International Journal of Computer Vision 49(2/3), 215–227, 2002
© 2002 Kluwer Academic Publishers. Manufactured in The Netherlands.

Helmholtz Stereopsis: Exploiting Reciprocity for Surface Reconstruction

TODD E. ZICKLER
Electrical Engineering, Yale University, New Haven, CT 06520-8285, USA
zickler@yale.edu

PETER N. BELHUMEUR
Computer Science, Columbia University, New York, NY 10027, USA
belhumeur@cs.columbia.edu

DAVID J. KRIEGMAN
Computer Science and Engineering, University of California, San Diego, LaJolla, CA 94143-1290, USA
kriegman@cs.ucsd.edu

Received December 6, 2001; Revised December 6, 2001; Accepted March 1, 2002

Abstract. We present a method—termed Helmholtz stereopsis—for reconstructing the geometry of objects from a collection of images. Unlike existing methods for surface reconstruction (e.g., stereo vision, structure from motion, photometric stereopsis), Helmholtz stereopsis makes no assumptions about the nature of the bidirectional reflectance distribution functions (BRDFs) of objects. This new method of multinocular stereopsis exploits Helmholtz reciprocity by choosing pairs of light source and camera positions that guarantee that the ratio of the emitted radiance to the incident irradiance is the same for corresponding points in the two images. The method provides direct estimates of both depth and surface normals, and consequently weds the advantages of both conventional stereopsis and photometric stereopsis. Results from our implementation lend empirical support to our technique.

Keywords: BRDF, reflectance, surface reconstruction, stereo, photometric stereo, Helmholtz reciprocity

1. Introduction

In this paper, we present Helmholtz stereopsis, a novel method for reconstructing the geometry of a surface that has arbitrary and unknown surface reflectance. This method does *not* make the ubiquitous assumption that the reflectance is Lambertian or of some other parametric form, and it enables the reconstruction of surfaces for which the reflectance is anisotropic, and for which it varies from point to point across the surface. Helmholtz stereopsis works by exploiting the symmetry of surface reflectance—pairs of light source and camera positions are chosen to guarantee that the relationship between pixel values at corresponding image points depends only on the shape of the surface (and is independent of the reflectance).

At a suitable scale, reflectance is accurately described by the bidirectional reflectance distribution function, or BRDF (Nicodemus et al., 1977). The BRDF of a surface point, denoted $f_r(\hat{\mathbf{i}}, \hat{\mathbf{e}})$, is the ratio of the outgoing radiance to the incident irradiance. Here, $\hat{\mathbf{i}}$ is the direction of an incident light ray, and $\hat{\mathbf{e}}$ is the direction of the outgoing ray. These are typically written as directions in a coordinate frame attached to the tangent plane of the surface. It is not an arbitrary four dimensional function since, in general, it is symmetric about the incoming and outgoing angles: $f_r(\hat{\mathbf{i}}, \hat{\mathbf{e}}) = f_r(\hat{\mathbf{e}}, \hat{\mathbf{i}})$. This symmetry condition is a

generalization of a principle of reciprocity described by Helmholtz (1925, p. 231) and is commonly referred to as Helmholtz reciprocity. (It has been pointed out in the literature that reciprocity may not hold for some surfaces (Helmholtz, 1925; Clarke and Parry, 1985; Snyder et al., 1998)—a discussion of these cases is outside the scope of this paper).

In computer vision and computer graphics, models are used to simplify the BRDF. In computer vision, the assumption that surfaces are Lambertian is the basis for most reconstruction techniques. In computer graphics, the vast majority of rendered images use the Phong reflectance model which is composed of an ambient term, a diffuse (Lambertian) term and an ad hoc specular term (Phong, 1975). While the isotropic Phong model captures the reflectance properties of plastics over a wide range of conditions, it does not effectively capture the reflectance of materials such as metals and ceramics, particularly when they have rough surfaces or a regular surface structure (e.g., parallel grooves). Much less common are a number of physics-based parametric models (Oren and Nayar, 1996; Torrance and Sparrow, 1967; Cook and Torrance, 1981; He et al., 1992; Koenderink et al., 1999; Ashikhmin et al., 2000), and each of these only characterizes a limited class of surfaces. A recent alternative to parametric models is the measurement of the BRDF and its representation by a suitable set of basis functions (Koenderink and van Doorn, 1996). In contrast to these approaches, this paper is concerned with surfaces with *arbitrary* BRDFs—those for which we have no information a priori. (Of course, this includes all of the BRDF models mentioned above.)

To see how Helmholtz reciprocity can be used for stereopsis, consider obtaining a pair of images as shown in Fig. 1. The first image is captured while the object is illuminated by a single point light source, and the second image is captured once the camera and light source positions have been swapped. That is, the camera's center of projection is moved to the former location of the light source, and vice versa. By acquiring images in this manner, Helmholtz reciprocity ensures that, for any visible scene point, the ratio of the emitted radiance (in the direction of the camera) to the incident irradiance (from the direction of the light source) is the same for both images. This is not true for general stereo pairs that are acquired under fixed illumination (unless the BRDFs of the surfaces are Lambertian).

We will show that three or more pairs of images acquired in this manner provide a matching constraint,

Figure 1. The setup for acquiring a pair of images that exploits Helmholtz reciprocity. First an image is acquired with the scene illuminated by a single point source as shown on the left. Then, a second image is acquired after the positions of the camera and light source are exchanged as shown on the right.

which leads to a multinocular stereo imaging geometry. These images contain sufficient information to establish a constraint that can be used to solve the correspondence problem (and thereby solve for depth). In addition, they contain sufficient information to directly estimate the surface normal at each point without taking derivatives of either the images or the depth map. The direct estimation of surface orientation is similar to photometric stereopsis, but here the BRDF may be unknown and arbitrary.

