Shape Estimation Using Polarization and Shading from Two Views

Gary A. Atkinson, *Student Member*, *IEEE*, and Edwin R. Hancock

Abstract—This paper presents a novel method for 3D surface reconstruction that uses polarization and shading information from two views. The method relies on the polarization data acquired using a standard digital camera and a linear polarizer. Fresnel theory is used to process the raw images and to obtain initial estimates of surface normals, assuming that the reflection type is diffuse. Based on this idea, the paper presents two novel contributions to the problem of surface reconstruction. The first is a technique to enhance the surface normal estimates by incorporating shading information into the method. This is done using robust statistics to estimate how the measured pixel brightnesses depend on the surface orientation. This gives an estimate of the object material reflectance function, which is used to refine the estimates of the surface normals. The second contribution is to use the refined estimates to establish correspondence between two views of an object. To do this, surface patches are extracted from each view, which are then aligned by minimising an energy functional based on the surface normal estimates and local topographic properties. The optimum alignment parameters for different patch pairs are then used to establish stereo correspondence. This process results in an unambiguous field of surface normals, which can be integrated to recover the surface depth. Our technique is most suited to smooth nonmetallic surfaces. It complements existing stereo algorithms since it does not require salient surface features to obtain correspondences. An extensive set of experiments, yielding reconstructed objects and reflectance functions, are presented and compared to ground truth.

Index Terms—Polarization imaging, surface shape recovery, stereo, reflectance function estimation, patch alignment.

✦

1 INTRODUCTION

THE problem of recovering the 3D shape of an object from one or more views is a central topic in computer vision and is given the generic term *shape-from-X*. Perhaps the best-known single view technique is *shape-from-shading* (SFS), where variations in image brightness are used to estimate the field of surface normal directions and, ultimately, the 3D surface geometry. A multi-image extension of SFS is *photometric stereo*, where several images are used with a fixed camera position but varying light source directions.

In this paper, we present a novel two-view technique for surface reconstruction based on polarization analysis. The technique uses Fresnel theory to recover the surface normals under conditions of diffuse reflection and establishes the necessary two-view correspondence using surface topography information. We use polarization measurements obtained from a standard digital camera and a linear polarizing filter to accomplish three tasks. First, we estimate the reflectance properties of the material by statistically determining the relationship between the surface orientation (as estimated using polarization) and the pixel brightness. Second, we establish correspondence between the data from the two views using a novel surface-matching algorithm. Finally, the estimated surface normals are used for shape reconstruction.

• *The authors are with the Department of Computer Science, University of York, York YO10 5DD, UK. E- mail: {atkinson, erh}@cs.york.ac.uk.*

1.1 Related Work

1.1.1 Shape-from-Shading and Stereo

One of the most extensively studied surface recovery methods is SFS [1]. The foundation of the technique is the relationship between the image pixel brightness and the local surface orientation. SFS aims to estimate the surface normal for each pixel using a single gray-scale image, subject to boundary and smoothness constraints. The height of the surface can then be determined from the resulting field of surface normals (the needle map) using surface integration methods [2], [3].

Despite many attempts to apply it to real-world problems, SFS has found only few successful applications. One reason for this is that, with unknown light sources, an inherent ambiguity known as *bas-relief* [4] is present, which allows differently shaped objects to produce the same intensity image. Another reason for the lack of success is that the relationship between the incident light source direction and the distribution of reflected light is generally complicated and unknown. This relationship, which is defined quantitatively later on in Section 3, is termed the *reflectance function* or *bidirectional reflectance distribution function (BRDF)*. The traditional approach adopted in SFS is to assume that *Lambert's law* applies. This means that the observed image brightness simply depends on the cosine of the angle between the surface normal and the light source direction. It has long been known, however, that Lambert's law only works well for matte surfaces and breaks down for rough or shiny surfaces [5].

Treuille et al. [6] overcome the problem of unknown BRDF by placing spherical reference objects in the scene. Surface orientation is then estimated by applying the *orientation-consistency* assumption. That is, the surface orientations at

points on the unknown object are deduced from those points on the reference object that have identical intensities. Robles-Kelly and Hancock [7] estimate the BRDF from a single image without reference objects by calculating a mapping from a single image onto a Gauss sphere using the cumulative distribution of intensity gradients. A similar orientation-consistency assumption to that used by Treuille et al. is then applied. Ragheb and Hancock [8] use theoretical reflectance models [5] to "correct" the images, so that they appear to be Lambertian. Conventional SFS algorithms are then applied to determine the 3D geometry.

In order to overcome the underconstrained nature of a single-view vision, many researchers have turned to multiple-view techniques [9], [10]. If a point in an image from one view is known to correspond to a point in an image from a second view, then triangulation can be used to determine the distance to the camera. The correspondence problem, however, is difficult and is a major hurdle in stereo vision. Jin et al. [11] use shading information from multiple views to formulate stereo vision in terms of region correspondence. Many other techniques rely on matching salient image features [12].

In *Photometric stereo* [13], a fixed camera is used with varying but known light source directions. In certain situations, photometric stereo is possible using three images. The case where the surface is known to be Lambertian is an example of this. However, at least four images are necessary if more complicated reflectance functions are present. Also, more than three images are generally needed to deal with shadows and interreflections and to estimate the albedo. Zickler et al. [14] have used a technique called *Helmholtz stereopsis*, where the light source and camera are interchanged. This has the advantage that only two images are needed but requires more precisely controlled experimental conditions.

1.1.2 Polarization Methods

An alternative method for shape analysis relies on the principle that light is partially polarized as a result of surface reflection [15, Section 8.6]. The angle of polarization provides an ambiguous estimate of the surface normal azimuth angle (the angle of the projection of the surface normal onto the image plane), whereas the degree of the polarization constrains the zenith angle (the angle between the surface normal and the viewing direction) [16]. The polarization process is a result of interactions between the incident electromagnetic waves and the directional electron charge density in the medium. An important advantage of *polarization vision* over SFS is that the azimuth angle is easily recovered up to an ambiguity of 180 degrees. One must know, however, whether the reflection is a direct mirror-like reflection (specular reflection) or a result of subsurface interactions (diffuse reflection). If the reflection type is not known, then there are four possible angles separated by 90 degrees.

A technique that has become a standard in polarization vision is to take digital images of an object using a digital camera mounted with a linear polarizer rotated to different orientations. This was used, for example, by Miyazaki et al. [17] to recover the 3D geometry of transparent objects using specular reflection—a class of material that makes SFS impossible. In an earlier paper [18], the emission of an infrared light was also used. Wolff [19] made the first attempt to combine the polarization data from more than one view, with the aim to estimate the orientation of a

plane. On the other hand, Miyazaki et al. [20] and Drbohlav and Šára [21] apply the theory of polarization to recover shape using *diffuse* reflection. Rahmann and Canterakis [22], [23] attempt to account for both reflection types.

In addition to shape reconstruction, polarization has found applications in several other areas of computer vision too. Wolff and Boult [16], Nayar et al. [24], and Umeyama [25] have devised ways to separate diffuse from specular reflection. Wolff and Boult [16] experimented with segmenting images according to dielectric/metallic reflectance properties. Drbohlav and Šára [26] have enhanced photometric stereo by the addition of constraints from polarization. Clark et al. [27] demonstrate the usefulness of polarization in triangulation-based range scanning, where it is used to distinguish true laser stripes from interreflections. Schechner et al. have demonstrated how polarization can enhance images taken in poor viewing conditions such as haze [28] or underwater [29]. Finally, Shibata et al. [30] use polarization and a range scanner to recover reflectance functions.

The main disadvantage of polarization vision is the experimental difficulty in acquiring the required data. The standard method, mentioned above, requires a minimum of three images from each view, each with a different polarizer angle. This is a time-consuming process and limits the possible applications. It is, however, generally less of a practical hindrance than the steps needed to acquire the images used in a photometric stereo and Helmholtz stereopsis, since fewer illumination conditions or views are normally required for polarization methods. Wolff [31] and others have improved matters a little by developing *polarization cameras*. These devices use liquid crystals to rapidly switch the axis of the polarizing filter. The disadvantage here is that the data has a greater susceptibility to noise. Miyazaki et al. [32] improved the polarization camera by using PLZT (from lead (Pb) lanthanum (La) zirconium (Zr) titanium (Ti)) to switch the polarization state of reflected light.

1.2   Contribution

The main aim of this paper is to develop a new method for stereo shape recovery for featureless surfaces. Examples of such surfaces include shiny porcelain, plastics, and painted surfaces. Our motivation is that we cannot rely on traditional stereo techniques that depend on salient surface features for correspondence. We therefore propose a novel method for establishing correspondence that does not rely on feature matching. To accomplish this task, we use polarization to analyze the topographical surface structure and use this information to perform patch matching in order to establish correspondence. To ensure that the results are accurate, we incorporate shading information into the algorithm by means of an estimated reflectance function. This is used to improve the accuracy of the surface normal estimates. Fig. 1 shows an overview of the method diagrammatically and should be referred to throughout the paper.

The proposed method first makes ambiguous estimates of the surface normals independently from each of the two views assuming that light is diffusely reflected by the object. This is done by taking images of the object with a linear polarizer mounted on the camera lens and rotated to three different orientations. The normal estimates are then enhanced by incorporating shading information. We do this by robustly estimating the relationship between the surface zenith angles and the pixel brightnesses using global

Use polarization to estimate the zenith angles from (5) and the two possible azimuth angles from the polarization angle.

Refine zenith angle estimates using the intensity and the reflectance function

Correspondence to be established between resulting patches

| View 1 estimates | View 2 estimates |

Reflectance function estimation [§3.3]

Enhanced data | Enhanced data

Segment into patches | Segment into patches [§4.1]

Calculate alignment parameters and matching costs for patch pairs using surface normals and topography [§4.2-3]

Acquire correspondence from patch pairs with lowest cost [§4.3]

Use interpolation to estimate remaining correspondences [§4.4]

Integrate unambiguous surface normals to recover depth [Frankot and Chellappa 1988]

Depth

Fig. 1. Flow diagram of the proposed shape recovery method. The stages of the algorithm in the large box are shown in more detail in Fig. 15.

statistics. This is done because the initial estimates of the surface normals are heavily contaminated by noise and interreflections. For this paper, we concentrate on the case where the light source and the camera directions are identical so that the brightness is independent of the azimuth angle. We also assume that the reflectance is isotropic.

The enhanced, but still ambiguous, surface normals are then fully constrained by establishing correspondence between the two views using patch matching to optimize a cost function. The cost function for this matching process is based on both the normal directions and the local surface topography. The latter property is characterized by the view invariant quantity referred to as the *shape index* [33]. Surface normals are then integrated to obtain depth using the Frankot-Chellappa method [2]. In principle, triangulation is more accurate, but integration is less sensitive to errors, and our contribution is primarily about surface *normal* estimation.

Our method differs from previous shape recovery techniques in the following respects: Most importantly, our method is the first polarization-based technique to apply a global reflectance function estimation algorithm to incorporate shading information into the recovery process. Miyazaki et al. [17] use two views and estimate surface orientation from specular reflections. This means that the object had to be placed inside a "cocoon" with several external light sources so that specular reflections occur over the entire object. We also aim to recover denser correspondence here. In [20], Miyazaki et al. are able to use only a single view and therefore have the advantage over our method that less specific measurement conditions are needed. However, they only partially solve the azimuth angle ambiguity, making restrictive assumptions about the object's shape histogram. The methods devised by Rahmann and Canterakis [22], [23] use only the azimuth angles to establish correspondence and allow for both specular and diffuse reflection. This potentially results in less noisy reconstructions but at the cost of discarding large amounts of information.

