Helmholtz Stereopsis: Exploiting Reciprocity for Surface Reconstruction⋆

Todd Zickler[1], Peter N. Belhumeur[2], and David J. Kriegman[3]

[1] Center for Comp. Vision & Control, Yale University, New Haven, CT 06520-8285
zickler@yale.edu
[2] Department of Computer Science, Columbia University, New York, NY 10027
belhumeur@cs.columbia.edu
[3] Beckman Institute, University of Illinois, Urbana-Champaign, Urbana, IL 61801
kriegman@uiuc.edu

Abstract. We present a method – termed Helmholtz stereopsis – for reconstructing the geometry of objects from a collection of images. Unlike most existing methods for surface reconstruction (e.g., stereo vision, structure from motion, photometric stereo), 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 field of surface normals, and consequently weds the advantages of both conventional and photometric stereopsis. Results from our implementations lend empirical support to our technique.

1 Introduction

In this paper, we present a method for reconstructing the geometry of a surface that has arbitrary and unknown surface reflectance, as described by the bidirectional reflectance distribution function (BRDF) [17]. This method does *not* make the assumption that the BRDF is Lambertian or of some other parametric form, and it enables the reconstruction of surfaces for which the BRDF is anisotropic and spatially varying. Helmholtz stereopsis works by exploiting the symmetry of surface reflectance – it chooses pairs of light source and camera positions that guarantee that the relationship between pixel values at corresponding image points depends only on the surface shape (and is independent of the BRDF).

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
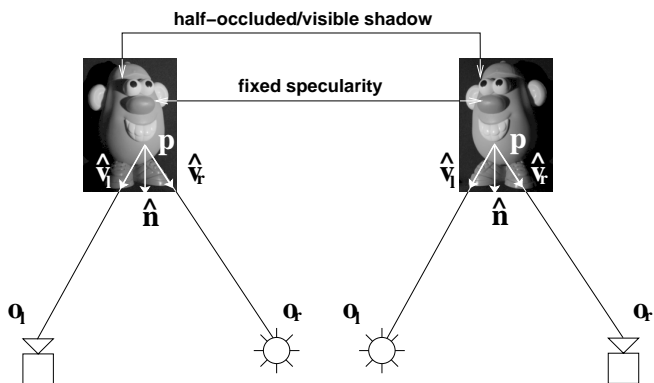

Fig. 1. The setup for acquiring a pair of images that exploits Helmholtz reciprocity. First an image is acquired with the scene illuminated by a single point as shown on the left. Then, a second image is acquired after the positions of the camera and light source are exchanged as shown on the right.

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, as 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 was first enunciated by Helmholtz ([10], p. 231) and is commonly referred to as Helmholtz reciprocity.

To see how 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. The second image is captured once the camera and light source have been exchanged. 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 mutually 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 BRDF of the surface is Lambertian.)

In [14], it was shown that three or more pairs of images acquired in this manner are needed to establish a matching constraint, which leads to a multinocular stereo imaging geometry. It was also shown that in addition to this constraint being useful for searching depth, it contains sufficient information to directly estimate the surface normal at each point without taking derivatives of either the images or the depth map. This is similar to conventional photometric stereopsis, but here the BRDF may be unknown and non-parametric.

The fact that the normal field can be estimated in addition to depth is quite significant in that it allows the surface to be reconstructed in regions of constant image brightness. For these places on the object's surface where traditional stere-

opsis fails to recover any information, Helmholtz stereopsis is able to recover the field of surface normals. This information can then be integrated – as it is in photometric stereopsis – to recover the surface shape.

We should point out that in [14], Helmholtz stereopsis was introduced as one of two methods for reconstructing surfaces with arbitrary BRDFs. In that paper, only a simple binocular demonstration of the matching constraint was shown, one which could be applied under a very restricted imaging situation (fronto-parallel surfaces and small baselines); in contrast, this paper provides a full multinocular description, analysis, and implementation of Helmholtz stereopsis.