The paper is organized as follows. In the next section, we derive the relationship between image irradiance values at corresponding pixels in a reciprocal pair of images, and demonstrate a special case in which we can recover depth from a single reciprocal pair. Section 3 describes the complete multinocular reciprocity-based method in detail. Since the method combines the advantages of conventional multinocular stereopsis (direct estimation of depth) with those of photometric stereopsis (direct estimation of surface normals), the similarities and differences of these methods are summarized in Section 4 (see Fig. 4). Finally, Section 5 describes the experimental results of our implementation.

2. Reciprocal Image Pairs

In order to examine the relationship between irradiance values at corresponding image points, consider the imaging geometry shown in the left half of Fig. 1. As shown in that figure, $\mathbf{o}_l$ and $\mathbf{o}_r$ denote the positions of the camera and light source, respectively. We also denote by $\mathbf{p}$ and $\hat{\mathbf{n}}$ a point on the surface and its associated

unit normal vector. The unit vectors $\hat{\mathbf{v}}_\mathrm{l} = \frac{1}{|\mathbf{o}_\mathrm{l}-\mathbf{p}|}(\mathbf{o}_\mathrm{l}-\mathbf{p})$, and $\hat{\mathbf{v}}_\mathrm{r} = \frac{1}{|\mathbf{o}_\mathrm{r}-\mathbf{p}|}(\mathbf{o}_\mathrm{r}-\mathbf{p})$ denote the directions from $\mathbf{p}$ to the camera and light source, respectively. Given this system, the image irradiance at the projection of $\mathbf{p}$ is given by

$$i_\mathrm{l} = f_r(\hat{\mathbf{v}}_\mathrm{r}, \hat{\mathbf{v}}_\mathrm{l})\frac{\hat{\mathbf{n}}\cdot\hat{\mathbf{v}}_\mathrm{r}}{|\mathbf{o}_\mathrm{r}-\mathbf{p}|^2} \tag{1}$$

where $\hat{\mathbf{n}}\cdot\hat{\mathbf{v}}_\mathrm{r}$ gives the cosine of the angle between the direction to the light source and the surface normal, $\frac{1}{|\mathbf{o}_\mathrm{r}-\mathbf{p}|^2}$ is the $1/r^2$ fall-off from a unit-strength, isotropic point light source, and f_r is the BRDF.

Now, consider the reciprocal case in which the light source is positioned at $\mathbf{o}_\mathrm{l}$, and the camera observes $\mathbf{p}$ from $\mathbf{o}_\mathrm{r}$. In this case, the image irradiance is

$$i_\mathrm{r} = f_r(\hat{\mathbf{v}}_\mathrm{l}, \hat{\mathbf{v}}_\mathrm{r})\frac{\hat{\mathbf{n}}\cdot\hat{\mathbf{v}}_\mathrm{l}}{|\mathbf{o}_\mathrm{l}-\mathbf{p}|^2}. \tag{2}$$

Because of Helmholtz reciprocity, we have $f_r(\hat{\mathbf{v}}_\mathrm{r}, \hat{\mathbf{v}}_\mathrm{l}) = f_r(\hat{\mathbf{v}}_\mathrm{l}, \hat{\mathbf{v}}_\mathrm{r})$, and we can eliminate the BRDF term in the above two equations to obtain

$$\left(i_\mathrm{l}\frac{\hat{\mathbf{v}}_\mathrm{l}}{|\mathbf{o}_\mathrm{l}-\mathbf{p}|^2} - i_\mathrm{r}\frac{\hat{\mathbf{v}}_\mathrm{r}}{|\mathbf{o}_\mathrm{r}-\mathbf{p}|^2}\right)\cdot\hat{\mathbf{n}} = \mathbf{w}(d)\cdot\hat{\mathbf{n}} = 0. \tag{3}$$

In this equation, i_l and i_r are irradiance measurements obtained from a radiometrically calibrated camera. Also, for geometrically calibrated cameras and a value for the binocular disparity (or equivalently the depth d), the values for $\mathbf{o}_\mathrm{l}$ and $\mathbf{o}_\mathrm{r}$ are known, and the values for $\mathbf{p}$, $\hat{\mathbf{v}}_\mathrm{l}$, and $\hat{\mathbf{v}}_\mathrm{r}$ can be computed (we write $\mathbf{w}(d)$ to denote this fact). It follows that only the surface normal $\hat{\mathbf{n}}$ and the depth d are unknown. Note that the vector $\mathbf{w}(d)$ lies in the plane defined by $\mathbf{p}$, $\mathbf{o}_\mathrm{r}$ and $\mathbf{o}_\mathrm{l}$ (the epipolar plane).

Equation (3) provides a constraint on pixel values of corresponding image points, and unlike similar constraints used by conventional stereopsis, this constraint is independent of the BRDF—it depends solely on the shape of the object (the depth d and surface normal $\hat{\mathbf{n}}$). However, given that there are three degrees of freedom and only a single constraint, we cannot, in general, recover this information from a single pair of images. A multinocular constraint that enables recovery of both the depth and the normal field is developed in Section 3.

There are two more things to note about the constraint in (3). First, in deriving this constraint we ignore interreflection effects. Second, an interesting historical note is that 60 years ago (long before the definition of the BRDF was introduced), Minnaert (1941) derived a special case of this constraint that was used to increase the number of lunar reflectance measurements that could be made from Earth.

