not intersect the first one. Express this surface parametrically as $\mathbf{s}_2(\phi_2, \psi_2)$ and the corresponding measured images intensities as $i_2(\phi_2, \psi_2)$.

For every light source position from the first (inner) surface $\mathbf{s}_1(\phi_1, \psi_1)$, there is a light source on the second (outer) surface $\mathbf{s}_2(\phi_2, \psi_2)$ where the ray from \mathbf{p} through $\mathbf{s}_1(\phi_1, \psi_1)$ is identical to the ray from \mathbf{p} through $\mathbf{s}_2(\phi_2, \psi_2)$; see Figure 1. We can express this correspondence of light sources on the two surfaces as a change of coordinates $\phi_2(\phi_1, \psi_1; \lambda)$ and $\psi_2(\phi_1, \psi_1; \lambda)$. This change of coordinates depends on the unknown location of \mathbf{p} , and so it is parameterized by the depth λ .

For such a pair of light sources $\mathbf{s}_1(\phi_1, \psi_1)$ and $\mathbf{s}_2(\phi_2, \psi_2)$, the value of the apparent BRDF $\rho(\hat{\mathbf{d}}, \hat{\mathbf{r}})$ is the same, and so the image intensities are related by

$$\begin{aligned} i_2(\phi_2, \psi_2) &= \frac{[d_1(\mathbf{s}_1(\phi_1, \psi_1), \lambda)]^2}{[d_2(\mathbf{s}_2(\phi_2, \psi_2), \lambda)]^2} i_1(\phi_1, \psi_1) \\ &= i_2(\phi_2(\phi_1, \psi_1; \lambda), \psi_2(\phi_1, \psi_1; \lambda)) \end{aligned} \quad (3)$$

This relation between the intensities for corresponding light sources can be used to form an objective that is a function of the depth λ :

$$\begin{aligned} \mathcal{O}(\lambda) = \int \int & [d_2^2(\lambda) i_2(\phi_2(\phi_1, \psi_1; \lambda), \psi_2(\phi_1, \psi_1; \lambda)) \\ & - d_1^2(\lambda) i_1(\phi_1, \psi_1)]^2 d\phi_1 d\psi_1. \end{aligned} \quad (4)$$

The depth λ is then found by minimizing $\mathcal{O}(\lambda)$, which would be zero without any image noise. At such a minimum, we have found a depth where the correspondence between light sources on the two surfaces leads to image intensities which are consistent.

2.1 Implementation

We have implemented this method using synthetic images created by a ray tracing package (VORT). In the reported experiments, a teapot rests on a Lambertian support plane. Phong shading is used to model the reflectance function of the teapot, with parameters chosen for a Lambertian teapot, a plastic teapot, and a metal teapot. Naturally, the reconstruction method has no knowledge of the underlying structure of the BRDF. Figure 2 shows a collection of sample images of a plastic teapot.

The development of (4) is based on measuring the image intensity as a continuous function of the light source location. In practice, we obtain images by sampling the two surfaces of point light sources; in our examples the two surfaces were taken to be spheres about the object. In general for some sampled light source $\mathbf{s}_1(\phi_1, \psi_1)$, there is no corresponding sampled light source