In the next section, we describe the reciprocity-based method and follow in Sect. 4 with experimental results of our current implementation. Since the method combines the advantages of conventional multinocular stereopsis (estimation of depth) with those of photometric stereo (estimation of the normal field), we have summarized the similarities and differences in Sect. 3.

2   Helmholtz Stereopsis

In this section we describe our method for reconstructing surfaces with arbitrary BRDFs using a form of multinocular stereopsis. For a more detailed derivation, the reader is referred to [14]. As stated in the introduction, the method differs from standard stereopsis in that it is an *active* reconstruction technique – in addition to changing the camera position, the illumination of the scene is manipulated to exploit Helmholtz reciprocity.

Before describing Helmholtz stereopsis, it will be helpful to provide a framework for general N-view 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 can 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 is a one-parameter family of n-point sets $\{\mathbf{q}_1, \ldots, \mathbf{q}_n\}$ in the other images that could correspond to $\mathbf{q}$. We can 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 for 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 of the surface.

A single reciprocal pair of images is gathered by interchanging the positions of the light source and the camera's center of projection as shown in Fig. 1. As shown in [14], because of Helmholtz reciprocity, given such an image pair the image irradiance values measured at corresponding pixels in the left and right images satisfy the constraint

$$\left(i_\mathrm{l} \frac{\hat{\mathbf{v}}_\mathrm{l}}{|\mathbf{o}_\mathrm{l} - \mathbf{p}|^2} - i_\mathrm{r} \frac{\hat{\mathbf{v}}_\mathrm{r}}{|\mathbf{o}_\mathrm{r} - \mathbf{p}|^2} \right) \cdot \hat{\mathbf{n}} = \mathbf{w}(d) \cdot \hat{\mathbf{n}} = 0, \tag{1}$$

where, i_l and i_r are image irradiance measurements (obtained from a radiometrically calibrated camera), $\mathbf{o}_\mathrm{l}$ and $\mathbf{o}_\mathrm{r}$ are the left and right camera centers, $\mathbf{p}$ is a surface point, and $\hat{\mathbf{v}}_\mathrm{l}$ and $\hat{\mathbf{v}}_\mathrm{r}$ are unit vectors that point from $\mathbf{p}$ to the camera centers (see Fig. 1). For calibrated cameras and a hypothesized value for the depth d, values for these vectors and points can be computed (we write $\mathbf{w}(d)$ to denote this fact), and only the surface normal $\hat{\mathbf{n}}$ is 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).

If we capture N_P of these reciprocal pairs (each from a different pair of positions $\mathbf{o}_\mathrm{l}, \mathbf{o}_\mathrm{r}$), we will have N_P linear constraints of this form. Let $\mathbf{W}(d)$ be the $N_\mathrm{P} \times 3$ matrix where the ith row is given by $\mathbf{w}_i(d) = i_{\mathrm{l}i} \frac{\hat{\mathbf{v}}_{\mathrm{l}i}}{|\mathbf{o}_{\mathrm{l}i} - \mathbf{p}|^2} - i_{\mathrm{r}i} \frac{\hat{\mathbf{v}}_{\mathrm{r}i}}{|\mathbf{o}_{\mathrm{r}i} - \mathbf{p}|^2}$. Then the set of constraints from (1) can be expressed as

$$\mathbf{W}(d)\,\hat{\mathbf{n}} = 0. \tag{2}$$

Clearly, for the 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 the depth. Note that at least three camera/light source positions are needed to exploit this constraint. An implementation of a system that uses this constraint for surface reconstruction will be presented in Sect. 4. Before that, we present a comparison of this method with some existing reconstruction techniques.

3   Comparison with Existing Methods

In principle, Helmholtz Stereopsis has a number of advantages when compared to conventional multinocular stereopsis and photometric stereopsis. Figure 2 summarizes these. 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.