The main weakness of our contribution is that we currently require a somewhat restrictive arrangement of the equipment. Most notably, we require that the object is rotated. In future work, we aim to reduce this restriction by moving the camera (or using two cameras simultaneously) but keeping the light source and object fixed. This would be more difficult since the reflectance function would be a function of both azimuth and zenith angles for at least one of the views.

By contrast to photometric stereo, our method has the following advantages: First, photometric stereo requires three or more images with different light source directions and constant camera positions, whereas our method needs only two views with the light source and camera in the same direction. Second, we do not rely on an assumed reflectance function but, instead, estimate it using global image statistics. Finally, because photometric stereo uses a fixed camera position, there is no obvious way to perform 360 degrees object reconstruction. Our method by contrast, involves a patch-matching procedure, which can be used to combine surfaces from any number of views.

The remainder of the paper is organized as follows: Section 2 introduces the standard Fresnel reflectance theory and explains how this leads to a means to constrain surface normals. Section 3 describes the proposed reflectance function estimation technique and presents results for objects of different shapes and materials. Section 4 presents the multiview polarization-based reconstruction method. Real-world examples of reconstructed shapes are provided in Section 5. Section 6 concludes the paper.

2 FRESNEL THEORY AND POLARIZATION VISION

In this section, we present an overview of the standard background theory necessary to understand our novel method. The work is primarily based on the Fresnel theory, which is used to describe how light is polarized when transmitted or reflected from an interface between two media with different refractive indices. We then show how this has been used in computer vision and define some of the quantities used in the remainder of the paper.
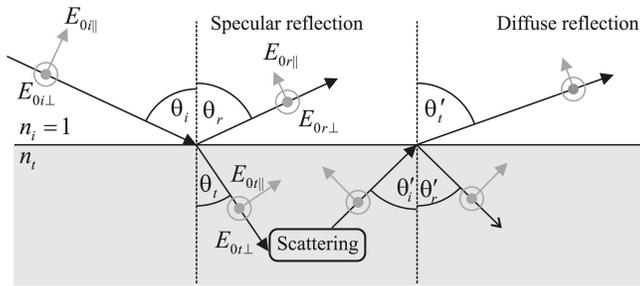
Fig. 2. Reflection of a ray of light by a medium. Definitions of angles and electric fields are shown.
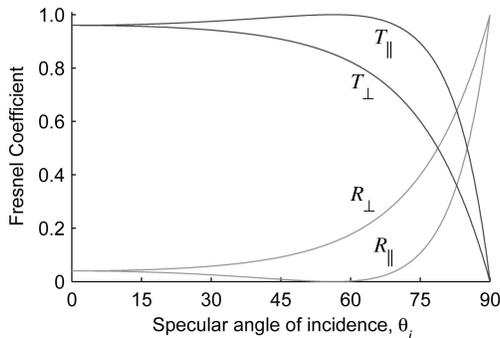
Fig. 3. Reflection and transmission coefficients for a dielectric ($n = 1.5$).

2.1 Fresnel Coefficients

Consider the specular reflection of a ray of light from a surface point, as shown in Fig. 2. Assume that the surface is a smooth interface between a dielectric and air. The Fresnel equations [15, Section 4.6] give the ratios of the reflected wave amplitude to the incident wave amplitude. They can be applied to incident light that is linearly polarized perpendicular to or parallel to the plane of incidence (the plane of the paper in Fig. 2). The ratios depend upon the angle of incidence θ_i and the refractive index n_t of the reflecting medium. We assume that the object is placed in air ($n_i = 1$). Since the incident light can always be resolved into two perpendicular components, the Fresnel equations are applicable to all incident polarization states. For the work reported here, we use unpolarized incident light.

For the geometry shown in Fig. 2, the Fresnel reflection coefficients are

$$r_\perp(n_i, n_t, \theta_i) \equiv \frac{E_{0r\perp}}{E_{0i\perp}} = \frac{n_i \cos \theta_i - n_t \cos \theta_t}{n_i \cos \theta_i + n_t \cos \theta_t}, \qquad (1)$$

$$r_\parallel(n_i, n_t, \theta_i) \equiv \frac{E_{0r\parallel}}{E_{0i\parallel}} = \frac{n_t \cos \theta_i - n_i \cos \theta_t}{n_t \cos \theta_i + n_i \cos \theta_t}. \qquad (2)$$

Equation (1) gives the reflection ratio for light polarized perpendicular to the plane of incidence, and (2) is for light polarized parallel to the plane of incidence. The angle θ_t can be obtained from the well-known Snell's law: $n_i \sin \theta_i = n_t \sin \theta_t$. Cameras do not measure the amplitude of a wave but the square of the amplitude or *intensity*. With this in mind, it can be shown that the *intensity coefficients*, which relate the reflected power to the incident power, are $R_\perp = r_\perp^2$ and $R_\parallel = r_\parallel^2$ [15, Section 4.6].
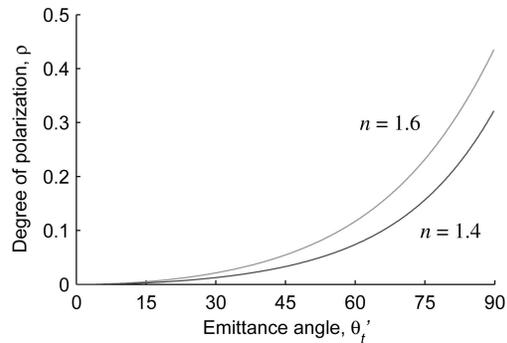
Fig. 4. Degree of polarization for diffuse reflection for two different refractive indices.

Fig. 3 shows the Fresnel intensity coefficients for a typical dielectric as a function of the angle of the incident light. Both reflection and transmission coefficients are shown, where the latter refers to the ratio of transmitted to incident power (the transmission coefficients are simply $T_\perp = 1 - R_\perp$ and $T_\parallel = 1 - R_\parallel$).

2.2 Polarization Analysis

The work reported here relies on taking a succession of images of objects with a polarizer mounted on the camera rotated to different angles. The measured pixel brightness at a given point varies with polarizer angle θ_{pol} according to the *Transmitted Radiance Sinusoid* (TRS)

$$I(\theta_{pol}) = \frac{I_{max} + I_{min}}{2} + \frac{I_{max} - I_{min}}{2} \cos(2\theta_{pol} - 2\phi), \qquad (3)$$

where I_{max} and I_{min} are the maximum and minimum intensities in the sinusoid, respectively, and ϕ is the angle of polarization of the light or the *phase* angle.

The *degree of polarization* is defined to be

$$\rho = \frac{I_{max} - I_{min}}{I_{max} + I_{min}}. \qquad (4)$$

It is possible to combine (4) and the Fresnel theory to obtain an expression for the degree of polarization in terms of the refractive index and the zenith angle [16]. The expression, however, is only applicable to specular reflection since the process that causes polarization by diffuse reflection is different, as explained below.

During diffuse reflection [16], a portion of the incident light penetrates the surface and is scattered internally. Some of the light is then refracted back into air and is partially polarized in the process. Fig. 2 illustrates the stages of a diffuse reflection schematically. Snell's law and the Fresnel *transmission* coefficients (shown in Fig. 3) can be combined to derive the following equation for the degree of polarization:

$$\rho(n, \theta) = \frac{(n - 1/n)^2 \sin^2 \theta}{2 + 2n^2 - (n + 1/n)^2 \sin^2 \theta + 4 \cos \theta \sqrt{n^2 - \sin^2 \theta}}. \qquad (5)$$

Here, and for the remainder of the paper, we call the refractive index of the reflecting medium n and use the typical value of $n = 1.4$. The dependence of the degree of polarization ρ on the zenith angle θ is shown in Fig. 4.
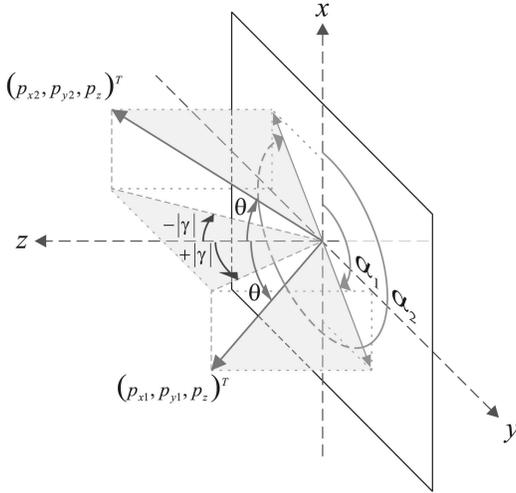
Fig. 5. Definition of angles and the two possible surface normals. γ is used in Section 4. p_{x1} and p_{y1} can be calculated using (6) with $\alpha = \alpha_1$. For p_{x2} and p_{y2}, $\alpha = \alpha_2$ is used.

The azimuth angle of the surface normal is also intimately related to the Fresnel equations. As Fig. 3 shows, the component of the internally scattered light polarized parallel to the plane containing the surface normal and reflected ray is transmitted back into the air most efficiently. The orientation of this plane is therefore the phase angle and is equivalent to the surface azimuth angle. However, since two polarization angles separated by 180 degrees are equivalent, the azimuth angle can only be determined up to an ambiguity. Fig. 5 defines the angles and directions used throughout this paper. We denote the phase angle of the light as ϕ, whereas the two candidates for the surface azimuth angle α are $\alpha_1 = \phi$ and $\alpha_2 = \phi + 180$ degrees. For a fully constrained surface normal, the Cartesian components p_x, p_y, and p_z are given by

$$\begin{pmatrix} p_x \\ p_y \\ p_z \end{pmatrix} = \begin{pmatrix} \cos\alpha\sin\theta \\ \sin\alpha\sin\theta \\ \cos\theta \end{pmatrix}. \qquad (6)$$

The experimental arrangement used for this work is shown in Fig. 6. We concentrate on the case where $\theta_L \approx 0$, that is, *retroreflection*. Our experiments are performed in dark room conditions with a single-point light source. For all the work

Fig. 7. Polarization image of a porcelain bear model. (a) Intensity. (b) Phase. (c) Degree of polarization (darker areas have higher values for (c)).

presented in this paper, gray-scale images of intensities I_0, I_{45}, and I_{90} were acquired with the polarizer oriented at $0, 45$, and 90 degrees, respectively, where 0 degrees corresponds to vertical alignment. The following equations were then used to calculate the intensity, phase, and degree of polarization, which collectively form the *polarization image* [31]:

$$I = I_0 + I_{90}, \qquad (7)$$

$$\phi = \begin{cases} \frac{1}{2}\arctan\left(\frac{I_0 + I_{90} - 2I_{45}}{I_{90} - I_0}\right) & \text{if} \quad I_{90} < I_0 < I_{45} \\ \quad'' \quad\quad\quad + 180° & \text{if} \quad I_{90} < I_0 \quad \text{and} \quad I_{45} < I_0 \\ \quad'' \quad\quad\quad + 90° & \text{otherwise}, \end{cases}$$
$$(8)$$

$$\rho = \frac{I_{90} - I_0}{(I_{90} + I_0)\cos 2\phi}. \qquad (9)$$

An example of the three components of a polarization image for a white porcelain object (a bear model) is shown in Fig. 7.