As mentioned above, in general we cannot recover information about the surface from a single pair of images. However, in the next section we discuss a case in which a single reciprocal pair *can* provide enough information for depth reconstruction.

2.1. *A Special Case: Fronto-Parallel Objects*

This section describes a special case in which we can recover the depth of the scene from a single reciprocal pair. While the limitations may be too great to make this special case useful in practice, it demonstrates some important properties of Helmholtz stereopsis.

Consider the reciprocity constraint in (3). When the stereo rig has a small baseline relative to the scene depth, we can write

$$|\mathbf{o}_\mathrm{l}-\mathbf{p}|^2 \approx |\mathbf{o}_\mathrm{r}-\mathbf{p}|^2, \tag{4}$$

and if the surfaces are nearly fronto-parallel, we have

$$\hat{\mathbf{n}}\cdot\hat{\mathbf{v}}_\mathrm{l} \approx \hat{\mathbf{n}}\cdot\hat{\mathbf{v}}_\mathrm{r} \approx 1. \tag{5}$$

Using these approximations the matching constraint (3) reduces to

$$i_\mathrm{l} = i_\mathrm{r}. \tag{6}$$

That is, correspondence can be established simply by comparing pixel intensities across the epipolar lines in the two images just as in standard stereo vision algorithms. Recall that unlike standard stereo, we have lit the scene differently for the two images.

Figure 2(a) shows a reciprocal image pair that satisfies these assumptions. Note that the specularities occur at the same locations in both images, as predicted by Helmholtz reciprocity. Thus, the specularities become features in both images which can actually aid in establishing correspondence. Also note that shadowed regions correspond identically to half-occluded regions in both images—if a point is in shadow in the left image, it is not visible in the right image, and vice versa.

To establish correspondence between the two images shown in Fig. 2(a), we implemented the "World II" stereo algorithm described by Belhumeur (1993).

Figure 2. Result of stereo matching applied to a reciprocal pair: (a) a stereo pair of images acquired by swapping the camera and light source, and (b) the disparity map.

Figure 3. Result of stereo matching applied to a conventional stereo pair: (a) a stereo pair from the same camera positions as in Fig. 2, but under fixed lighting; and (b) the disparity map.

We chose this algorithm both because it is intensity based (not edge-based) and because it implicitly resolves half-occluded regions by linking them to depth discontinuities. The result for our implementation of this algorithm applied to the stereo pair in Fig. 2(a) is shown in Fig. 2(b).

We also gathered a standard stereo pair (as shown in Fig. 3(a)) in which the lighting remained fixed for both the left and right images. The stereo pair in Fig. 3(a) differs from that in Fig. 2 only in the illumination—the positions of the cameras and the scene geometry are identical. The result for our implementation of the same algorithm applied to the standard stereo pair is shown in Fig. 3(b). Note that we used the same procedure to establish correspondences for the new pair of images. Although the accuracy of the stereo matching may have been improved by pre-filtering the images, we avoided this to make the point that image intensity is very much viewpoint dependent.

There are two things to note about the results. First, the reciprocal images in Fig. 2 have significant specularities, but they remain fixed in the images and do not hinder stereo matching. Contrast this with the images in Fig. 3. These also have specularities (as seen on the frame and on the glass) and non-Lambertian effects, but these effects change between images and significantly hinder matching. Second, there is little texture on the background wall, yet the reciprocal images allow the stereo algorithm to estimate the depth discontinuity at the boundary of the picture frame, because the half-occluded regions and visible shadows are in correspondence.

The properties of Helmholtz stereopsis are further discussed in Section 4, but first we will develop a multinocular constraint based on (3) that will allow the recovery of depth and surface normals for general surfaces.

3. Helmholtz Stereopsis

In this section we describe our method for reconstructing surfaces with arbitrary BRDFs using a form of multinocular stereopsis. Before describing Helmholtz stereopsis, however, it will be helpful to provide a framework for general n-view stereopsis. (This is a generalization of the correspondence problem in conventional binocular stereopsis.) Consider n calibrated

cameras whose centers of projection are located at $\mathbf{o}_c$ for $c = 1, \ldots, n$. Define a camera centered at $\mathbf{o}_\mathrm{p}$ to be the principal camera. This camera is used to parametrize the depth search, and while it could be one of the cameras located at $\mathbf{o}_c$, it need not be a physical camera (i.e., it can be virtual). Given a point $\mathbf{q}$ in the principal image, there exists a one-parameter family of n-point sets $(\mathbf{q}_1, \ldots, \mathbf{q}_n)$—one point in each of the n images—that could correspond to $\mathbf{q}$. We parametrize this family by the depth d, and by defining a discrete set of possible values for $d (d \in D = \{d_0, \ldots, d_{\mathrm{N_D}}\})$ we can index this family of n-point sets, $Q(d) = \{\mathbf{q}_c(d), c = 1, \ldots, n\}$.

A multinocular matching constraint provides a method of deciding, given a set of image intensities measured at the points $Q(d)$, whether or not the hypothesized depth value d could correspond to a true surface point. In the case of traditional dense stereo, the surface is assumed to be Lambertian, and the constraint is simply $I_1(\mathbf{q}_1(d)) = I_2(\mathbf{q}_2(d)) = \cdots = I_n(\mathbf{q}_n(d))$ where $I_c(\mathbf{q}_c)$ is the intensity at point $\mathbf{q}_c$ in the image centered at $\mathbf{o}_c$. (Note that many other stereo methods exist in which the constraint involves filtered intensities as opposed to the image intensities themselves.)