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

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

Fig. 2. A comparison of Helmholtz stereopsis with conventional dense multinocular and photometric stereopsis. A more detailed discussion of the subtleties of the entries in this table is given in Sect. 3.

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 arbitrary, 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. These methods provide an estimate of the field of surface normals which is then integrated to recover the surface depth. Similar to conventional dense multinocular stereopsis, many photometric methods assume that the BRDF is Lambertian [12, 19,21]. The methods that do not make this assumption either assume that the BRDF (or the reflectance map) is completely known a priori, or can be specified using a small number of parameters [9,11,16,20]. These parametric BRDFs are often derived from physical models of reflectance and are therefore 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.

In [13], a hybrid method with controlled lighting and object rotation 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).

The assumptions made about surface reflectance for three reconstruction techniques – conventional, photometric, and Helmholtz stereopsis – are summarized diagrammatically in Fig. 3. Note that many natural surfaces actually have
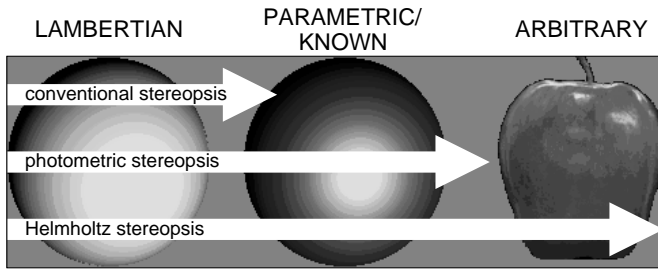
Fig. 3. A summary of the assumptions made about surface reflectance by three reconstruction techniques. Both conventional dense multinocular stereopsis and photometric stereopsis assume the BRDF is either Lambertian or of some other parametric form. Yet many natural surfaces (e.g. human skin, the skin of a fruit, glossy paint) do not satisfy these assumptions. In contrast to the other methods, Helmholtz stereopsis makes no assumption about the BRDF.

surface reflectance in the rightmost region of the figure and are thus excluded from reconstruction by conventional techniques.

In Helmholtz stereopsis, because the relationship between the 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 (fixed illumination) stereo images, these radiometric events become reliable features, and they actually simplify the matching problem.

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 [5], or through the use of an energy functional [1]. 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 at each point based on 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
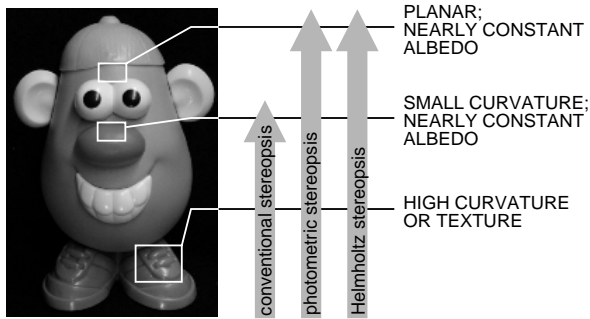
Fig. 4. A summary of the surface properties required for Lambertian surface reconstruction by conventional, photometric and Helmholtz stereo techniques. Even when the BRDF is Lambertian, conventional stereopsis is only capable of recovering surface geometry in regions of texture (i.e. varying albedo) or high curvature (i.e. edges). Neither photometric stereopsis nor Helmholtz stereopsis suffer from this limitation.

image-based rendering and image-based modeling, a good estimate of the normal field is critical for computing intensities and accurately measuring reflectance properties.

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 [8,15,18,2] or through regularized surface evolution [6].) 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. 4.