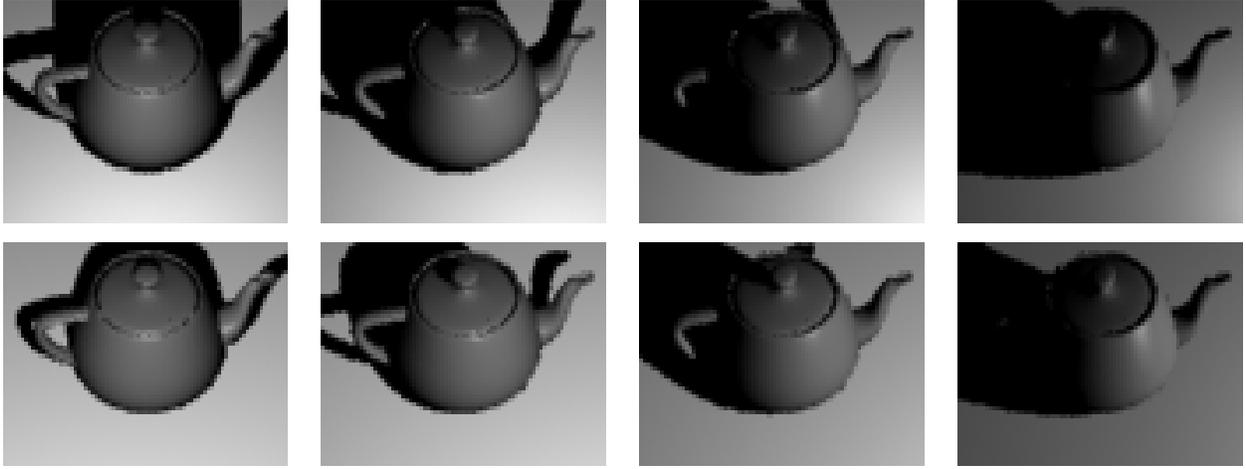


Figure 2: The top row shows four input images from the inner sphere of light sources. The bottom row shows four input images from the outer sphere of light sources. Notice the attached and cast shadows as well as the specularities.

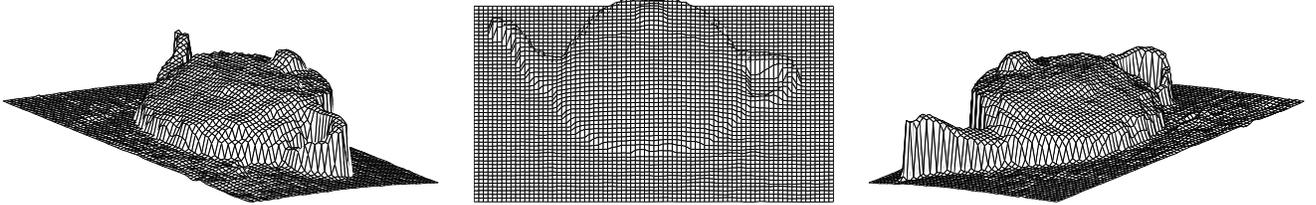


Figure 3: Three views of the reconstructed plastic teapot. (Caveat: when plotted using matlab, the scales in the x, y and z directions are arbitrary resulting in a global affine transformation of the shape in these figures.)

at $\mathbf{s}_2(\phi_2(\phi_1, \psi_1; \lambda), \psi_2(\phi_1, \psi_1; \lambda))$. So, instead, we must interpolate from the available samples, and this is done as follows. A Delaunay triangulation of the sampled light sources on the second surface is performed. Given $\mathbf{s}_1(\phi_1, \psi_1)$ and an estimated depth λ , we find the intersection $\mathbf{s}'_2(\phi_1, \psi_1)$ of the ray defined by $\mathbf{p}(\mathbf{q})$ and $\mathbf{s}_1(\phi_1, \psi_1)$ with one of the triangles in the triangulation of the second surface. From the intensity values corresponding to the vertices and the coordinates of the light source of the vertices, bilinear interpolation is used to approximate the intensity $i'_2(\mathbf{s}'_2)$.

The integral in (4) becomes a summation over n sampled light sources whose locations are $\mathbf{s}(\phi_j, \psi_j)$ on the first surface and with corresponding pixel intensity $i_1(\phi_j, \psi_j)$. This leads to the discrete form of the objective function,

$$\mathcal{O}'(\lambda) = \sum_{j=1}^n [d_2^2(\lambda) i'_2(\phi_2(\phi_j, \psi_j; \lambda), \psi_2(\phi_j, \psi_j; \lambda)) - d_1^2(\lambda) i_1(\phi_j, \psi_j)]^2. \quad (5)$$