For all of the results in this paper, we used a Nikon D70 digital single-lens reflex (SLR) camera, with an exposure of $1/30$ second and an aperture of $f5.6$. Raw images were stored in gray scale using 256 levels. When the reflectance function estimation technique described below is used, we found that the contribution to the error of the final zenith angle estimate due to camera noise is small compared to the overall uncertainty. (An inexact refractive index estimate and other factors make a greater contribution). This is not always true for the azimuth angle estimates, so we use the average of four

Fig. 6. Schematic of a polarization image acquisition system. If diffuse reflection is being studied, then the best results are obtained when the lamp is collimated as this reduces reflections from the environment. These effects are minimal when our reflectance function is used.

frames before any processing is conducted. We assume that the reflected radiance is proportional to the measured pixel brightness. In fact, experiments show that this is not quite true and there is potential for improving the algorithm by taking the camera response function into consideration.

2.3 Discussion

One advantage of exploiting diffuse reflection compared to specular reflection is that the relationship between the degree of polarization and the zenith angle is one to one for the former reflection type but one to two for the latter [16]. Another advantage is that, assuming the light becomes completely depolarized after surface penetration, then the polarization state of the incident light can be arbitrary for diffuse reflection.

The main difficulties with using diffuse reflection are as follows: First, the data is generally much noisier for diffuse reflection due to the weaker polarizing effects. Second, although the dependence upon the refractive index is not strong for either reflection type, it is weaker for specular reflection. Finally, interreflections between points are mainly specular and cause erroneous orientation estimates if the theory for diffuse reflection is applied.

Both diffuse and specular reflections are affected by roughness. At the microscopic scale, a rough surface can be regarded as a set of facets with Gaussian random orientation. Due to external scattering between these facets and after subsurface scattering followed by re-emission, light impinging on a single element of the camera charge-coupled device (CCD) chip has a distribution of polarization angles. The degree of polarization is therefore observed to be reduced, resulting in an underestimate of the zenith angle. The azimuth angle estimate remains accurate provided that the mean facet orientation matches the local mean surface orientation.

3 REFLECTANCE FUNCTION ESTIMATION

In this section, we present our method for calculating the reflectance function using the zenith angles estimated from the polarization data (5) and the intensity images. The aim is to improve the accuracy of the zenith angle estimates and to reduce the effects of noise, interreflections, and surface roughness. Specifically, we show how the intensity depends on the zenith angle for the case where the camera and light source lie in the same direction. This allows the zenith angles to be estimated from the measured intensities via a lookup table. The method discussed in this section corresponds to the upper part of Fig. 1.

3.1 Preliminaries

The BRDF $f(\theta_i, \alpha_i, \theta_r, \alpha_r)$ of a particular material is the ratio of the reflected radiance $L_r(\theta_r, \alpha_r)$ to the incident irradiance $L_i(\theta_i, \alpha_i)$ for the given illumination and viewing directions. It is measured per unit solid angle per unit foreshortened area and is given by

$$f(\theta_i, \alpha_i, \theta_r, \alpha_r) = \frac{L_r(\theta_r, \alpha_r)}{L_i(\theta_i, \alpha_i) \cos \theta_i \ \mathrm{d}\omega}, \qquad (10)$$

where θ and α, respectively, denote the zenith and azimuth angles, and the subscripts i and r, respectively, denote the directions of light incidence and reflectance [34, Section 4.2].

For this paper, we work under retroreflection conditions. This means that the light source and camera directions are identical at any given visible point on the surface. In addition, if we assume that the BRDF is isotropic, then the reflected radiance is independent of the azimuth angle α. Furthermore, the zenith angles in (10) are equal ($\theta_i = \theta_r$). Therefore, the BRDF depends only on θ and can be reduced to $f(\theta)$. As a result, the reflected radiance is

$$L_r(\theta) = f(\theta) L_i(\theta) \cos \theta \ \mathrm{d}\omega. \qquad (11)$$

If we assume that the target object subtends a small angle with respect to the optical axis of the camera, then the irradiance on the sensing element in the camera (CCD chip) is

$$L_{\mathrm{CCD}} = \frac{\pi}{4} \left(\frac{A_L}{A_{\mathrm{II}}} \right)^2 L_r \propto L_r, \qquad (12)$$

where A_L is the lens diameter, and A_{II} is the distance between the lens and the image plane [34, Section 4.2]. As mentioned earlier, we also assume that the camera response function is linear, that is, pixel gray-scale values $I \propto L_r$ (or L_{CCD}).

In effect, we are estimating a unidimensional reflectance distribution function. Since the measured pixel intensity is independent of the azimuth angle under retroreflection conditions, the azimuth angle ambiguity does not complicate the reflectance function estimation process. The incident radiance $L_i(\theta)$ is not generally known since we work with an uncalibrated light source. We are therefore estimating the reflected radiance up to an unknown constant. However, this constant is unimportant for shape recovery, since we only require the relationship between the surface normal and the measured *pixel brightness*.

For our shape recovery algorithm presented in Section 4, we do not use the reflectance function $f(\theta)$ directly, but the *radiance function* $L_r(\theta)$, which incorporates the $\cos \theta$ term (11). That is, $L_r(\theta) \propto f(\theta) \cos \theta$. For the rest of this section, we therefore concentrate on the estimation of the radiance function. It is clearly trivial to calculate the reflectance function up to an unknown constant with the radiance function to hand.

Our technique makes use of the 2D histogram of zenith angles and intensities (that is, the observed distribution of the gray-scale values with the zenith angles). The histogram is defined formally as follows: From the polarization data, we have a set of Cartesian pairs (zenith angles and measured pixel brightness), $\mathcal{D} = \{(\theta_d, I_d); d = 1, 2 \ldots |\mathcal{D}|\}$, where θ_d is the zenith angle estimated from (5), and I_d is the measured pixel gray-scale value at the pixel indexed d. We wish to approximate the radiance function $L_r(\theta)$ in terms of a set of discretely sampled Cartesian pairs, $\hat{L}_r = \{(\theta_k, I_k); k = 1, 2 \ldots k_{\max}\}$, where $L_r(\theta_k) = I_k$, and we choose $k_{\max} = 100$.

The histogram contents for bin (i, j) are given by $HC_{i,j} = |\mathcal{BC}_{i,j}|$, where $\mathcal{BC}_{i,j}$ is the subset of $\mathcal{D}$ that contains the Cartesian pairs falling into the relevant bin. This subset is given by

$$\mathcal{BC}_{i,j} = \left\{ (\theta_d, I_d); \begin{array}{ll} \theta_d & \in \quad \left[\theta_i^{(\mathrm{bin})} - w_\theta/2, \ \theta_i^{(\mathrm{bin})} + w_\theta/2 \right), \\ I_d & \in \quad \left[I_j^{(\mathrm{bin})} - w_I/2, \ I_j^{(\mathrm{bin})} + w_I/2 \right) \end{array} \right\}, \qquad (13)$$

where w is the bin size that is arbitrarily chosen to be $w_\theta = 90°/k_{\max}$ and $w_I = 255/k_{\max}$. $\theta_i^{(\mathrm{bin})}$ and $I_i^{(\mathrm{bin})}$ are equally
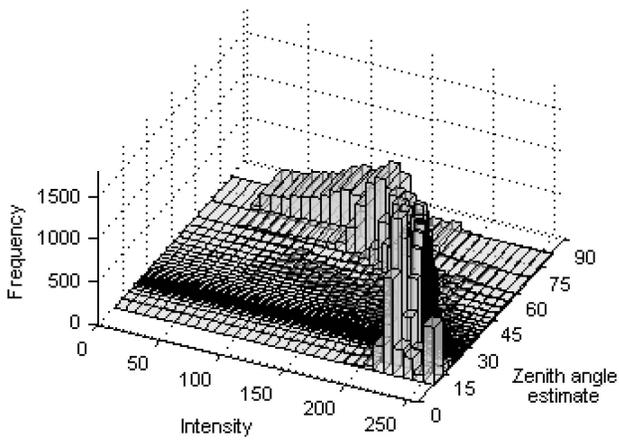
Fig. 8. Histogram showing the frequency of occurrence of intensity-zenith angle pairs for the polarization image of the bear model in Fig. 7.

spaced between 0 and 90 degrees and between 0 and 255, respectively. Fig. 8 shows a histogram of the intensities and the zenith angles for the porcelain bear model (for this figure only, unequal bin sizes are used to aid visualization).

3.2 Histogram Interpretation

Before introducing our reflectance function estimation method, we describe a simple experiment that illustrates the structure of the histogram defined above for a real-world example. We also show how the histogram can be interpreted using physical considerations. The footprint of the histogram from Fig. 8 is shown in Fig. 9. Note that there are a small but significant number of pixels that do not fall on or even near the peak frequency curve and that the curve itself is broader than one might initially expect.

Figs. 9b, 9c, and 9d show gray-scale images of the bear model. Pixels falling into the boxes marked in Fig. 9a are highlighted. The first box includes pixels falling onto the peak frequency curve of the histogram. Pixels that fall into boxes 2 and 3 clearly do not follow the general trend and make the task of shape recovery and reflectance function estimation more difficult.

The peak frequency curves found in the histograms are broadened due to roughness, interreflections, and noise. As explained in Section 2.3, roughness tends to depolarize light, reducing the estimate of the zenith angle. The bear model used in Fig. 9a is relatively smooth, but later, we show examples of rougher surfaces, where the curve is broadened further. The more distant outliers in the histograms are caused by interreflections. These occur where light from a source is specularly reflected between two points on the surface and then toward the camera. This interreflection process obeys the theory for specular reflection, although a small diffuse component will generally be present also.

The exact effect of an interreflection depends upon its strength. If it is weak, then the diffuse component still dominates. The degree of polarization and, hence, the zenith angle estimate will be reduced since the phase angle of specularly reflected light is perpendicular to that of diffuse reflections. This process can be seen in box 2 of Fig. 9. Here, a small specular component is present due to reflections from the table on which the object rests. For strong interreflections, the specular component dominates. Since the polarizing properties of specular reflection are greater than that for diffuse reflection, we have a situation where the degree of polarization exceeds that which would be expected for any purely diffuse reflection (box 3). Careful examination of the ears of the bear model in Fig. 9c shows a region *surrounding* an interreflection where the specular component is weaker.

Notice that, apart from the few cases where the strong interreflection limit is met (as in box 3), the degree of polarization is always equal to or less than the value expected for a given zenith angle. This is substantiated by the fact that the exact curve (measured using an object of the same material but known to have cylindrical shape) approximately follows the outer envelope of the histogram. This was verified with other objects.