Using this framework, we can proceed to develop a matching constraint for reciprocal image pairs. What is unique to Helmholtz stereopsis, is that this constraint is independent of the BRDF, and that it allows the direct recovery of both the depth *and* surface normals.

Suppose we capture N_P reciprocal pairs of images as described in Section 2, and suppose that each of these pairs is captured from a different pair of positions $(\mathbf{o}_{\mathrm{l}j}, \mathbf{o}_{\mathrm{r}j})$, $j = 1, \ldots, N_\mathrm{P}$. We can form N_P linear constraints like that in (3). Define $\mathbf{W}(d) \in \mathbb{R}^{N_\mathrm{P} \times 3}$ to be the matrix in which the jth row is given by $\mathbf{w}_j(d)^\mathsf{T} = (i_{\mathrm{l}j} \frac{\hat{\mathbf{v}}_{\mathrm{l}j}}{|\mathbf{o}_{\mathrm{l}j} - \mathbf{p}|^2} - i_{\mathrm{r}j} \frac{\hat{\mathbf{v}}_{\mathrm{r}j}}{|\mathbf{o}_{\mathrm{r}j} - \mathbf{p}|^2})^\mathsf{T}$. Then the set of constraints from (3) can be expressed as

$$\mathbf{W}(d)\,\hat{\mathbf{n}} = 0. \tag{7}$$

Clearly, for a correct depth value $d^\star$, the surface normal lies in the null space of $\mathbf{W}(d^\star)$, and it can be estimated from a noisy matrix using singular value decomposition. In addition, $\mathbf{W}(d^\star)$ will be rank 2, and this can be used as a necessary condition when searching for the surface depth (i.e., for solving the correspondence problem). Note that at least three camera/light source pairs are needed to exploit this constraint.

An implementation of a system that uses this constraint for surface reconstruction is discussed in Section 5. Next, we present a comparison of Helmholtz stereopsis with some existing reconstruction techniques.

4. Comparison with Existing Methods

In principle, Helmholtz Stereopsis has a number of advantages when compared to conventional multinocular stereopsis and photometric stereopsis. This section compares these methods in four separate categories. A summary of the information in this section is contained in Fig. 4. While our implementation may not fully reveal these advantages (we do not make explicit use of available half-occlusion indicators for detecting depth discontinuities), we believe that future refinements will.

4.1. Assumed BRDF

Most conventional dense stereo reconstruction methods assume that scene radiance is independent of viewing direction, i.e. that surface reflectance is Lambertian. However, the majority of surfaces are not Lambertian and therefore violate this assumption. For these surfaces, large-scale changes in scene radiance occur as specularities shift with viewpoint, and small-scale changes occur everywhere on the surface. In addition, if the BRDF is spatially varying, these changes may occur differently at every point on the surface. Using traditional dense stereopsis, establishing correspondence in this situation is difficult, if at all possible. Most sparse, or feature-based, stereo methods also rely (albeit less heavily) on the Lambertian assumption—if the BRDF is arbitarary, the detected feature points may be viewpoint or lighting dependent.

Whereas viewpoint is manipulated in conventional stereopsis, in photometric stereopsis, the viewpoint remains fixed while the illumination is varied. Photometric stereo methods provide an estimate of the field of surface normals which is then integrated to recover the surface depth. Similar to conventional multinocular stereopsis, many photometric methods assume that the BRDF is Lambertian (Langer and Zucker, 1994; Silver, 1980; Woodham, 1981). The methods that do not make this assumption either assume that the BRDF is completely known a priori, or can be specified using a small number of parameters (Hayakawa, 1994; Ikeuchi and Horn, 1981; Nayar et al., 1990; Tagare and deFigueiredo, 1991). As mentioned in the introduction, these parametric BRDFs are often derived from

Property / Method	Assumed Reflectance	Surface Information Recovered			Handles Cast Shadows	Handles Half-Occlusion	Active/ Passive
		Textured Albedo	Constant Albedo	Depth Discontinuities			
Photometric Stereopsis	Lambertian or Known	Surface Normals	Surface Normals	No	No	NA	Active
Multinocular Stereopsis	Lambertian	Depth	Nothing	Sometimes	Yes	No/Yes	Passive
Helmholtz Stereopsis	Arbitrary	Depth + Surface Normals	Surface Normals	Yes	Yes	Yes	Active

Figure 4. A comparison of Helmholtz stereopsis with conventional multinocular and photometric stereopsis. A detailed discussion of the entries in this table is given in Section 4.

physical models of reflectance and are restricted to a limited class of surfaces. When the form of the BRDF is unknown, or when the form of the BRDF is spatially varying, there is insufficient information to reconstruct both the geometry and the BRDF.

Lu and Little (1999) presented a hybrid method with controlled lighting and object rotation that was used to estimate both surface structure and a non-parametric reflectance map. This is similar to our method in that it: (1) is an active imaging technique that exploits changes in viewpoint *and* illumination; and (2) considers a general, non-parametric BRDF. However, the method requires that the BRDF is both isotropic and uniform across the surface (the present method makes no such assumptions).

Another reconstruction method for surfaces with arbitrary BRDFs was introduced (along with our preliminary work on Helmholtz stereopsis) by Magda et al. (2001). In addition to recovering depth, the method also enables the recovery of a 2-D slice of the *apparent BRDF* (a coupling of the reflectance and orientation information) at each point on the surface. It does not, however, enable the explicit recovery of the normal field, and it requires many more images of the object.