There is no reason to expect (5) to be convex, but fortunately it is only a function of one variable λ , and we have bounds on λ given by smallest diameter of the two surfaces of light

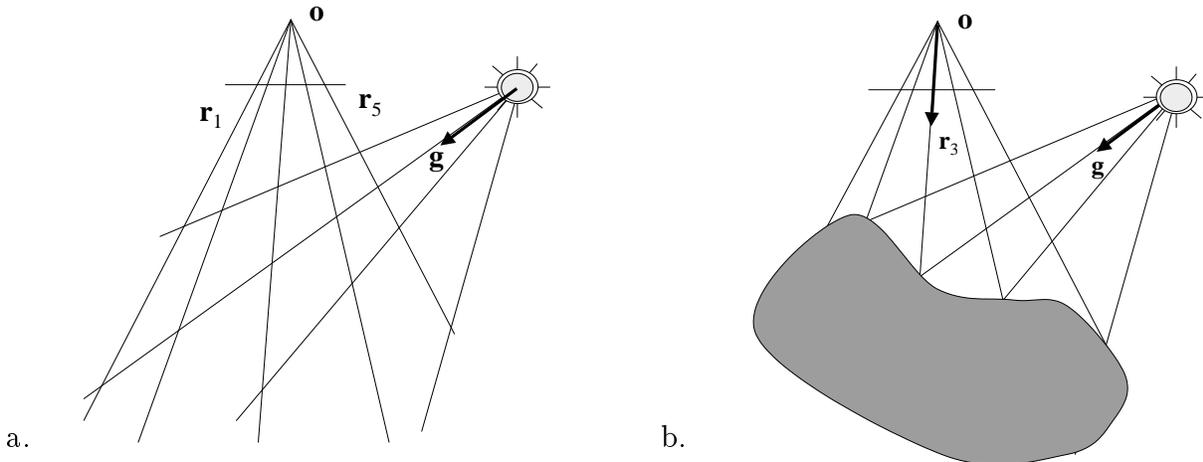


Figure 4: a. Without knowing the scene geometry, there is no way to determine which line of sight $\mathbf{r}_1 \cdots \mathbf{r}_5$ corresponds to an illuminating ray, say \mathbf{g} ; b. However, if the scene geometry is known, then one can establish the correspondence of \mathbf{g} with \mathbf{r}_3 .

sources. To estimate the minimum of $\mathcal{O}'(\lambda)$, we evaluate \mathcal{O}' for ten values of λ , select the minimal one, and then perform local optimization on $\mathcal{O}'(\lambda)$. Since $\mathcal{O}'(\lambda)$ is independent for each pixel, the depth of each pixel $\lambda(\mathbf{q})$ can be estimated independently.

In Fig. 2, we display eight of the 400 VORT generated images used to reconstruct the plastic teapot. For the inner surface of light sources, a total of 320 monochrome images with 100x100 pixels were used. For the outer surface of light sources, a total of 80 monochrome images with 100x100 pixels were used. Figure 3 shows results of the reconstruction as a wire-frame displayed from three different viewpoints. Similar tests were performed on a purely Lambertian teapot and a highly specular (metal) teapot, but the results were so similar that they did not warrant inclusion in this paper.

3 Image-based rendering

At SIGGRAPH 1996, two papers introduced a novel approach to image-based rendering of natural 3-D scenes from arbitrary viewpoints [7, 13]. Rather than modeling object geometry and the reflectance function across the surface as is traditionally done in computer graphics, the approach is based on directly representing the radiance in all directions emanating from a scene under constant illumination. As discussed in [11], the set of light rays is a four-dimensional manifold. Under static illumination, the radiance along a ray in free space is constant. Note that this reduces the 5-D plenoptic function to 4-D [5]. Now consider surrounding a scene by a closed smooth convex surface. By moving a camera with its two-dimensional image plane over the entire surface (a 2-D manifold), one can sample the intensity along every ray emanating from the surface. In doing this, one obtains a function on the 4-D ray space \mathcal{L} which has been called the Lumigraph [7] or light field [13].