3.3 The Proposed Method

Based on the premise above, the radiance function is estimated by fitting a curve to the outer envelope of the intensity-zenith angle histogram. The proposed method allows for *some* strong (box 3 type in Fig. 9) specular interreflections, but we do assume that these do not cover a large fraction of the surface. Only one view is strictly necessary for this section, although the method described in
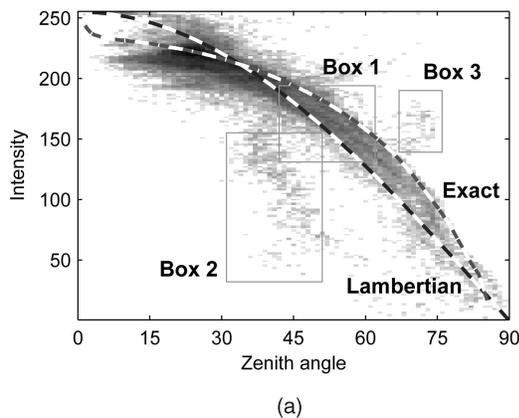
Fig. 9. (a) Footprint of the histogram in Fig. 8 where higher frequencies are shown darker (logarithmic scale). The Lambertian relationship and the exact reflectance curve are also shown. (b), (c), and (d) Gray-scale images where highlighted pixels fall into (b) box 1, (c) box 2, and (d) box 3.

(a)

(b)
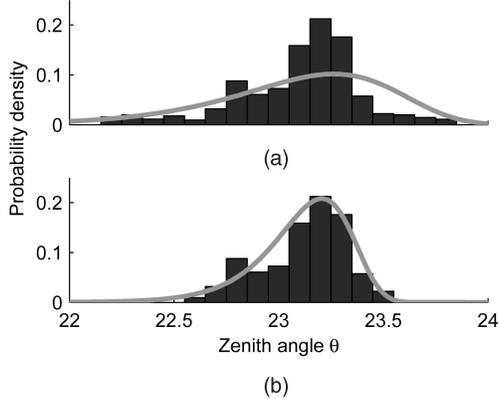
Fig. 10. Histogram of data along the fitting direction shown in Fig. 11a. (a) Initially, the fit is poor. (b) When the low-probability data is discarded using (19), the remaining data closely follows a Weibull distribution.

Section 4 requires two views (with identical lighting). For this reason, we include pixels from both views in the set $\mathcal{D}$. Experiments to estimate the radiance function from a single view were found to give similar results.

An initial estimate of the radiance function $\hat{L}_r = \{(\theta_k^{(1)}, I_k^{(1)})\}$ can be obtained using the mean intensity for regularly spaced zenith angle columns of the histogram, that is

$$\theta_k^{(1)} = \theta_k^{(\text{bin})}, \qquad I_k^{(1)} = \frac{\sum_{j=1}^{k_{\max}} HC_{k,j} I_j^{(\text{bin})}}{\sum_{j=1}^{k_{\max}} HC_{k,j}}. \qquad (14)$$

The same number of bins $(k_{\max})$ are used for sampling zenith angles and intensities. It turns out that our algorithm is not critically dependent on this initial estimate and, so, we aim for maximum computational efficiency at this point.

Next, we need the outer envelope of the histogram. The algorithm obtains this by extracting a one-dimensional (1D) histogram slice from $HC_{i,j}$ along the straight line that is perpendicular to the initial estimate at each point k. Histogram values along this line are taken from a linear interpolation of $HC_{i,j}$. Fig. 10a shows such a histogram for the straight line shown in Fig. 11a. This histogram data is then fitted to a Weibull probability density function [35, Chapter 1], which is given by

$$g(q \mid \sigma, \zeta, \mu) = \begin{cases} \frac{\zeta}{\sigma} \left(\frac{q-\mu}{\sigma}\right)^{\zeta-1} \exp\left(-\left(\frac{q-\mu}{\sigma}\right)^{\zeta}\right) & q \geq \mu \\ 0 & q < \mu, \end{cases} \qquad (15)$$

where σ is a scale parameter, ζ is a shape parameter, and μ is a location parameter (not the mean). The corresponding cumulative distribution function is

$$G(q \mid \sigma, \zeta, \mu) = \begin{cases} 1 - \exp\left(-\left(\frac{q-\mu}{\sigma}\right)^{\zeta}\right) & q \geq \mu \\ 0 & q < \mu. \end{cases} \qquad (16)$$

The Weibull distribution is primarily used in lifetime and failure analysis. It is used here since we expect most pixels to fall near the main curve in the 2D intensity-zenith angle histogram. However, due to roughness and interreflections, we also expect more pixels to be located on the left-hand side of the crest (in Fig. 8) than the right-hand side. The 1D histogram slice shown in Fig. 10 justifies this since there is a sharper cutoff on the right-hand side of the histogram.

The fitting process yields parameters $\sigma_{\text{opt}}^{(1)}$, $\zeta_{\text{opt}}^{(1)}$, and $\mu_{\text{opt}}^{(1)}$. One option to obtain the Cartesian points in our radiance function $\hat{L}_r = \{(\theta_k^{(\text{fin})}, I_k^{(\text{fin})})\}$ is to use

$$\theta_k^{(\text{fin})} = \underset{\theta}{\arg\max}\, g\left(\theta \mid \sigma_{\text{opt}}^{(1)}, \zeta_{\text{opt}}^{(1)}, \mu_{\text{opt}}^{(1)}\right), \qquad (17)$$

$$I_k^{(\text{fin})} = I_k - \frac{\theta_{k+1} - \theta_{k-1}}{I_{k+1} - I_{k-1}}\left(\theta_k^{(\text{fin})} - \theta_k\right), \qquad (18)$$

where (18) is on the straight line perpendicular to the initial radiance function estimate. However, as Fig. 10a shows, the data is not always unimodal and contains outliers, especially in the presence of interreflections. We therefore add the following robust fitting step. If τ is the lower limit of a bin on the 1D histogram and w is the bin size, then the data is discarded for bins that satisfy

$$\int_{\tau}^{\tau+w} g\left(\theta \mid \sigma_{\text{opt}}^{(1)}, \zeta_{\text{opt}}^{(1)}, \mu_{\text{opt}}^{(1)}\right) d\theta < 0.1, \qquad (19)$$

where w is chosen such that the number of bins is fixed at $k_{\max}$.

The probability density function parameters are then re-estimated iteratively until convergence to give $\sigma_{\text{opt}}^{(\text{it})}$, $\zeta_{\text{opt}}^{(\text{it})}$, and $\mu_{\text{opt}}^{(\text{it})}$. Condition (19) is applied after each iteration. We found,

(a)

(b)
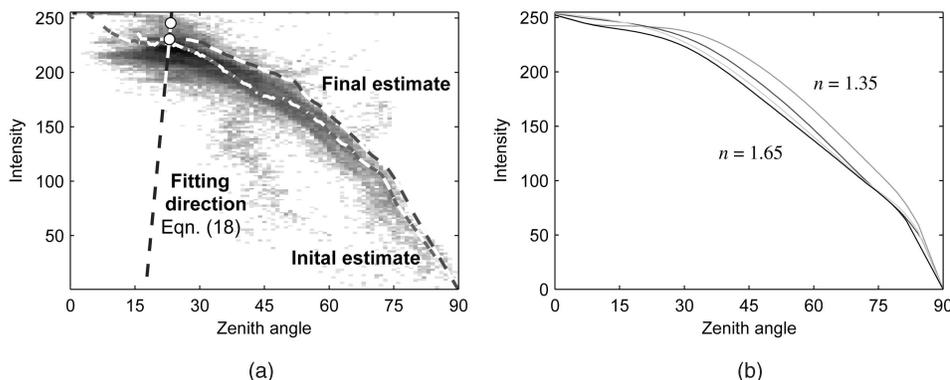
Fig. 11. (a) Estimation of the radiance function. The upper circle on the fitting curve is an estimate of the function before the robust data reduction step is performed (see (19) and Fig. 10). The lower circle indicates the final estimate. (b) Effect of different refractive indices on the smoothed reflectance function estimate.
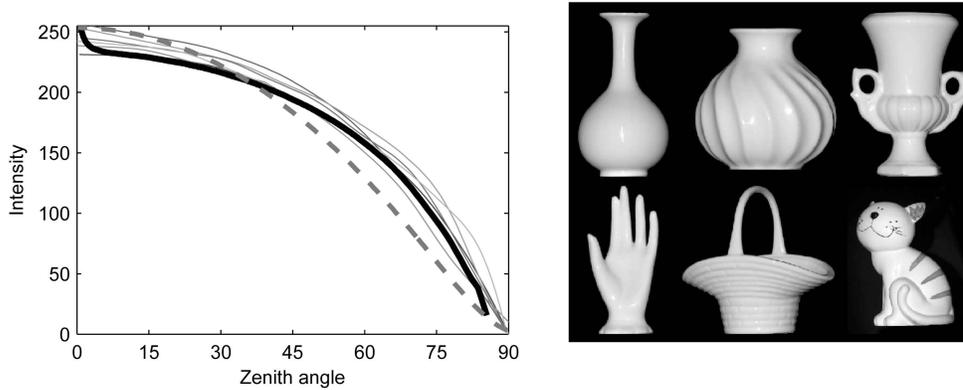
Fig. 12. Estimates of the radiance function of the selection of porcelain objects shown to the right (and the bear model from Fig. 9) compared to the exact curve (thick line) and the Wolff model (broken line). The recovered curves were linearly extrapolated to $\theta = 0$.

empirically, that the parameters converge after only few (~ 4) iterations.

The outer envelope can then be taken as

$$\theta_k^{(\text{fin})} = \theta : G\left(\theta \mid \sigma_{\text{opt}}^{(\text{it})}, \zeta_{\text{opt}}^{(\text{it})}, \mu_{\text{opt}}^{(\text{it})}\right) = 0.99, \qquad (20)$$

where (18) is used to obtain the associated $I_k^{(\text{fin})}$. In other words, the point where the cumulative probability density function reaches 99 percent is regarded as the outer envelope.

The choice of 99 percent is somewhat arbitrary, and it is possible that the algorithm would give better results if the value was fine tuned depending on the material (for example, if the material is known, then an experiment with a sphere or cylinder may help to obtain an optimum value). This is particularly true for rougher surfaces. In the case of smooth surfaces with no noise, the fitted PDF is very narrow so the exact choice of threshold has little impact on results.

Finally, the discretely sampled radiance function $\hat{L}_r$ is smoothed using moving averages. Since the relationships between the intensity and zenith angles are usually simple for retroreflection (monotonic, continuous function, continuous derivative, and so forth), we can apply the smoothing at a moderately intense level. Fig. 11b shows the result of the smoothing process. The figure also shows the effect of different refractive indices on the estimated reflectance function between two extreme values. Note that the curves are similar for high refractive indices but deviate further for lower values. Finally, we apply the constraint that $\lim_{\theta \to 90^\circ} L_r(\theta) = 0$ to estimate the reflectance function at very large zenith angles [5].

3.4 Measured Reflectance Functions

A set of porcelain objects, together with estimates of their radiance functions, are shown in Fig. 12. The objects are made from similar types of porcelain and all have smooth surfaces. The cat model is also partially painted and is used to investigate how areas of different albedo affect algorithm performance. The first point to note is that all the objects gave similar radiance functions, which closely matched the exact curve.

Fig. 13 shows that pixels of the painted areas of the porcelain cat model fall well within the envelope of the histogram. This meant that the algorithm recovered the radiance function of the porcelain, whereas the smaller painted areas were disregarded by the robust fitting method.