Depth Discontinuities

Depth discontinuities present difficulties for both traditional and photometric stereopsis. As mentioned above, when there is a depth discontinuity, it does not make sense to integrate the normal field that is output by photometric stereo techniques; photometric stereopsis cannot resolve depth discontinuities at the boundaries of objects. 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. 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 the lighting setup is such that that the shadowed and half-occluded regions are in correspondence. To see this, consider the reciprocal pair shown in Fig. 1. When a surface point is in shadow in the left image, it is not visible in the right image, and vice versa. The shadowed regions in the images of a Helmholtz pair can therefore be used to locate depth discontinuities. If one uses a stereo matching algorithm that exploits the presence of half-occluded regions for determining depth discontinuities [1,3,4,8], then these shadowed regions may significantly enhance the quality of the depth reconstruction.

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.

4   Implementation and Results

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

- reconstructions of surfaces with arbitrary BRDFs (surfaces that are not Lambertian or approximately Lambertian)
- the recovery of both surface depth and the normal field
- reconstructions of surfaces in regions of constant image brightness

Capturing Reciprocal Images

To evaluate the use of Helmholtz reciprocity we have implemented a system that enables the acquisition of multiple reciprocal image pairs with a single camera and light source. These are mounted on a wheel as shown schematically in Fig. 5a. 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, the two images will form a reciprocal pair, and corresponding image irradiance values will satisfy the constraint in (1). It is clear that we can capture any number of these pairs by rotating the wheel through $360°$ while stopping to capture images at reciprocal positions.
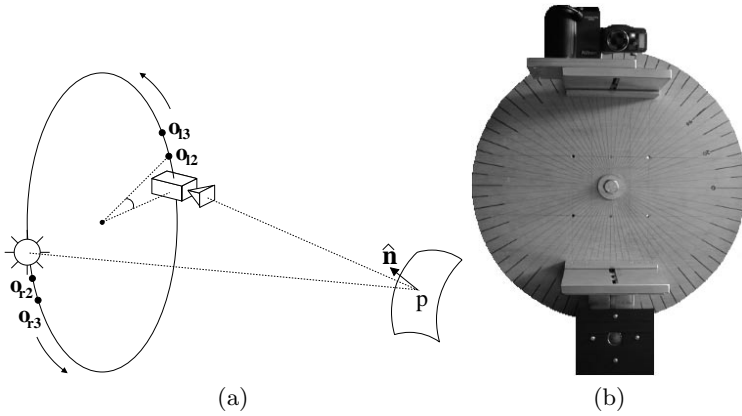
(a)   (b)

Fig. 5. (a) A wheel is used to capture multiple reciprocal image pairs employing a single camera and light source. By rotating the wheel through $360\,^{\circ}$, 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 100W frosted incandescent bulb fitted with a small aperture.

A picture of such a system is shown in Fig. 5b. The camera is a Nikon Coolpix 990, and the light source consists of a standard 100W 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. 6. For all results shown in this paper the diameter of the wheel was 38cm and the distance from the center of the wheel to the scene was approximately 60cm. 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.

Using the Matching Constraint

In Sect. 2 we described a matrix constraint equation that can be used to recover the surface depth and orientation corresponding to each point $\mathbf{q}$ in the principal view. How this constraint is 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,
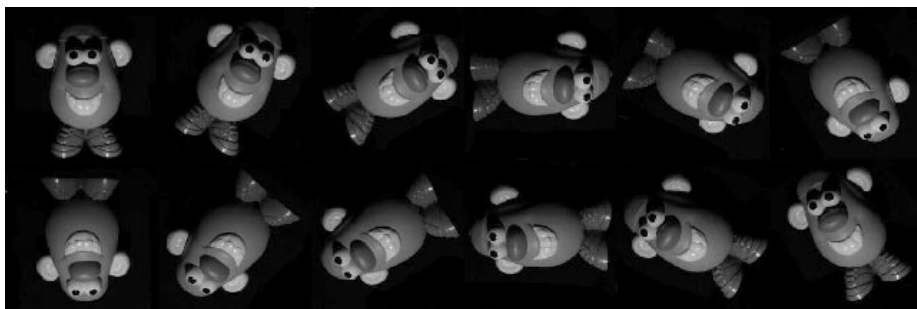
Fig. 6. An example of 6 reciprocal images pairs captured using our rig. Image pairs are arranged vertically.

we have chosen one particularly simple implementation which will be described here. Results for four different surfaces follow in the next section.