For any viewpoint \mathbf{o} outside of the surface, an image can be synthesized by considering the radiance of all of the rays passing through \mathbf{o} . The set of rays passing through \mathbf{o} is simply a two-dimensional subset of \mathcal{L} , and radiance of those rays that intersect the image plane are used to compute the irradiance of the synthesized image. This turns rendering into a problem of

simply indexing into a representation of \mathcal{L} rather than a ray tracing or radiosity problem. The advantages of such an approach are that the representation is constructed directly from images without needing reconstruction or correspondence, that no assumptions about the surface BRDF are required, and that interreflections do not need to be computed since they have occurred physically when the images were acquired. The main challenges and issues of this approach are interpolation from the finite number of sample images, compression of the lumigraph/light field representation which is constructed from a very large number of images, and accurate camera localization during modeling.

In [7, 13] the illumination must be constant during modeling, and all synthesized images are valid only under the same illumination; this complicates rendering scenes composed of both traditional geometric models and lumigraphs/light fields under general lighting. It is natural to ask whether one could directly “turn the lumigraph/light field around” and synthesize images under fixed pose, but variable lighting. As described in [11], the space of source rays illuminating a scene is four-dimensional. Like the rays passing through a camera’s optical center, the set of light rays emanating from a point light source is two-dimensional. Hence, by moving an isotropic point source over a closed surface (a 2-D manifold) bounding a scene, images can be acquired for all possible source rays crossing this surface. As in the reconstruction method described in Section 2, let the surface of point source locations be given by $\mathbf{s}(\phi, \psi)$ and the images be given by $\mathbf{I}(\phi, \psi)$;

Now consider synthesizing an image from the same viewpoint, but under completely different lighting conditions. The applied lighting is a function on the 4-D light ray space. For a single illumination ray \mathbf{g} , we can find the intersection of the ray with the surface of point sources. The intersection is a point light source location $\mathbf{s}(\phi, \psi)$ and there is corresponding image $\mathbf{I}(\phi, \psi)$. There exists a light ray emanating from the point source $\mathbf{s}(\phi, \psi)$ coincident with \mathbf{g} which intersects the scene, and sheds light onto some image pixel. However, it is not evident which pixel of $\mathbf{I}(\phi, \psi)$ corresponds to the intersection of the illuminating ray and the surface. This problem is illustrated in Figure 4.a. If the depth were known, this dilemma of determining the correspondence between an image pixel and an illuminating ray can be resolved as illustrated in Figure 4.b.

These observations lead us to a method for rendering an image of a scene illuminated by a point light source \mathbf{s} which does not necessarily lie on the surface defined by $\mathbf{s}(\phi, \psi)$. The method can clearly be extended to render scenes under other light sources, e.g., area sources, strip sources, point sources at infinity, etc. First, the method of Section 2 is used to reconstruct the depth $\lambda(\mathbf{q})$ of the surface corresponding to each pixel. As in Section 2, the surface of light source locations, say $\mathbf{s}_1(\phi_1, \psi_1)$, is triangulated. To determine the intensity of a pixel \mathbf{q} for point light source \mathbf{s} , we find the intersection \mathbf{s}' of the ray defined by $\mathbf{p}(\mathbf{q})$ and \mathbf{s} with one of the triangles in the above mentioned triangulation. See Figure 5. If the intersection happens to be a vertex, then the intensity of the pixel in the corresponding measurement image could be used. Since this is rarely the case, we instead interpolate the intensities of corresponding pixels associated with the three vertices to estimate the intensity $i'(\mathbf{q})$ for a fictitious point light source at \mathbf{s}' . Because the solid angle of the surface corresponding to \mathbf{q} as seen by the \mathbf{s} and \mathbf{s}' depends on the squared distance, the pixel is rendered with

$$i(\mathbf{q}) = \frac{\|\mathbf{p} - \mathbf{s}'\|^2}{\|\mathbf{p} - \mathbf{s}\|^2} i'(\mathbf{q}). \quad (6)$$