Note, however, that the algorithm does not always recover the radiance function of the most prevalent color/material. For example, suppose that most of the cat model was red but that some areas were white. The results would then be strongly affected by both red and white areas. Further investigation of textured objects is out of the scope of this paper. In future work, we hope to segment images according to spectral composition, allowing several reflectance functions to be estimated.

Fig. 12 also compares the results obtained to the physics-based model developed by Wolff [36], which accounts for multiple subsurface scattering using a Fresnel attenuation factor (based on the same theory as that presented in Section 2). This model uses a scattering theory originally derived for radiative transfer to model the distribution of scattered light within a medium. According to Wolff, the BRDF is

$$f(\theta_i, \theta_r, n) = \varrho\,(1 - F(\theta_i, n))\left(1 - F\left(\arcsin\left(\frac{\sin\theta_r}{n}\right), \frac{1}{n}\right)\right), \qquad (21)$$

where ϱ is called the total diffuse albedo, and F is the Fresnel function given by

$$F(\theta_i, n) = \frac{1}{2}\frac{\sin^2(\theta_i - \theta_t)}{\sin^2(\theta_i + \theta_t)}\left(1 + \frac{\cos^2(\theta_i + \theta_t)}{\cos^2(\theta_i - \theta_t)}\right), \qquad (22)$$

where $n = (\sin\theta_i)/(\sin\theta_t)$ (Snell's law).
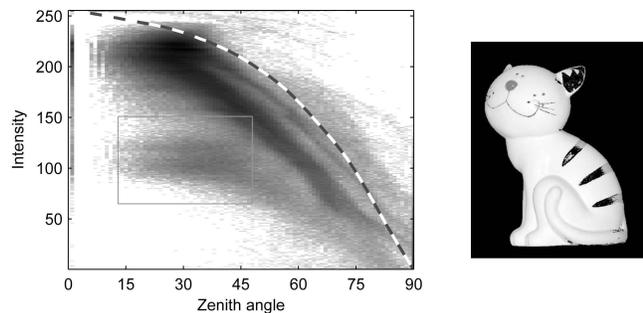
Fig. 13. Histogram for the partially painted porcelain cat. The areas falling into the box are shown in black with the rest of the image brightened for clarity (see Fig. 12 for the original image). The estimated radiance function is also shown.
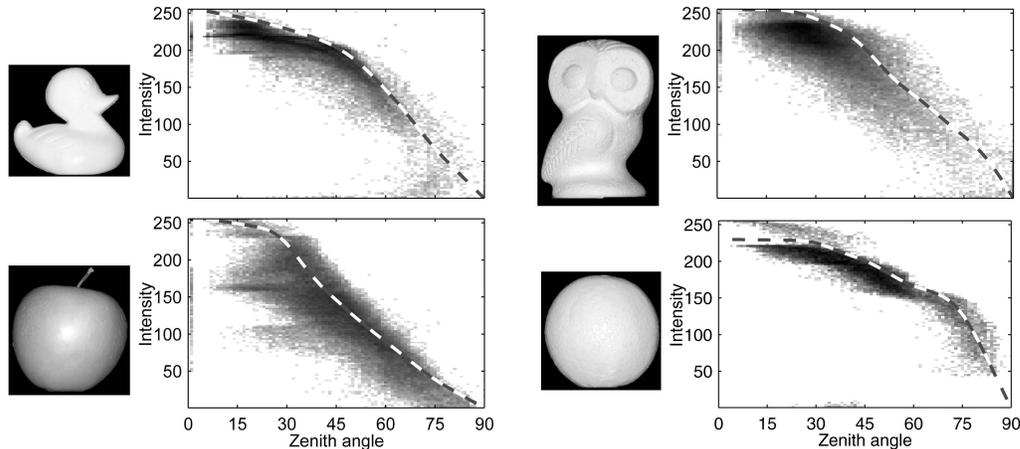
Fig. 14. Histograms and radiance functions for a slightly rough plastic duck, a plaster owl, an apple, and an orange.

Since we are studying the case where $\theta_i = \theta_r$, the two Fresnel terms in (21) are equal. Allowing for the foreshortening effects of orthographic imaging, (21) then simplifies to

$$I = \kappa \cos \theta (1 - F(\theta, n))^2, \qquad (23)$$

where κ is a normalizing constant. We use this equation in Fig. 12, where κ is selected to give the best agreement with the exact curve, which is also shown in the figure. Note that the use of reflectance models requires estimates of various surface or material parameters, whereas our method needs only an approximate value of the refractive index.

We have also applied our method to various slightly rough surfaces. Examples of which are shown in Fig. 14. Results are generally reasonable, although the estimated radiance curves tend to be least reliable for small zenith angles, where the radiance function is often overestimated.

4 RECONSTRUCTION FROM TWO VIEWS

In this section, we describe our method for shape estimation from two views. The goal is to establish and use dense stereo correspondence to resolve azimuth angle ambiguities present in the surface normal estimates from a single view. The first step is to extract surface patches from the polarization images. These patches are then aligned using an optimization scheme to obtain correspondence indicators. The error criterion is similar to those developed by Cross and Hancock [37] using the EM algorithm and Chui and Rangarajan [38] using softassign.

The input to this stage of the algorithm is the set of ambiguous surface normal azimuth angles from the phase images together with their zenith angles from the estimated radiance function and the raw intensity images. We use the setup shown in Fig. 6 with polarization images taken from two views. For the second view, the turntable is rotated by $\theta_{\rm rot}$ about the x-axis (see Fig. 6 for the coordinate axis convention). In this paper, we use a known value of $\theta_{\rm rot} = 20°$.

The set of surface patches are obtained by segmenting the two polarization images into regions[1] according to the estimated surface normals. The segmentation is performed

using a bithreshold technique, where normals falling into an angular interval are selected from each image. The angular interval is shifted by $\theta_{\rm rot}$ between images to account for the object rotation. Details of this procedure can be found in Section 4.1.

The correspondence problem is then reduced to matching patches from the two views. The solution is to align patch pairs for which a potential correspondence exists. We adopt a dual-step matching procedure in which both correspondence indicators and alignment parameters are sought. The cost function uses the correspondence indicators to gate contributions to a sum-of-squares alignment error for corresponding points. The alignment error measures the difference in the surface normal components and a scale invariant measure of surface topography (the shape index). The matching cost is calculated for all potentially corresponding patch pairs, and the combination that gives the lowest total cost is taken as the set of patch correspondences. The derivation of the cost function that we adopt is given in Section 4.2. The matching algorithm itself and how it leads to unambiguous azimuth angles is described in detail in Section 4.3.

After this process, some areas of the images remain without detected correspondence. Monotonic interpolation is then used to estimate correspondences for such areas. This stage of the method is detailed in Section 4.4. Finally, the Frankot-Chellappa surface integration method is applied to recover a depth map from the field of unambiguous surface normals. In the remainder of this section, we furnish details of the steps described above. Refer back to Fig. 1 for a schematic of the overall algorithm structure.

4.1 Segmentation

The purpose of this stage of the algorithm is to segment the two images into regions that are suitable for establishing correspondence. The segmentation is performed according to the angle defined by

$$\gamma = \arctan(\sin \alpha \tan \theta). \qquad (24)$$

This is the angle between the viewing direction and the projection of the surface normal onto the horizontal plane (the y-z plane in Fig. 5). If the correct azimuth angle is greater than 180 degrees, then γ is negative, that is, the surface normal vector points to the left from the viewer's perspective. However, we do not know the sign of γ at this

---

1. We refer to each unconnected 2D image segment as a "region" and use the term "patch" to refer to the 3D data (that is, the surface normals that can be used to estimate surface geometry) for each of these segments.
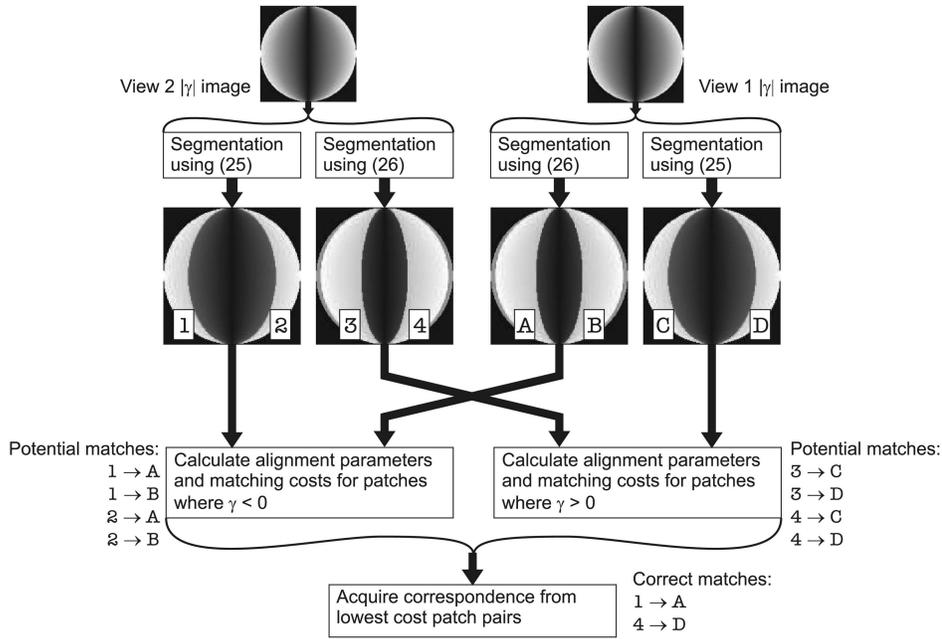
Fig. 15. Flow diagram of the segmentation and correspondence part of the algorithm, using a synthetic sphere as an example. This figure corresponds to the large box in Fig. 1.

point so we can only calculate the modulus of the angle from the estimated phase, that is, $|\gamma| = \arctan(\sin\phi\tan\theta)$.

The zenith angle θ is determined by the radiance function (11), estimated using the technique from Section 3.3. Let the minimum angle for reliable $|\gamma|$ estimates be γ_{lim}. Empirically, we found that γ_{lim} was approximately 20 degrees for perfectly smooth objects. For rougher surfaces, a reasonable reconstruction was possible using this value, but results were better when $\gamma_{\text{lim}} = 30°$ was used. We know that for the regions where $\gamma < 0$ for both images, then, where the pixels in the right-hand image (view 2 in Fig. 6) satisfy the condition

$$\gamma_{\text{lim}} + \theta_{\text{rot}} < |\gamma| \le 90°, \quad (25)$$

the corresponding pixels in the left-hand image (view 1) must satisfy the condition

$$\gamma_{\text{lim}} < |\gamma| \le 90° - \theta_{\text{rot}}. \quad (26)$$

Conversely, where $\gamma > 0$, then for the regions of the right-hand image that satisfy the condition in (26), the pixels in the corresponding left-hand image regions must satisfy condition in (25).

Our region segmentation algorithm makes two image passes for each view. On the first pass, the right-hand image is

thresholded using the angle condition in (25), whereas the left-hand image is thresholded using the condition in (26). On the second pass, the angle conditions are interchanged between the left and right-hand views. This two-pass thresholding gives two sets of regions per view. Fig. 15 illustrates the possible correspondence combinations between region sets. In this way, we impose angle interval constraints on the allowable region correspondences. Fig. 16 shows a real-world example of the result of the segmentation process.