The assumptions made about surface reflectance for three reconstruction techniques—conventional, photometric, and Helmholtz stereopsis—are summarized diagrammatically in Fig. 5. Note that many natural surfaces actually have surface reflectance in the rightmost region of the figure and cannot be accurately reconstructed by conventional techniques.

In Helmholtz stereopsis, because the relationship between image intensities of corresponding points does not depend on viewpoint, non-Lambertian radiometric events such as specularities appear fixed to the surface of the object. In contrast with conventional stereo (fixed illumination) images, these radiometric events become reliable features, and they actually simplify the correspondence problem.

Figure 5. A summary of the assumptions made about surface reflectance by three reconstruction techniques. Both conventional multinocular stereopsis and photometric stereopsis assume the BRDF is Lambertian or of some other known parametric form. Yet, many natural surfaces (e.g., human skin, the skin of a fruit, glossy paint) do not satisfy these assumptions. In contrast to the other methods, Helmholtz stereopsis makes no assumption about the BRDF.

4.2. *Recovered Surface Information*

In conventional binocular or multinocular stereopsis, depth is readily computed. Typically, the output of the system is a discrete set of depth values at pixel or sub-pixel intervals—a depth map. In most cases, unless a regularization process is used to smooth the depth estimates, the normal field found by differentiating the recovered depth map will be very noisy. Instead of direct differentiation of the depth map, regularized estimates of the normal field can be obtained, for example, based on an assumption of local planarity (Deverney and Faugeras, 1994), or through the use of an energy functional (Belhumeur, 1996). In contrast to these methods, photometric stereopsis provides a direct estimate of the field of surface normals which is

then integrated (in the absence of depth discontinuities) to obtain a surface. Helmholtz stereopsis is similar to photometric stereopsis (and different from the regularization techniques used in conventional stereopsis) in that the normal field is directly estimated on a point-by-point basis using the photometric variation across reciprocal image pairs.

In this way, Helmholtz stereopsis combines the advantages of conventional and photometric methods by providing both a direct estimate of the surface depth *and* the field of surface normals. It also provides information about the location of depth discontinuities (see below). Note that for applications such as image-based rendering and image-based modeling, a good estimate of the normal field is critical for computing intensities and accurately measuring reflectance properties.

Figure 6. A summary of the surface properties required for Lambertian surface reconstruction by conventional and Helmholtz stereo techniques. Even when the BRDF is Lambertian, conventional stereopsis is only capable of recovering surface geometry in regions of texture (i.e., varying albedo) or high curvature (i.e., edges). Neither photometric stereopsis nor Helmholtz stereopsis suffer from this limitation.

4.3. Constant Intensity Regions

Dense stereo and motion methods work best when the surfaces are highly textured; when they are not textured, regularization is needed to infer the surface. This can be achieved, for example, using a statistical prior (Geiger et al., 1992; Matthies, 1992; Poggio et al., 1985; Belhumeur, 1996) or through surface evolution (Faugeras and Keriven, 1998). Sparse stereo and motion methods also have difficulty in these regions. These methods only reconstruct the geometry of corresponding feature points, so by their nature, they cannot directly reconstruct smoothly curving surfaces whose reflectance properties are constant. In contrast, photometric stereo techniques and Helmholtz stereopsis are unaffected by lack of texture, since they can effectively estimate the field of normals which is then integrated to recover depth (see Fig. 6).

4.4. Depth Discontinuities

Depth discontinuities present difficulties for both traditional and photometric stereopsis. When there is a depth discontinuity, it does not make sense to integrate the normal field that is output by photometric stereo techniques. Likewise, traditional stereo algorithms often have trouble locating depth discontinuities. This difficulty arises for two reasons. First, if a background object has regions of constant intensity and the discontinuity in depth occurs within one of these regions, it is quite difficult to reliably locate the boundary of the foreground object. Second, depth discontinuities induce half-occlusion in adjacent regions of the image, and these regions, which are not visible in at least one of the images, often confuse the matching process.

Helmholtz stereopsis simplifies the task of detecting depth discontinuities since, as seen in the example in Fig. 2, the lighting setup is such that the shadowed and half-occluded regions are in correspondence. The shadowed regions in the images of a Helmholtz pair can therefore be used to locate depth discontinuities. As shown in that example, if one uses a stereo matching algorithm that exploits the presence of half-occluded regions for determining depth discontinuities (Belhumeur, 1993; Belhumeur and Mumford, 1992; Cox et al., 1992; Geiger et al., 1992), these shadowed regions may significantly enhance the quality of the depth reconstruction.

4.5. Active vs. Passive Imaging

Like photometric stereopsis and unlike conventional stereopsis, Helmholtz stereopsis is active. The scene is illuminated in a controlled manner, and images are acquired as lights are turned on and off. Clearly a suitable optical system can be constructed so that the camera and light source are not literally moved, but rather a virtual camera center and light source are co-located. Alternatively, as will be shown in the next section, a simple system can be developed that captures multiple reciprocal image pairs with a single camera and a single light source.

5. Implementation and Results

In the previous sections a number of claims were made about the capabilities of Helmholtz stereopsis as a reconstruction technique. This section describes an implementation of a Helmholtz stereo system, and gives results that support those claims. Specifically, in this section, we give examples of:

- the reconstruction of surfaces with arbitrary, spatially varying BRDFs (surfaces that are neither Lambertian nor approximately Lambertian)
- direct recovery of both surface depth and the field of surface normals
- the reconstruction of surfaces in regions of constant image brightness

5.1. Capturing Reciprocal Images

To demonstrate Helmholtz stereopsis, we constructed a system that enables the acquisition of multiple reciprocal image pairs with a single camera and a single light source. These are mounted on a wheel as shown schematically in Fig. 7(a). First, suppose an image is captured with the wheel in the position shown in this figure. If the wheel is rotated by 180° and another image is captured, these two images will form a reciprocal pair, and corresponding image irradiance values will satisfy the constraint in (3). It is clear that we can capture any number of reciprocal pairs by rotating the wheel through 360° while stopping to capture images at reciprocal positions.