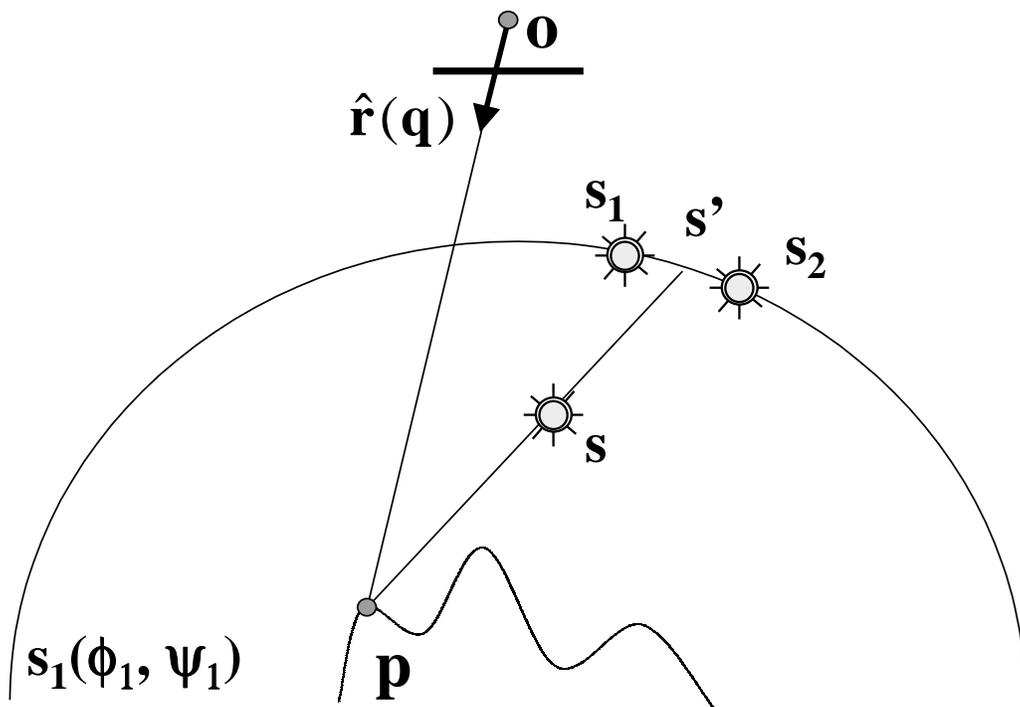


Figure 5: To render pixel \mathbf{q} (determine its intensity $i(\mathbf{q})$) for point light source \mathbf{s} which does not lie on $\mathbf{s}_1(\phi_1, \psi_1)$, the reconstructed point \mathbf{p} is required. The intersection \mathbf{s}' of the ray from \mathbf{p} through \mathbf{s} with the triangulated surface of light sources is determined. Based on the vertices (sample light sources) of the triangle containing \mathbf{s}' (e.g., \mathbf{s}_1 and \mathbf{s}_2 in this figure), the image intensity $i(\mathbf{q})$ is determined by interpolating the measured pixel intensities in the images formed under light sources located at the vertices.

For each pixel in the synthesized images, this procedure can be performed independently. Images for multiple point light sources $\{\mathbf{s}_1, \dots, \mathbf{s}_n\}$ can be synthesized simply through superposition of the images $I_1 \dots I_n$ formed for each light source, weighted by the relative strength of the light sources.

We have implemented this approach and tested it using the same synthetic images mentioned above. In the first test, we created a movie in which the light source starts off at infinity to the left of the teapot, heads in a direction parallel to the image plane and toward the optical axis, takes a 90° turn along the optical axis, and heads back out to infinity. The path of the light source is shown in the top of Fig. 6, and sample frames from the movie are shown in bottom of Fig. 6. Notice the motion of the specularity and how the shape of the shadows change as the light source moves. In the second test, we used our method to synthetically generate two images of the teapot with a diffuse light source, see Fig. 7.a. Finally in the third test, we used our method to generate two images of the teapot with an anisotropic spotlight, see Fig. 7.b.