4.2 Cost Functions

As already mentioned, the algorithm seeks patch correspondences so as to minimize a cost (or energy) function. In this section, we first describe a rudimentary cost function based purely on local patch reconstructions. We then introduce a more complex cost function based directly on the estimated surface normals and the scale invariant topographic quantity known as the shape index.

4.2.1 A Rudimentary Cost Function

This cost function is based on the similarity between local patch reconstructions obtained by applying the Frankot-Chellappa method to recover surface depth h from needle maps [2]. Fig. 17 shows patch reconstructions from two of the larger regions from Figs. 16a and 16b, using $\gamma < 0$.

Consider the alignment of two patches defined by sets of points $\mathcal{U} = \{\mathbf{u}_a; a = 1, 2 \ldots |\mathcal{U}|\}$ and $\mathcal{V} = \{\mathbf{v}_b; b = 1, 2 \ldots |\mathcal{V}|\}$, where $\mathbf{u}_a = (x_a^{(\text{u})}, y_a^{(\text{u})}, z_a^{(\text{u})})^T$ and $\mathbf{v}_b = (x_b^{(\text{v})}, y_b^{(\text{v})}, z_b^{(\text{v})})^T$. The alignment is performed by applying a transformation function J to patch $\mathcal{U}$ to obtain the set of transformed positions $\hat{\mathcal{U}} = \{\hat{\mathbf{u}}_a\}$, where $\hat{\mathbf{u}}_a = (\hat{x}_a^{(\text{u})}, \hat{y}_a^{(\text{u})}, \hat{z}_a^{(\text{u})})^T$.

For our case, where the target object has undergone a known rotation about the vertical axis (x-axis), the transformation contains a y-translation of Δy, a z-translation of Δz, and a rotation about the x-axis of θ_{rot}. For our
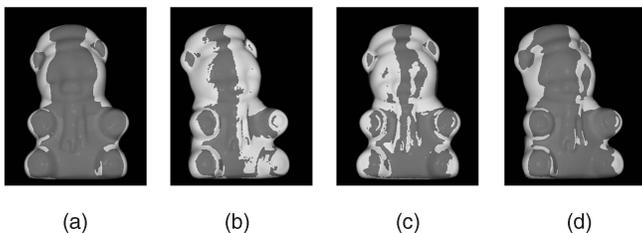
Fig. 16. Segmentation of a real-world image of a porcelain bear model. In (a) and (d), the highlighted regions obey (25), whereas in (b) and (c), they obey (26).
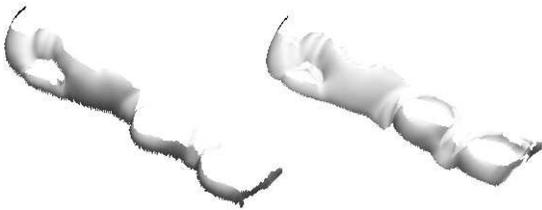
Fig. 17. Patch reconstructions of the leftmost regions of Figs. 16a and 16b.

Fig. 18. Illustration of how matrices $M(\overline{\Theta})$ and $M'(\overline{\Theta})$ are constructed. The points shown are from a horizontal cross-section of two patches. All broken lines correspond to elements of $M(\overline{\Theta})$ set to unity. For $M'(\overline{\Theta})$ only the dark broken lines correspond such elements. The elements of $M(\overline{\Theta})$ are shown to the right. For the reduced correspondence matrix, all elements are zero except for those highlighted.

particular transformation, therefore, $\hat{x}_a^{(u)} = x_a^{(u)} \; \forall a \in \{1, 2 \ldots |\mathcal{U}|\}$. We represent the transformation parameters as $\overline{\Theta} = (\Delta y, \Delta z, \theta_{\rm rot})^T$, where $\theta_{\rm rot}$ is known. Our task then becomes that of estimating the vector of parameter values $\overline{\Theta}$ for each potentially corresponding patch pair that minimizes an energy functional.

The rudimentary energy functional for aligning patches $\mathcal{U}$ and $\mathcal{V}$ is

$$\varepsilon_{\mathcal{UV}}(\overline{\Theta}) = \frac{\sum_{a=1}^{|\mathcal{U}|} \sum_{b=1}^{|\mathcal{V}|} m_{ab}(\overline{\Theta}) |\mathbf{v}_b - \hat{\mathbf{u}}_a(\overline{\Theta})|^2}{\sum_{a=1}^{|\mathcal{U}|} \sum_{b=1}^{|\mathcal{V}|} m_{ab}(\overline{\Theta})}, \quad (27)$$

where $m_{ab}(\overline{\Theta})$ are the elements of the correspondence indicator matrix $M(\overline{\Theta})$, defined below. This is a modified version of the least squares alignment parameter estimation problem. The correspondence indicators exclude the contributions to squared alignment error from nonmatching points. Similar cost functions are obtained by Cross and Hancock [37] and Gold et al. [39]. In this paper, the elements of the correspondence indicator matrix are assigned as follows:

$$m_{ab}(\overline{\Theta}) = \begin{cases} 1 & \text{if} \begin{cases} a = \underset{a' \in \mathcal{A}(b)}{\mathrm{argmin}} \left| \hat{y}_{a'}^{(u)}(\overline{\Theta}) - y_b^{(v)} \right| & \textbf{or} \\ b = \underset{b' \in \mathcal{B}(a)}{\mathrm{argmin}} \left| \hat{y}_a^{(u)}(\overline{\Theta}) - y_{b'}^{(v)} \right| \end{cases} \\ 0 & \text{otherwise,} \end{cases} \quad (28)$$

where $\mathcal{A}(b) = \{a \in 1, 2 \ldots |\mathcal{U}| \mid x_a^{(u)} = x_b^{(v)}\}$ is the set of labels for points in $\mathcal{U}$ that lie in the same horizontal plane as point $\mathbf{v}_b$. Similarly, $\mathcal{B}(a) = \{b \in 1, 2 \ldots |\mathcal{V}| \mid x_b^{(v)} = x_a^{(u)}\}$ are the labels for points in $\mathcal{V}$ in the same plane as point $\mathbf{u}_a$. In other words, $m_{ab}(\overline{\Theta}) = 1$ if $\mathbf{v}_b$ is the closest point in $\mathcal{V}$ to $\hat{\mathbf{u}}_a(\overline{\Theta})$ in the y-direction within the same horizontal cross-section. Similarly, $m_{ab}(\overline{\Theta}) = 1$ also if $\hat{\mathbf{u}}_a(\overline{\Theta})$ is the closest point in $\hat{\mathcal{U}}(\overline{\Theta})$ to $\mathbf{v}_b$. The construction of the matrix $M(\overline{\Theta})$ is illustrated in Fig. 18. Note that it does not include a term with the difference in the z-position. This is because, in the final algorithm, the energy functional is given in terms of the surface normals, so the height reconstruction (difference in z-position) is not needed to establish correspondence.

The matrix $M(\overline{\Theta})$ is hence defined in a similar fashion to the correspondence matrix used by Chui and Rangarajan [38]. Note that Chui and Rangarajan assume that the correspondence is one to one, meaning that each row and each column of the final correspondence matrix has exactly one element set to unity. We do not enforce this constraint here.

Using the correspondence indicator matrix defined in (28) means that any points in nonoverlapping areas of the patches are inappropriately included in the energy calculation. We

therefore use a reduced correspondence matrix $M'(\overline{\Theta})$ with elements m'_{ab}, which is defined by

$$m'_{ab}(\overline{\Theta}) = \begin{cases} m_{ab}(\overline{\Theta}) & \text{if} \begin{cases} \underset{a' \in \mathcal{A}(b)}{\min}\left\{\hat{y}_{a'}^{(u)}(\overline{\Theta})\right\} < \; y_b^{(v)} \; < \underset{a' \in \mathcal{A}(b)}{\max}\left\{\hat{y}_{a'}^{(u)}(\overline{\Theta})\right\} \\ \textbf{and} \\ \underset{b' \in \mathcal{B}(a)}{\min}\left\{y_{b'}^{(v)}\right\} \; < \hat{y}_a^{(u)}(\overline{\Theta}) < \underset{b' \in \mathcal{B}(a)}{\max}\left\{y_{b'}^{(v)}\right\} \end{cases} \\ 0 & \text{otherwise.} \end{cases}$$
$$(29)$$

This reduced correspondence indicator matrix therefore discards points in the patches where there is no overlap, as shown in Fig. 18.

Note that this method is an alternative to the iterated closest point (ICP) algorithm [40]. We do not use ICP because that method would restrict our cost function to be based on the location of points in 3D space. Although the use of (27) specifically has no advantage over ICP in this respect, we would ideally like to base our cost primarily on the surface normals, rather than on patch reconstructions, to avoid error propagation.

4.2.2 An Improved Cost Function

We aim here to refine the cost function to include surface normal information. However, if *only* the surface normals are used, then we encounter a problem: It is possible that two surface points on different patches could have identical surface normal directions but very different local shapes. We therefore introduce the shape index into the energy functional to account for local surface topography.

The shape index s is a view-invariant quantity that describes the local topography of a surface and falls in the interval $[-1, +1]$. Where $s = -1$, the local surface shape is a concave spherical cup, where $s = 0$, a saddle point is found, and where $s = +1$, the surface takes the form of a spherical dome. Between these values, the surface takes intermediate topographic forms. The shape index is undefined for planar surfaces.

The starting point for the derivation of the shape index is the Hessian matrix, defined by

$$H = \begin{bmatrix} \partial^2 h/\partial x^2 & \partial^2 h/\partial x \partial y \\ \partial^2 h/\partial x \partial y & \partial^2 h/\partial y^2 \end{bmatrix} = \begin{bmatrix} \partial p'_x/\partial x & \partial p'_x/\partial y \\ \partial p'_y/\partial x & \partial p'_y/\partial y \end{bmatrix}, \quad (30)$$

where $h(x, y)$ is the surface height. p'_x and p'_y are the gradients of the surface and are related to the Cartesian components of the unit surface normal in (6) by $p'_x = -p_x/p_z$ and $p'_y = -p_y/p_z$.

The Hessian matrix can be calculated from the normals estimated using polarization data and the radiance function. Since we are performing a differentiation, the estimate of H is not entirely robust to noise. However, since the Hessian will not form the main part of our cost function (the normals themselves do), this problem is not severe.