For each point $\mathbf{q}$, and for each depth value $d \in D = \{d_1, d_2, \ldots, d_{N_D}\}$ we can construct a matrix $\mathbf{W_q}(d)$ as in (2). 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 $\mathbf{W}$ lies in the epipolar plane defined by $\mathbf{p}$, $\mathbf{o}_{li}$, and $\mathbf{o}_{ri}$, no two rows of $\mathbf{W}$ will be collinear, so rank$(\mathbf{W}) \geq 2$.)

In the presence of noise, $\mathbf{W}$ will generally be rank three, and we require a measure for the coplanarity of the row vectors $\mathbf{w}_i$. Since we know that rank$(\mathbf{W}) \geq 2$, a suitable measure (and one that works well in practice) is the ratio of the second to third singular values of $\mathbf{W}$. Given each matrix $\mathbf{W_q}(d)$, we compute the singular value decomposition, $\mathbf{W} = \mathbf{U\Sigma V}^\mathsf{T}$ where $\mathbf{\Sigma} = \mathrm{diag}(\sigma_1, \sigma_2, \sigma_3), \sigma_1 \geq \sigma_2 \geq \sigma_3$. Then, our support measure to be used in the depth search is the ratio

$$r_\mathbf{q}(d) = \frac{\sigma_2}{\sigma_3}. \tag{3}$$

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

The condition shown in (2) 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 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

$$d_{\mathbf{q}_\circ}^\star = \arg\max_{d \in D} \sum_{\mathbf{q} \in W} r_\mathbf{q}(d) \tag{4}$$

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

$$\hat{\mathbf{n}}_{\mathbf{q}_\circ}^\star = \arg\min_{\hat{\mathbf{n}}} \|\mathbf{W}_{\mathbf{q}_\circ}(d^\star)\hat{\mathbf{n}}\|^2, \|\hat{\mathbf{n}}\| = 1, \tag{5}$$

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 (4) 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 to detect 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 its simplicity, the surface reconstructions are quite good.

Results

Figures 7-10 show the results of applying this procedure. Each figure consists of: (a) one of the input images of the object, (b) the estimated depth, and (c) the recovered field of surface normals. Note that the viewpoints of the displayed input images are slightly different from the reconstruction viewpoints due to the use of a virtual principal camera.

Figure 7 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. 4, 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 matrix $\mathbf{W}_{\mathbf{q}}(d)$ will have approximately the same null space.

Figure 8 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 as well as the iris of the eyes) since these regions are very sensitive to noise.

Figures 9 and 10 show two more examples of surface reconstructions. Again, note that the recovered field of surface normals is 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 the normal field is accurately estimated, even in regions where the depth is not. To illustrate how surfaces can be reconstructed
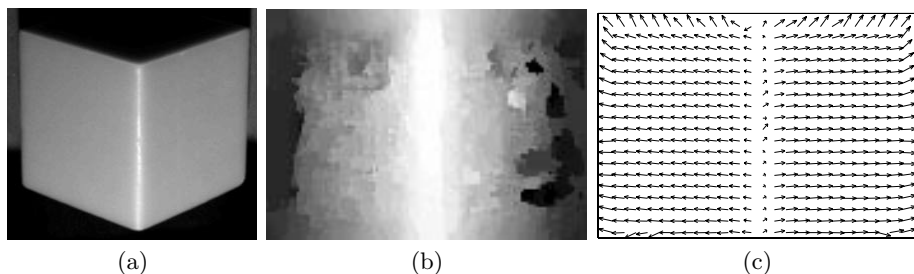
<p align="center">(a)   (b)   (c)</p>