4 Discussion

The major contributions of the paper are the following:

- We have presented a method for reconstructing the shape of an object from images in which the object is illuminated by single point light sources located at different positions. The

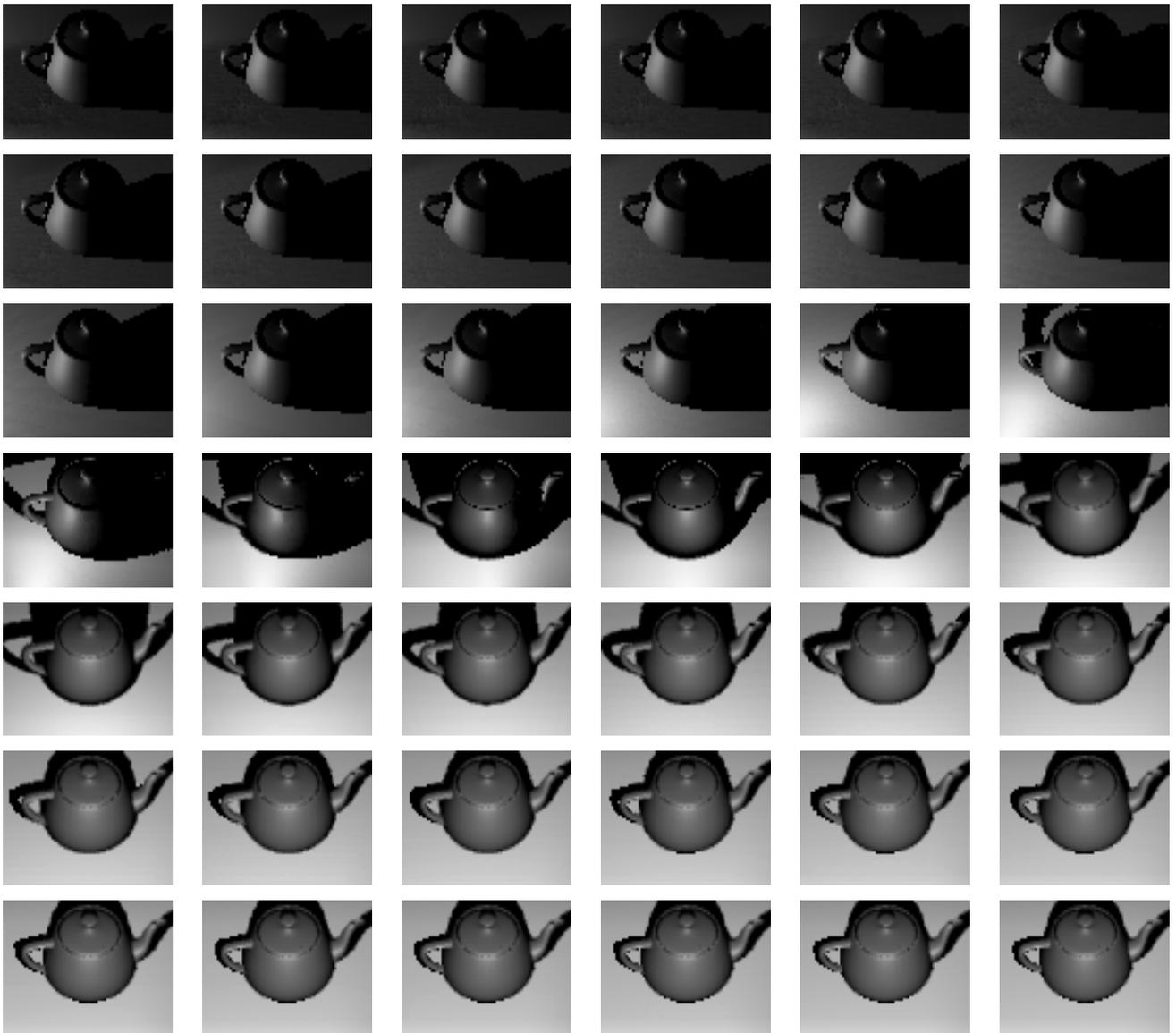
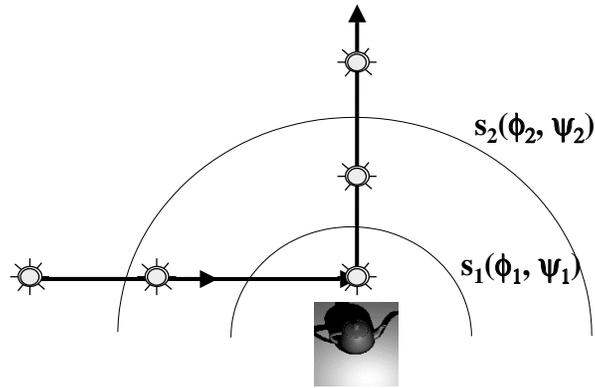