Using (30) directly leads to a view-biased estimate of the Hessian, due to a foreshortening effect. We therefore introduce the viewpoint-compensated Hessian derived by Woodham [41]

$$C = \frac{1}{\sqrt{(1 + p_x{}^2 + p_y{}^2)^3}} \begin{bmatrix} p_y^2 + 1 & -p_x p_y \\ -p_x p_y & p_x^2 + 1 \end{bmatrix} H. \quad (31)$$

The eigenvalues of this matrix are

$$\lambda_{\max} = -\frac{1}{2}(c_{11} + c_{22} - S), \quad (32)$$

$$\lambda_{\min} = -\frac{1}{2}(c_{11} + c_{22} + S), \quad (33)$$

where $S = \sqrt{(c_{11} - c_{22})^2 + 4(c_{21}c_{12})}$, and c_{ij} denotes elements of C. $\lambda_{\max}$ and $\lambda_{\min}$ are the principle curvature directions at the surface point and are the sole variables in the definition of shape index

$$s = \frac{2}{\pi}\arctan\left(\frac{\lambda_{\max} + \lambda_{\min}}{\lambda_{\max} - \lambda_{\min}}\right). \quad (34)$$

In our energy functional, we motivate the use of the difference in shape index between surface points as follows: Consider two surface points with identical normal directions with one lying on a spherical dome ($s = +1$) and the other on a spherical cup ($s = -1$). The difference in shape index is then $+2$, which increases the matching cost. Based on this observation, we propose the following form, which incorporates both the shape indices and the surface normals

$$\varepsilon_{uv}(\overline{\Theta}) = \frac{\sum_{a=1}^{|\mathcal{U}|} \sum_{b=1}^{|\mathcal{V}|} m'_{ab}(\overline{\Theta}) P_{ab}(\overline{\Theta}) \Delta s_{ab}^2}{\sum_{a=1}^{|\mathcal{U}|} \sum_{b=1}^{|\mathcal{V}|} m'_{ab}(\overline{\Theta})}, \quad (35)$$

where $P_{ab}(\overline{\Theta}) = (p_{x_b} - \hat{p}_{x_a}(\overline{\Theta}))^2 + (p_{y_b} - \hat{p}_{y_a}(\overline{\Theta}))^2 + (p_{z_b} - \hat{p}_{z_a}(\overline{\Theta}))^2$ and $\Delta s_{ab} = s_b - s_a$.

The "hat" notation is used to represent quantities having undergone transformation J as before. Note that the shape index does not depend on the transformation parameters since it is rotation invariant. Matrix $M'(\overline{\Theta})$ is calculated in the same way as before, using the Frankot-Chellappa method to determine the nearest neighbors in the y-direction. Note that the reconstruction is effectively used only as a means to account for foreshortening (that is, so that patch $\mathcal{U}$ can be rotated by θ_{rot}).

4.3 Final Algorithm

The optimum alignment parameters for patches $\mathcal{U}$ and $\mathcal{V}$ are given by

$$\overline{\Theta}_{uv} = \underset{\overline{\Theta}}{\operatorname{argmin}}\, \varepsilon_{uv}(\overline{\Theta}). \quad (36)$$

For each pair of corresponding patches $\mathcal{W}$ and $\mathcal{W}'$, the correspondence matrix is $M(\overline{\Theta}_{ww'})$. The matching cost $\varepsilon_{uv}(\overline{\Theta}_{uv})$ is calculated for all potential matches. The algorithm sorts the list of costs into an ascending order (that is, the most confident match first) and confirms the matches sequentially but rejects those that do not satisfy the following consistency constraint:

If a point at the y-position $y_a^{(u)}$ corresponds to a point at $y_b^{(v)}$, then two other corresponding points at $y_{a'}^{(u)}$ and $y_{b'}^{(v)}$ must satisfy

$$\operatorname{sgn}\left(y_{a'}^{(u)} - y_a^{(u)}\right) = \operatorname{sgn}\left(y_{b'}^{(v)} - y_b^{(v)}\right). \quad (37)$$

Due to occlusion, not all patches have a correspondence at all, so we ideally require a cost threshold. Any matches with a cost above the threshold are then discarded. At present, we are yet to develop an adaptive means of setting this threshold.

The algorithm was first tested using (35) as the energy functional. One problem encountered when using (35) is that an area of a patch may be occluded in one view but not in the other. Another issue is that, even though (35) only needs the Frankot-Chellappa method as a convenient way to account for foreshortening, it still introduces a distortion in the patch shape and, therefore, a nonoptimal correspondence matrix. We reduce the impact of these problems by replacing (35) with the following functional, which takes the median energy of the separate costs for each horizontal cross-section of the patches:

$$\varepsilon_{uv}(\overline{\Theta}) = \frac{1}{K(\overline{\Theta})}\underset{x \in \ell}{\operatorname{median}}\left(\frac{\sum_{a=1}^{|\mathcal{U}|} \sum_{b=1}^{|\mathcal{V}|} m''_{ab}(\overline{\Theta}, x) P_{ab}(\overline{\Theta}) \Delta s_{ab}^2}{\sum_{a=1}^{|\mathcal{U}|} \sum_{b=1}^{|\mathcal{V}|} m''_{ab}(\overline{\Theta}, x)}\right). \quad (38)$$

Here, $\ell = \{x_u\} \bigcap \{x_v\}$ is the set of unique x-positions covered by both patches and

$$m''_{ab}(\overline{\Theta}, x) = \begin{cases} m'_{ab}(\overline{\Theta}) & \text{if } x_a^{(u)} = x_b^{(v)} = x \\ 0 & \text{otherwise.} \end{cases} \quad (39)$$

We have also introduced the quantity $K(\overline{\Theta})$, which rewards patch overlap

$$K(\overline{\Theta}) = \frac{\sum_{a=1}^{|\mathcal{U}|} \sum_{b=1}^{|\mathcal{V}|} m'_{ab}(\overline{\Theta})}{\sum_{a=1}^{|\mathcal{U}|} \sum_{b=1}^{|\mathcal{V}|} m_{ab}(\overline{\Theta})}. \quad (40)$$

It is not necessary for the algorithm to calculate the cost for every conceivable pair of patches. This is because we know that the two pixels in any given correspondence pair lie on the same horizontal scan line for our particular experimental arrangement. Instead, it only calculates the cost of patch pairs where there is some overlap in the vertical coordinates (that is, $\ell \neq \oslash$). For such pairs, $\varepsilon_{uv}(\overline{\Theta})$ is minimized using the Nelder-Mead method [42].

We have also experimented with introducing an affine skew parameter into the vector $\overline{\Theta}$. This is to further reduce the impact of patch distortions arising from the local reconstruction process. The skew is particularly useful for long and thin patches such as those in Fig. 17 since imperfect surface normal estimates have a greater detrimental effect in such cases. It is straightforward to add additional parameters to the translation vector using our method to increase the generality of the technique.

Fig. 19. (a) Azimuth angles of disambiguated regions of the bear model. (b) Final estimates of azimuth angles for entire image.

After the patch matching is complete, we can disambiguate the azimuth angles for the regions where correspondence was established. We do this using the knowledge that the patches reconstructed using $\gamma > 0$ have $\alpha < 180°$ (that is, $\alpha = \alpha_1$ in Fig. 5) and patches where $\gamma < 0$ have $\alpha > 180°$ ($\alpha = \alpha_2$). Fig. 19a shows these unambiguous regions for the bear model images. We still have two estimates of these unambiguous surface normals, one from each view. Our final value is taken from the view where the zenith angle was greatest since the raw polarization data would have contained less noise for this estimate.

4.4 Correspondence for Remaining Areas

Large areas of the images remain without correspondence after the above matching procedure and, thus, remain ambiguous. To assign a value for the azimuth angle for these areas, we use *piecewise cubic Hermite interpolating polynomials* (PCHIP) [43]. This form of interpolation is similar to the spline technique in that a continuous curve is generated that passes exactly through the data points. The difference is that the spline maintains continuous first *and* second derivatives, whereas PCHIP allows discontinuities in the latter. The advantage of this, from our point of view, is that PCHIP ensures that a monotonic curve is generated, thus, ensuring point correspondences satisfy the condition in (37).

The algorithm effectively plots the y-positions of corresponding points from each view against each other and applies PCHIP to interpolate between correspondences. This method only provides relatively crude *correspondence* for areas not in the vicinity of the regions disambiguated previously. This is one reason for us currently relying on integration methods to recover depth instead of triangulation. However, it is sufficient as a means of selecting the most reliable azimuth angle (α_1 or α_2) for most image pixels.

To demonstrate the effects of disambiguation errors and to further motivate our choice of the Frankot-Chellappa algorithm for shape reconstruction from surface normals,
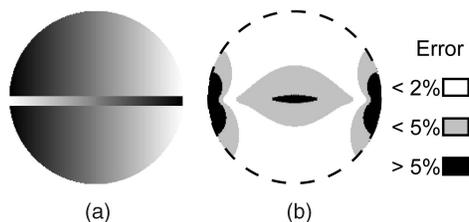
consider Fig. 20. For this figure, a synthetic sphere was created, and the surface normals were calculated at each point. For a slice of the sphere, the signs of the p_x and p_y values were changed to simulate the effects of a correspondence error. The surface normals were then integrated using the Frankot-Chellappa method to obtain a depth map. The difference between the original sphere and the recovered depth is shown in the figure and is clearly relatively small, given the magnitude of the error introduced. Had triangulation been used with such an error, then the reconstruction would be more severely deformed.

We do not use the results of PCHIP for areas of the image where $|\gamma| < \theta_{\rm rot}$. Instead, we propagate from surrounding areas and align surface normals to the mean of local 3×3 windows. This is advantageous since these areas are the most susceptible to noise due to lower degrees of polarization. Fig. 19b shows the unambiguous azimuth angles for the full image for each view.

5 RESULTS

Fig. 21 shows the recovered depth maps of some of the objects used in Section 3. They were obtained by applying the Frankot-Chellappa algorithm to the recovered needle maps. A reasonable estimate is made in each of the six cases. There are two poorly recovered object areas. First, the paws of the bear model do not protrude as they should due to interreflections. Although using the estimated radiance function helped to accurately recover the zenith angles here, the azimuth angles were incorrect for these areas. Second, the handle of the basket does not arch from one side of the object to the other as a result of the needle map integration method. Results for three other objects of different materials are shown

Fig. 20. Illustration of the effects of an incorrect disambiguation in recovered height. (a) Image of p_y for a synthetic sphere with a disambiguation error introduced. (b) Magnitude of the error in recovered height expressed as a percentage of the radius.
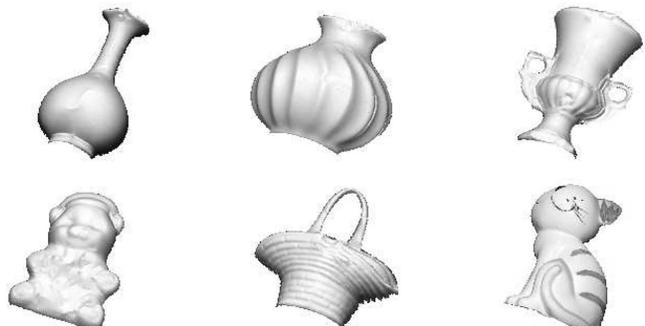
Fig. 21. Recovered 3D geometry of some of the objects in Figs. 7 and 12. The texture on the cat model has been mapped back onto the surface.

Fig. 22. Recovered shape of objects of different materials with the original textures mapped onto the surface.

Fig. 24. Ground-truth profiles of (a) the porcelain urn and (b) the orange (solid lines). The broken lines show estimates of the profile. The bottom curve in (b) was estimated without using the radiance function.

in Fig. 22. Here, the original red, green, blue (RGB) images have been used to reapply the texture onto the reconstruction. This final step is for aesthetic purposes only since we do not have full illumination control.