Figure 7. (a) A wheel is used to capture multiple reciprocal image pairs employing a single camera and a single light source. By rotating the wheel through 360°, any number of fixed-baseline pairs can be captured. For example, images captured at $\mathbf{o}_{l2}$ and $\mathbf{o}_{r2}$ will form a reciprocal pair. (b) An example of the wheel design shown in (a). The light source consists of a standard 100 W frosted incandescent bulb fitted with a small aperture.

A picture of such a system is shown in Fig. 7(b). The camera is a Nikon Coolpix 990, and the light source consists of a standard 100 W frosted incandescent bulb fitted with a small aperture. The camera is both geometrically and radiometrically calibrated. The former means that the intrinsic parameters and the extrinsic parameters of each camera position are known, while the latter means that we know the mapping from scene radiance values to pixel intensities (including optical fall-off, vignetting, and the radiometric camera response). Since the lamp is not an ideal isotropic point source, it also requires a radiometric calibration procedure in which we determine its radiance as a function of output direction.

An example of a set of images captured using this system is shown in Fig. 8. For all results shown in this paper the diameter of the wheel was 38 cm and the distance from the center of the wheel to the scene was approximately 60 cm. The reconstructions were performed from the viewpoint of a virtual principal camera located at the center of the wheel. We chose this camera to be orthographic to ensure uniform sampling of object space.

5.2. Using the Matching Constraint

In Section 3, we derived a matrix constraint that can be used to recover the surface depth and orientation corresponding to each pixel $\mathbf{q}$ in the principal view. How this constraint should be used was not specified; there are a number of possible methods, many of which can be adapted from conventional stereo algorithms. Our goal is to demonstrate the feasibility of Helmholtz stereopsis in general, so a discussion of possible methods is outside the scope of this paper. Instead, we have chosen one particularly simple implementation which will be described here. Results for four different surfaces follow in the next section.

For each pixel $\mathbf{q}$, and for each depth value $d \in D = \{d_1, d_2, \ldots, d_{N_D}\}$ we can construct a matrix $\mathbf{W}_{\mathbf{q}}(d)$ as in (7). If the hypothesized depth corresponds to a true surface point, this matrix will be rank 2, and the surface normal will be uniquely determined as the unit vector that spans its 1-D null space. (Note that since each row of $\mathbf{W}$ (we denote these $\mathbf{w}_j^{\mathsf{T}}$) lies in the epipolar plane defined by $\mathbf{p}$, $\mathbf{o}_{lj}$, and $\mathbf{o}_{rj}$, no two rows of $\mathbf{W}$ will be collinear, so rank($\mathbf{W}$) ≥ 2).

In the presence of noise, $\mathbf{W}$ is generally rank 3, and we require a measure for the coplanarity of the row vectors $\mathbf{w}_j^{\mathsf{T}}$. Since we know that rank($\mathbf{W}$) ≥ 2, a suitable

Figure 8. An example of 6 reciprocal images pairs captured using the rig described in Fig. 7. Reciprocal image pairs are arranged vertically.

measure (and one that works well in practice) is the ratio of the second to third singular values of $\mathbf{W}$. Given a matrix $\mathbf{W}_{\mathbf{q}}(d)$, we compute the singular value decomposition $\mathbf{W} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^{\mathsf{T}}$ where $\boldsymbol{\Sigma} = \mathrm{diag}(\sigma_1, \sigma_2, \sigma_3)$, $\sigma_1 \geq \sigma_2 \geq \sigma_3$. Then, our support measure used in the depth search is the ratio

$$r_{\mathbf{q}}(d) = \frac{\sigma_2}{\sigma_3}. \tag{8}$$

Note that at correct depth values, the ratio $r_{\mathbf{q}}(d)$ will be large.

The condition shown in (7) is a necessary condition that will be satisfied by true values of surface depth, but it is not sufficient. One way to resolve the ambiguity is to make some assumptions about the shape of the surface. (The BRDF remains arbitrary). One of the simplest methods, analogous to SSD matching in conventional binocular stereo, is to assume that the surface depth is locally constant. In the search for the depth at principal image point $\mathbf{q}_\circ$, we consider the ratio $r_{\mathbf{q}}(d)$ at this point as well as at points in a small rectangular window W around $\mathbf{q}_\circ$. Then, the estimated depth at this point is given by

$$d^\star_{\mathbf{q}_\circ} = \arg\max_{d \in D} \sum_{\mathbf{q} \in W} r_{\mathbf{q}}(d). \tag{9}$$

Once we have estimated the depth $d^\star$, the linear least-squares estimate of the normal is

$$\hat{\mathbf{n}}^\star_{\mathbf{q}_\circ} = \arg\min_{\hat{\mathbf{n}}} \|\mathbf{W}_{\mathbf{q}_\circ}(d^\star)\hat{\mathbf{n}}\|^2, \quad \|\hat{\mathbf{n}}\| = 1, \tag{10}$$

which is simply given by the right singular vector corresponding to the smallest singular value of $\mathbf{W}_{\mathbf{q}_\circ}(d^\star)$.