Fig. 7. This figure shows: (a) one of 36 input images (18 reciprocal pairs), (b) the recovered depth map, and (c) a quiver plot of the recovered normal field. As expected, even though we obtain a poor direct estimate of the depth, the surface normals are accurately recovered. Note that the image in (a) is taken from a position above the principal view used for reconstruction.
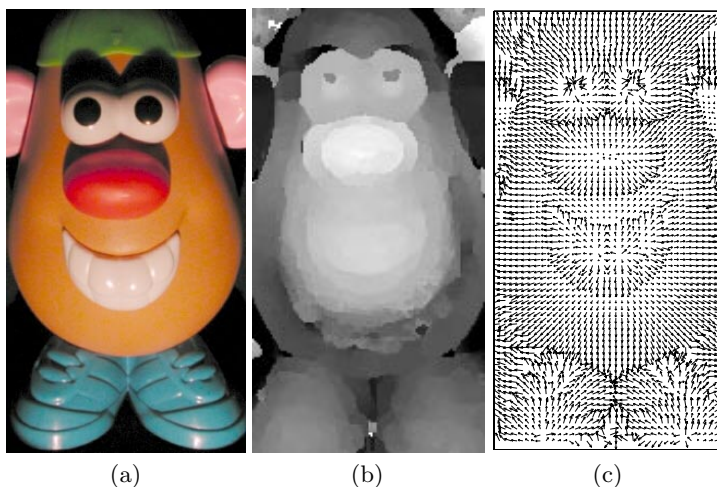

<p align="center">(a)   (b)   (c)</p>

Fig. 8. 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 normal field. 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.

in this way, we enforced integrability (using the method of Frankot and Chellapa [7] with a Fourier basis) and integrated the the vector fields shown in Figs. 7c and 10c. The results are shown in Figs. 11 and 12. As seen in these figures, the high frequency information provided by the surface normals enables the recovery of precise surfaces shape – more precise than what we could expect from most conventional n-view stereo methods. Note that it would be possible to obtain a similar reconstruction using photometric stereo, but this would require an assumed model for the reflectance at each point on the surface.
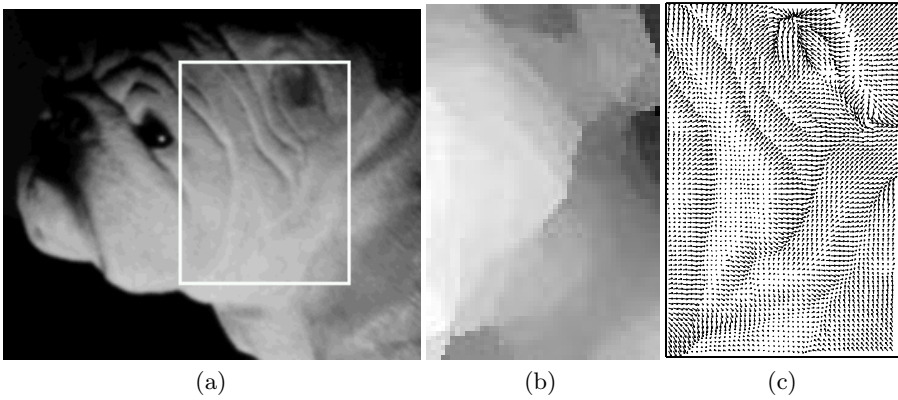
Fig. 9. A reconstruction of the marked interior region (a) of a ceramic figurine. Figures (b) and (c) are the depth map, and the quiver plot of the normal field. The low resolution of the depth map is caused by the 11×11 window used in the depth search, but this does not affect the accuracy of the estimated normal field. Again, 18 reciprocal image pairs were used.
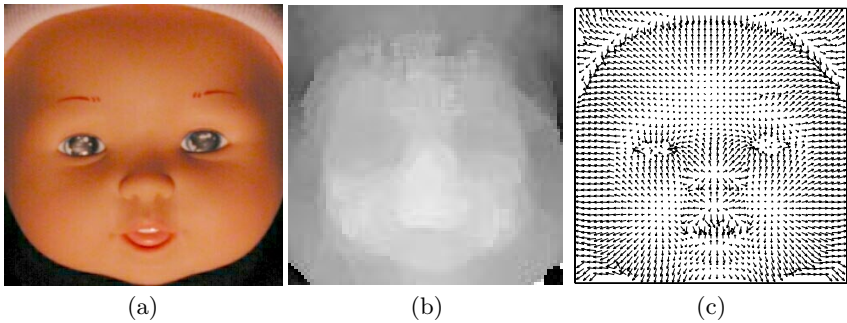