Figure 6: The rendered images are frames from a movie of a simulated moving point light source. As shown above, the point light source starts off far to the left of the teapot, heads in a direction parallel to the image plane and toward the optical axis, takes a 90° turn along the optical axis, and heads away from the teapot. The image corresponding to the turn is in the fourth row and fourth column.

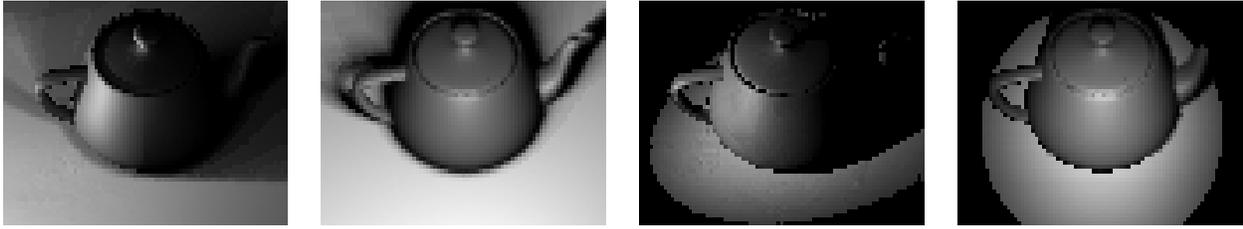


Figure 7: a. The left two images were synthetically generated using diffuse light source to the left and center of the teapot; b. The right two images were synthetically generated using an anisotropic spotlight to the left and center of the teapot.

method makes no assumptions about the shape of the object, the surface BRDF, or the presence or absence of shadows.

- We have presented a method for rendering novel images of the object under arbitrary lighting conditions. The method correctly handles shadowing without the need for ray tracing and can synthesize point, anisotropic, extended, or any other type of light source.

As yet, we have not implemented this method on images of real objects, since this requires a special rig for moving the light source. We are currently in the process of building this light source rig and hope to generate reconstructions and synthetically rendered images for a large database of objects.

There are many issues to explore: What is the relationship of the BRDF and geometry to the necessary sampling rate of light sources to yield effective reconstructions and renderings? What are efficient ways to compress the presumably redundant information for most scenes? What are fast ways to render images using the resulting representation? How can such methods be extended to handle different viewpoints as well as illumination. Note that while we have reconstructed a 3-D surface which could be viewed from an arbitrary viewpoint, we only recover a 2-D slice of the BRDF at each point. Are there principled means to extrapolate the apparent 4-D BRDF from the 2-D slice?