Note that the depth map that is recovered using (9) will have low resolution due to the assumption of local depth constancy. This initial estimate of the depth can be refined using the high frequency information provided by the field of surface normals. An example of this will be shown in the next section.

As a final note, this algorithm makes no attempt at detecting half-occluded regions even though this information is available through the visible shadows. We have chosen this method simply to demonstrate that reciprocity can be exploited for reconstruction. As shown in the next section, despite the simplicity of the method, the surface reconstructions are quite good.

5.3. *Results*

Figures 9–12 show the results of applying this procedure to four different objects. Each figure consists of: (a) one of the input images of the object, (b) the depth recovered using (9), and (c) the recovered field of surface normals. Note that the viewpoints of the displayed images differ slightly from the reconstruction viewpoints due to the use of a virtual principal camera.

Figure 9 is a demonstration of a surface reconstruction in the case of nearly constant image brightness. This surface (a wax candle) is a member of the class of surfaces described at the top of Fig. 6, and it is an example of a case in which conventional stereopsis has difficulty. Notice that Helmholtz stereopsis accurately estimates the normal field, even though the depth estimates are poor. The poor depth estimates are expected since at a principal image point $\mathbf{q}$, the ratio $r_{\mathbf{q}}(d)$ will be nearly constant for a small depth interval about the true surface depth. The normals are accurate, however, since each corresponding $\mathbf{W}_{\mathbf{q}}(d)$ will have nearly the same null space.

Figure 10 shows the results for a surface that is clearly non-Lambertian. The specularities on the nose, teeth and feet attest to this fact. Note that the reconstruction method is not expected to succeed in regions of very low albedo (e.g., the background and the iris

Figure 9. (a) One of 36 input images (18 reciprocal pairs), (b) the recovered depth map, and (c) a quiver plot of the recovered field of surface normals. As expected, even though we obtain a poor estimate of the depth due to lack of texture, the surface normals are accurately recovered. (Note that the image in (a) is taken from a position above the principal view used for reconstruction.)

Figure 10. As in the previous figure: (a) one of 34 input images (17 reciprocal pairs), (b) the recovered depth map, and (c) a quiver plot of the recovered field of surface normals. As evidenced by the specularities in (a), the surface is non-Lambertian. Regions of very small albedo (e.g., the iris of the eyes, the background) are sensitive to noise and erroneous results are expected there. Elsewhere, the depth and orientation are accurately recovered. A 9×9 window was used in the depth search.

Figure 11. A reconstruction for the marked interior region of a ceramic figurine shown in (a). (b) and (c) are the depth map, and normal field. The low resolution of the depth map is caused by the 11×11 window used in the depth search, but this does not affect the accuracy of the estimated surface normals. Eighteen reciprocal image pairs were used.

Figure 12. A reconstruction for the face of a plastic doll shown in (a). (b) and (c) are the estimated depth map and normal field. Eighteen reciprocal image pairs and a 9×9 window were used.

Figure 13. The surface that results from integrating the normal field shown in Fig. 9(c). Every third surface point is shown, and the surface is rotated for clarity.

of the eyes) since these regions are very sensitive to noise.

Figures 11 and 12 show two more examples of surface reconstructions. Again, note that the recovered surface normals are accurate despite the low resolution of the depth estimates, even in regions of nearly constant image brightness.

As mentioned at the end of the last section, it is possible to obtain a more precise surface reconstruction by integrating the estimated normal field. The examples above demonstrate that this field is accurately estimated, even in regions where the depth is not. To illustrate how surfaces can be reconstructed in this way, we enforced integrability (using the method of Frankot and Chellappa (1988) with a Fourier basis) and integrated the vector fields shown in Figs. 9(c) and 12(c). The results are shown in Figs. 13 and 14. As seen in these figures, the high resolution information provided by the surface normals enables the recovery of precise surface shape—more precise than what we would expect from most conventional n-view stereo methods. Note that it would be possible to obtain similar reconstructions using photometric stereopsis, but this would require an assumed model of reflectance at each point of the surfaces.

Figure 14. Three views of the surface that results from integrating the normal field shown in Fig. 12(c). To demonstrate the accuracy of the reconstruction, we have refrained from texture-mapping the albedo onto the recovered surface, and a real image taken from each corresponding viewpoint is displayed. The specularities on the doll's face clearly show that the surface is non-Lambertian.

6. Conclusion

This paper introduces Helmholtz stereopsis—a surface reconstruction method based on the principle of Helmholtz reciprocity. The method has two main advantages over both conventional n-view stereopsis and photometric stereopsis. First, it allows for the recovery of the shape of surfaces that have arbitrary, unknown, and possibly spatially varying BRDFs; and second, it provides point-wise estimates of both the depth and the surface normals.

This paper shows empirically that the reciprocity condition satisfied by the BRDF can be exploited for surface reconstruction, yet there are a number of possibilities for future work. The method could be adapted to locate depth discontinuities based on the correspondence between shadowed and half-occluded regions. This correspondence is a unique and powerful property of reciprocal image pairs. In addition, the analysis and construction of alternative imaging configurations may identify those that minimize ambiguities or are optimal from an estimation standpoint. While the results from the current implementation are promising, future improvements can only serve to produce more detailed and accurate reconstructions. We hope that this will open future avenues of research that exploit the reciprocity constraint.

Acknowledgments

The authors would like to thank Shree Nayar, Athos Georghiades, Melissa Koudelka, and Patrick Huggins

for helpful discussions regarding this work, as well as the reviewer for pointing out some relevant history regarding reciprocity. This work was supported by a Presidential Early Career Award IIS-9703134, NIH R01-EY 12691-01, NSF KDI-9980058, NSF ITR IIS-00-85864, and NSF CCR 00-86094.