Fig. 10. A reconstruction for the face of a plastic doll (a). Figures (b) and (c) are the depth map, and surface normals. 18 reciprocal image pairs and a 9×9 window were used.
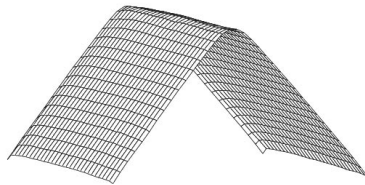

Fig. 11. The surface that results from integrating the normal field shown in Fig. 7c. A vertex for every third surface point is shown, and the surface is rotated for clarity.
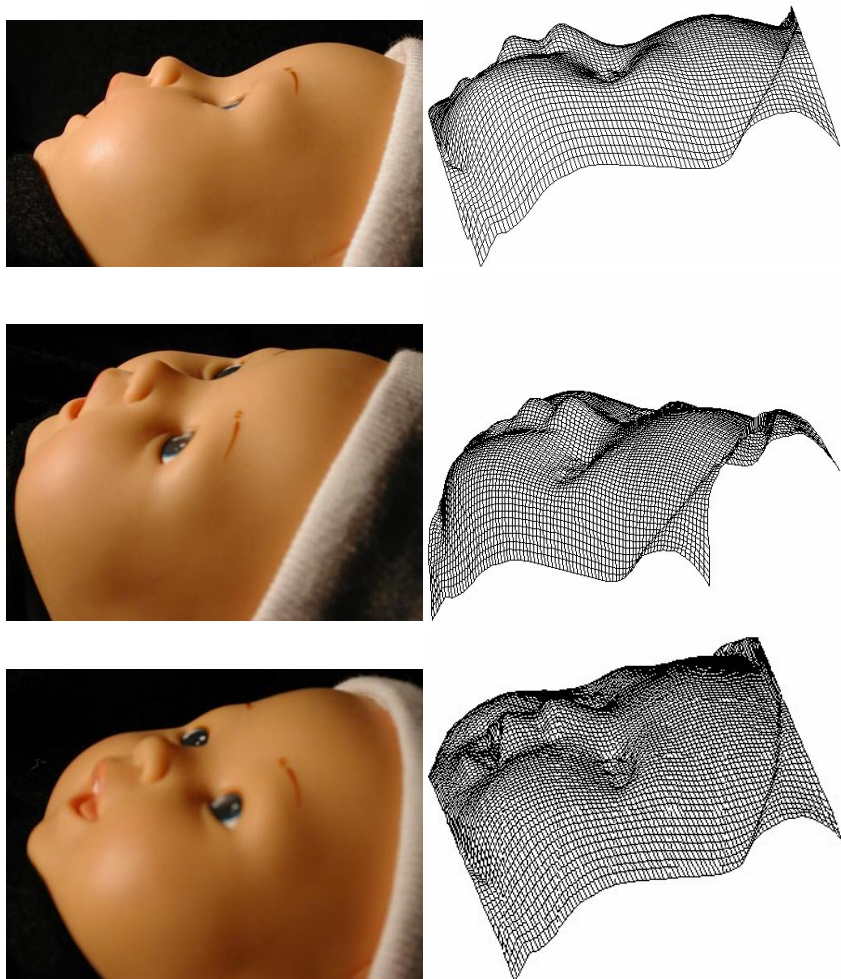
Fig. 12. Three views of the surface that results from integrating the normal field shown in Fig. 10c. To demonstrate the accuracy of the reconstruction, we have refrained from texture mapping the albedo onto the recovered surface, and a real image taken from each corresponding viewpoint is displayed. The specularities on the doll's face clearly show that the surface is non-Lambertian.