Currently, our method for reconstruction requires that the light source positions be known. It would be preferable to simply “wave” a light over the object during image acquisition, and then to simultaneously estimate the light source position and structure during reconstruction. While we are exploring methods for doing this, we wonder if there are inherent ambiguities such as the generalized bas-relief ambiguity found in reconstruction of Lambertian surfaces from images where the positions of the light source at infinity are unknown [2].

References

- [1] S. Avidan and A. Shashua. Novel view synthesis in tensor space. In *Proc. IEEE Conf. on Comp. Vision and Patt. Recog.*, pages 1034–1040, 1997.
- [2] P. Belhumeur, D. Kriegman, and A. Yuille. The bas-relief ambiguity. In *Proc. IEEE Conf. on Comp. Vision and Patt. Recog.*, pages 1040–1046, San Jaun, PR, 1997.
- [3] P. N. Belhumeur and D. J. Kriegman. What is the set of images of an object under all possible lighting conditions. *Int. J. Computer Vision*, 28(3):245–260, 1998.

- [4] P. Debevec, C. Taylor, and J. Malik. Modeling and rendering architecture from photographs: A hybrid geometry- and image-based approach. In *SIGGRAPH*, 1996.
- [5] E. Edelson and J. Bergen. Computational models of visual processing. In Landy and Movshon, editors, *The Plenoptic Function*. MIT Press, 1991.
- [6] Y. Genc and J. Ponce. Parameterized image varieties: A novel approach to the analysis and synthesis of image sequences. In *Int. Conf. on Computer Vision*, pages 11–16, 1998.
- [7] S. Gortler, R. Grzeszczuk, R. Szeliski, and M. Cohen. The lumigraph. In *SIGGRAPH*, pages 43–54, 1996.
- [8] H. Hayakawa. Photometric stereo under a light-source with arbitrary motion. *J. Optical Society of America A*, 11(11):3079–3089, Nov. 1994.
- [9] K. Ikeuchi. Determining surface orientations of specular surfaces by using the photometric stereo method. *IEEE Trans. Pattern Anal. Mach. Intelligence*, 3(6):661–669, 1981.
- [10] K. Kutulakos and J. Vallino. Calibration-free augmented reality. *IEEE Trans. Visualization and Computer Graphics*, 4(1):1–20, 1998.
- [11] M. Langer and S. Zucker. What is a light source? In *Proc. IEEE Conf. on Comp. Vision and Patt. Recog.*, pages 172–178, San Jaun, PR, 1997.
- [12] M. Langer and S. W. Zucker. Shape-from-shading on a cloudy day. *J. Opt Soc. Am.*, pages 467–478, 1994.
- [13] M. Levoy and P. Hanrahan. Light field rendering. In *SIGGRAPH*, pages 31–42, 1996.
- [14] J. Nimeroff, E. Simoncelli, and J. Dorsey. Efficient re-rendering of naturally illuminated environments. In *Eurographics Symposium on Rendering*, Darmstadt, Germany, June 1994.
- [15] R. Onn and A. Bruckstein. Integrability disambiguates surface recovery in two-image photometric stereo. *Int. J. Computer Vision*, 5(1):105–113, 1990.
- [16] S. Seitz and C. Dyer. View morphing. In *SIGGRAPH*, pages 21–30, 1996.
- [17] A. Sashua. On photometric issues in 3D visual recognition from a single image. *Int. J. Computer Vision*, 21:99–122, 1997.
- [18] W. Silver. *Determining Shape and Reflectance Using Multiple Images*. PhD thesis, MIT, 1980.
- [19] H. Tagare and R. deFigueiredo. A theory of photometric stereo for a class of diffuse non-lambertian surfaces. *IEEE Trans. Pattern Anal. Mach. Intelligence*, 13(2):133–152, February 1991.
- [20] T. Wong, P. Heng, S. Or, and W. Ng. Illuminating image-based objects. In *Proceedings of Pacific Graphics*, pages 69–78, Seoul, Oct. 1997.
- [21] R. Woodham. Analysing images of curved surfaces. *Artificial Intelligence*, 17:117–140, 1981.