References

Ashikhmin, M., Premoze, S., and Shirley, P. 2000. A microfacet-based BRDF generator. In *Proceedings of SIGGRAPH*, pp. 65–74.

Belhumeur, P. 1993. A binocular stereo algorithm for reconstructing sloping, creased, and broken surfaces in the presence of half-occlusion. In *Proc. Int. Conf. on Computer Vision*, pp. 431–438.

Belhumeur, P. 1996. A Bayesian approach to binocular stereopsis. *Int. Journal of Computer Vision*, 19(3):237–260.

Belhumeur, P. and Mumford, D. 1992. A Bayesian treatment of the stereo correspondence problem using half-occluded regions. In *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pp. 506–512.

Clarke, F. and Parry, D. 1985. Helmholtz reciprocity: Its validity and application to reflectometry. *Lighting Research and Technology*, 17(1):1–11.

Cook, R. and Torrance, K. 1981. A reflectance model for computer graphics. In *Proceedings of SIGGRAPH*, pp. 307–316.

Cox, I.J., Hingorani, S., Maggs, B.M., and Rao, S.B. 1992. Stereo without disparity gradient smoothing: A Bayesian sensor fusion solution. In *British Machine Vision Conference*, D. Hogg and R. Boyle (Eds.). Springer-Verlag: Berlin, pp. 337–346.

Deverney, F. and Faugeras, O. 1994. Computing differential properties of 3-D shapes from stereoscopic images without 3-D models. In *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pp. 208–213.

Faugeras, O. and Keriven, R. 1998. Variational principles, surface evolution, PDE's, level set methods, and the stereo problem. *IEEE Trans. Image Processing*, 7(3):336–344.

Frankot, R.T. and Chellappa, R. 1988. A method for enforcing integrability in shape from shading algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 10(4):439–451.

Geiger, D., Ladendorf, B., and Yuille, A. 1992. Occlusions and binocular stereo. In *Proc. European Conf. on Computer Vision*, pp. 425–433.

Hayakawa, K. 1994. Photometric stereo under a light source with arbitrary motion. *J. Opt. Soc. Am.* A, 11(11):3079–3089.

He, X.D., Heynen, P., Phillips, R., Torrance, K., Salesin, D., and Greenberg, D. 1992. A fast and accurate light reflection model. In *Proceedings of SIGGRAPH*, pp. 253–254.

Helmholtz, H.V. 1925. *Treatise on Physiological Optics*, vol. 1. Dover: New York.

Ikeuchi, K. and Horn, B. 1981. Numerical shape from shading and occluding boundaries. *Artificial Intelligence*, 17:141–184.

Koenderink, J. and van Doorn, A. 1996. Bidirectional reflection distribution function expressed in terms of surface scattering modes. In *Proc. European Conf. on Computer Vision*, pp. II:28–39.

Koenderink, J., van Doorn, A., Dana, K., and Nayar, S. 1999. Bidirectional reflection distribution function of thoroughly pitted surfaces. *Int. Journal of Computer Vision*, 31(2/3):129–144.

Langer, M. and Zucker, S.W. 1994. Shape-from-shading on a cloudy day. *J. Opt. Soc. Am.* A, 11(2):467–478.

Lu, J. and Little, J. 1999. Reflectance and shape from images using a collinear light source. *Int. Journal of Computer Vision*, 32(3):1–28.

Magda, S., Zickler, T., Kriegman, D., and Belhumeur, P., 2001. Beyond Lambert: Reconstructing surfaces with arbitrary BRDFs. In *Proc. Int. Conf. on Computer Vision*, pp. 391–398.

Matthies, L. 1992. Stereo vision for planetary rovers: Stochastic modeling to near real-time implementation. *Int. Journal of Computer Vision*, 8(1):71–91.

Minnaert, M. 1941. The reciprocity principle in lunar photometry. *Astrophysical Journal*, 93:403–410.

Nayar, S., Ikeuchi, K., and Kanade, T. 1990. Determining shape and reflectance of hybrid surfaces by photometric sampling. *IEEE J. of Robotics and Automation*, 6(4):418–431.

Nicodemus, F., Richmond, J., Hsia, J., Ginsberg, I., and Limperis, T. 1977. Geometric considerations and nomenclature for reflectance. Monograph 160, National Bureau of Standards (US).

Oren, M. and Nayar, S. 1996. Generalization of the Lambertian model and implications for machine vision. *Int. Journal of Computer Vision*, 14(3):227–251.

Phong, B. 1975. Illumination for computer-generated pictures. *Comm. of the ACM*, 18(6):311–317.

Poggio, T., Torre, V., and Koch, C. 1985. Computational vision and regularisation theory. *Nature*, 317:314–319.

Silver, W. 1980. Determining shape and reflectance using multiple images. Master's Thesis, MIT.

Snyder, W., Wan, Z., and Li, X. 1998. Thermodynamic constraints on reflectance reciprocity and Kirchoff's law. *Applied Optics*, 37(16): 3464–3470.

Tagare, H. and deFigueiredo, R. 1991. A theory of photometric stereo for a class of diffuse non-Lambertian surfaces. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 13(2):133–152.

Torrance, K. and Sparrow, E. 1967. Theory for off-specular reflection from roughened surfaces. *J. Opt. Soc. Am.*, 57(9):1105–1114.

Woodham, R. 1981. Analysing images of curved surfaces. *Artificial Intelligence*, 17:117–140.