5   Discussion

This paper presents a surface reconstruction method that does not make any assumptions about the BRDF of the observed surface. It combines the advantages of both conventional N-view stereopsis and photometric stereopsis in that it di-

rectly estimates both the depth and the field of surface normals of an object. In contrast to these methods, however, it can recover this geometric information for surfaces that have arbitrary, unknown, and possibly spatially varying BRDFs. The method is termed Helmholtz stereopsis, as it works by exploiting the physical principal known as Helmholtz reciprocity.

Helmholtz stereopsis is a form of active multinocular stereo, and as such, it requires that multiple images be captured under controlled illumination. This paper presented an implementation of a simple wheel design that is capable of gathering these images in a controlled manner with a single camera and a simple approximation to a point light source. The results from this implementation demonstrate its ability to recover surface shape.

The goal of this paper was to show empirically that the reciprocity condition satisfied by the BRDF can be exploited for surface reconstruction; there are a number of possibilities for future work. The rig that was used here was manually rotated and required a full geometric calibration a priori. One could imagine an automated system with a servo motor, a video camera replacing the digital still camera, and a self-calibration routine.

In addition, although the reciprocal image pairs described in the paper contain information that can be used for locating depth discontinuities, this information was not explicitly used in our implementation. The correspondence between shadowed and half-occluded regions is nevertheless a powerful source of information, one that we plan to exploit in the future.

Our model should be adapted to account for surface interreflections, and one could explore additional configurations for capturing Helmholtz pairs and a more direct method of combining depth and surface orientation information.

References

1. P. Belhumeur. A binocular stereo algorithm for reconstructing sloping, creased, and broken surfaces in the presence of half-occlusion. In *Int. Conf. on Computer Vision*, Berlin, Germany, 1993.
2. P. Belhumeur. A Bayesian approach to binocular stereopsis. *Int. J. Computer Vision*, 19, 1996.
3. P. Belhumeur and D. Mumford. A Bayesian treatment of the stereo correspondence problem using half-occluded regions. In *Proc. IEEE Conf. on Comp. Vision and Patt. Recog.*, pages 506–512, 1992.
4. I. J. Cox, S. Hingorani, B. M. Maggs, and S. B. Rao. Stereo without disparity gradient smoothing: a Bayesian sensor fusion solution. In *British Machine Vision Conference*, pages 337–346, 1992.
5. F. Deverney and O. Faugeras. Computing differential properties of 3-D shapes from stereoscopic images without 3-D models. In *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pages 208–213, 1994.
6. O. Faugeras and R. Keriven. Variational principles, surface evolution, PDE's, level set methods, and the stereo problem. *IEEE Trans. Image Processing*, 7(3):336–344, 1998.
7. R. T. Frankot and R. Chellappa. A method for enforcing integrability in shape from shading algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-10(4):439–451, 1988.

8. D. Geiger, B. Ladendorf, and A. Yuille. Occlusions in binocular stereo. In *Proc. European Conf. on Computer Vision*, Santa Margherita Ligure, Italy, 1992.

9. H. Hayakawa. Photometric stereo under a light-source with arbitrary motion. *J. Optical Society of America A*, 11(11):3079–3089, Nov. 1994.

10. H. v. Helmholtz. *Treatise on Physiological Optics*. Dover (New York), 1925.

11. 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.

12. M. Langer and S. W. Zucker. Shape-from-shading on a cloudy day. *J. Opt Soc. Am.*, pages 467–478, 1994.

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

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

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

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

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

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

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

20. 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.

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