Fig. 23a shows ground-truth cross-sections of the porcelain vase compared to the recovered height. The ground truth was calculated using the volume of revolution of the object contour. The comparison proves that the method works well for the simplest case, namely, that of a smooth object with basic geometry. The difference image between the ground truth and the recovered height is shown in Fig. 23b. Note that any symmetries in the true object geometry are not always preserved in the error. This is a result of small environmental interreflections and is possibly accentuated by imperfect polarizer calibration. For comparison, Fig. 23c shows the difference image when the raw zenith angle estimates were used (that is, those calculated directly from (5) without estimating the radiance function). The root-mean-square height errors were 8 units when the radiance function was used and 18 units when it was not used. Note that the depth of the vase is 150 units.

Fig. 24 shows a comparison of the ground truth profiles of the porcelain urn and the orange. The urn profile shows that, although the general shape is recovered, the height variations are less pronounced than they should be for objects with complicated geometry. The method performed better for the orange. Fig. 24b also shows the profile estimate of the orange without using the radiance function. Here, the radiance function was clearly essential for accurate shape estimation.

The method has been tested on a greater range and number of objects than in most of the previous contributions, and the results show several strengths and weaknesses. When the

algorithm is applied to objects with complicated geometry, our method is less accurate than those that use specular reflection such as that in [17]. This is mainly due to higher degrees of polarization associated with specular reflection. As mentioned previously, however, specular reflection is often difficult to measure for an entire surface or scene. Our method has a more general and complete disambiguation than that in [20] and allows reconstructions from different views but gives less accurate results in some cases. Compared to SFS, our method has stronger constraints, and the results are generally superior, although less general to illumination conditions. For cases where triangulation is possible, those methods are more accurate, provided that correspondences are reliable. For this reason, we will consider a combined integration and triangulation final step for future work.

6 CONCLUSION

We have devised a new shape recovery technique that combines polarization, shading, and stereo information. Polarization images are acquired from two views of an object on a turntable. The first part of our proposed algorithm estimates the reflectance function of the object under study. This is accomplished globally using the distributions of the initial zenith angle estimates and the pixel brightnesses. Using ambiguous surface normal estimates from the polarization images and refined using the estimated reflectance function, the proposed algorithm then establishes stereo correspondence between the views. This task is accomplished using an optimization scheme, where the cost function is based on the surface normals and the shape index. Our results show that the technique is applicable to smooth objects composed of a variety of materials with a known value of the refractive index. The fact that the refractive index must be
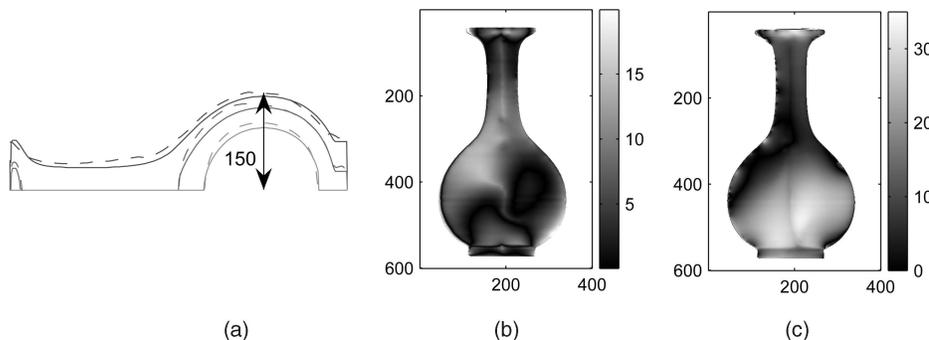
Fig. 23. (a) Three ground-truth cross-sections of the vase (solid lines) compared to the recovered height (broken lines). (b) Difference between the recovered height and ground truth when the radiance function is used to refine the zenith angles and (c) when the radiance function is not used. The scales to the right of the images indicate the absolute height error and are in the units of pixel lengths. Note that the scale is different for each image.

known is a weakness of the method, although only an approximate estimate is required.

Our method currently requires a special lighting configuration that limits potential applications. In future work, we intend to relax the retroreflection assumption to enable reflectance function estimation and shape recovery in more general illumination conditions. We aim to recover a greater fraction of the BRDF by using a fixed object and light source but different camera positions. This may also allow for the separation of specular and diffuse reflection components, reducing the effects of specularities and interreflections. Finally, estimating the camera response function may yield more accurate initial estimates of the surface normals.

The work may pave the way to the development of a sensor that can estimate object shape for smooth and slightly rough surfaces. Unlike structured light range scanners, our technique is nonintrusive. The surface matching aspect of the algorithm can be applied not only to patches of the same object, but also to different objects within a scene. It should also be possible for the sensor to classify materials within a scene according to their reflectance function.

REFERENCES

[1] R. Zhang, P.S. Tsai, J.E. Cryer, and M. Shah, "Shape from Shading: A Survey," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 21, pp. 690-706, 1999.

[2] R.T. Frankot and R. Chellappa, "A Method for Enforcing Integrability in Shape from Shading Algorithms," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 10, pp. 439-451, 1988.

[3] A. Agrawal, R. Chellappa, and R. Raskar, "An Algebraic Approach to Surface Reconstruction from Gradient Fields," *Proc. Int'l Conf. Computer Vision,* pp. 174-181, 2005.

[4] P.N. Belhumeur, D.J. Kriegman, and A.L. Yuille, "The Bas-Relief Ambiguity," *Proc. Computer Vision and Pattern Recognition,* pp. 1060-1066, 1997.

[5] L.B. Wolff, S.K. Nayar, and M. Oren, "Improved Diffuse Reflection Models for Computer Vision," *Int'l J. Computer Vision,* vol. 30, pp. 55-71, 1998.

[6] A. Treuille, A. Hertzmann, and S. Seitz, "Example-Based Stereo with General BRDFs," *Proc. European Conf. Computer Vision,* pp. 457-469, 2004.

[7] A. Robles-Kelly and E.R. Hancock, "Estimating the Surface Radiance Function from Single Images," *Graphical Models,* vol. 67, pp. 518-548, 2005.

[8] H. Ragheb and E.R. Hancock, "Surface Radiance Correction for Shape from Shading," *Pattern Recognition,* vol. 38, pp. 1574-1595, 2005.

[9] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision.* Cambridge Univ. Press, 2000.

[10] M.Z. Brown and G.D. Hager, "Advances in Computational Stereo," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 25, pp. 993-1008, 2003.

[11] H. Jin, A. Yezzi, and S. Soatto, "Stereoscopic Shading: Integrating Shape Cues in a Variational Framework," *Proc. Computer Vision and Pattern Recognitio,* pp. 169-176, 2000.

[12] D.G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," *Int'l J. Computer Vision,* vol. 60, pp. 91-110, 2004.

[13] R.J. Woodham, "Photometric Method for Determining Surface Orientation from Multiple Images," *Optical Eng.,* vol. 19, pp. 139-144, 1980.

[14] T. Zickler, P.N. Belhumeur, and D.J. Kriegman, "Helmholtz Stereopsis: Exploiting Reciprocity for Surface Reconstruction," *Int'l J. Computer Vision,* vol. 49, pp. 215-227, 2002.

[15] E. Hecht, *Optics,* third ed. Addison Wesley Longman, 1998.

[16] L.B. Wolff and T.E. Boult, "Constraining Object Features Using a Polarization Reflectance Model," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 13, pp. 635-657, 1991.

[17] D. Miyazaki, M. Kagesawa, and K. Ikeuchi, "Transparent Surface Modelling from a Pair of Polarization Images," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 26, pp. 73-82, 2004.

[18] D. Miyazaki, M. Saito, Y. Sato, and K. Ikeuchi, "Determining Surface Orientations of Transparent Objects Based on Polarisation Degrees in Visible and Infrared Wavelengths," *Optical Soc. Am. J. A,* vol. 19, pp. 687-694, 2002.

[19] L.B. Wolff, "Surface Orientation from Two Camera Stereo with Polarizers," *Proc. SPIE Conf. Optics, Illumination, Image Sensing for Machine Vision IV,* vol. 1194, pp. 287-297, 1989.

[20] D. Miyazaki, R.T. Tan, K. Hara, and K. Ikeuchi, "Polarization-Based Inverse Rendering from a Single View," *Proc. Int'l Conf. Computer Vision,* vol. 2, pp. 982-987, 2003.

[21] O. Drbohlav and R. Šára, "Using Polarization to Determine Intrinsic Surface Properties," *Proc. SPIE Conf. Polarization and Color Techniques in Industrial Inspection,* pp. 253-263, 1999.

[22] S. Rahmann and N. Canterakis, "Reconstruction of Specular Surfaces Using Polarization Imaging," *Proc. Computer Vision and Pattern Recognition,* pp. 149-155, 2001.

[23] S. Rahmann, "Reconstruction of Quadrics from Two Polarization Views," *Proc. Iberian Pattern Recognition and Image Analysis,* pp. 810-820, 2003.

[24] S.K. Nayar, X. Fang, and T. Boult, "Separation of Reflectance Components Using Colour and Polarization," *Int'l J. Computer Vision,* vol. 21, pp. 163-186, 1997.

[25] S. Umeyama and G. Godin, "Separation of Diffuse and Specular Components of Surface Reflection by Use of Polarization and Statistical Analysis of Images," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 26, pp. 639-647, 2004.

[26] O. Drbohlav and R. Šára, "Unambiguous Determination of Shape from Photometric Stereo with Unknown Light Sources," *Proc. Int'l Conf. Computer Vision,* pp. 581-586, 2001.

[27] J. Clark, E. Trucco, and L.B. Wolff, "Using Light Polarization in Laser Scanning," *Image and Vision Computing J.,* vol. 15, pp. 107-117, 1997.

[28] Y.Y. Schechner, S.G. Narashimhan, and S.K. Nayar, "Polarization-Based Vision through Haze," *Applied Optics,* vol. 42, pp. 511-525, 2003.

[29] Y.Y. Schechner and N. Karpel, "Clear Underwater Vision," *Proc. Computer Vision and Pattern Recognition,* pp. 536-543, 2004.

[30] T. Shibata, T. Takahashi, D. Miyazaki, Y. Sato, and K. Ikeuchi, "Creating Photorealistic Virtual Model with Polarization Based Vision System," *Proc. SPIE Int'l Symp. Optics and Photonics,* vol. 5888, pp. 25-35, 2005.

[31] L.B. Wolff, "Polarization Vision: A New Sensory Approach to Image Understanding," *Image and Vision Computing J.,* vol. 15, pp. 81-93, 1997.

[32] D. Miyazaki, N. Takashima, A. Yoshida, E. Harashima, and K. Ikeuchi, "Polarization-Based Shape Estimation of Transparent Objects by Using Raytracing and PLZT Camera," *Proc. SPIE Int'l Symp. Optics and Photonics,* vol. 5888, pp. 1-14, 2005.

[33] J.J. Koenderink and A.J. van Doorn, "Surface Shape and Curvature Scales," *Image and Vision Computing J.,* vol. 10, pp. 557-565, 1992.

[34] D.A. Forsyth and J. Ponce, *Computer Vision, A Modern Approach.* Prentice Hall, 2003.

[35] S. Kotz and S. Nadarajah, *Extreme Value Distributions: Theory and Applications.* Imperial College Press, 2000.

[36] L.B. Wolff, "Diffuse-Reflectance Model for Smooth Dielectric Surfaces," *Optical Soc. Am. J. A,* vol. 11, pp. 2956-2968, 1994.

[37] A.D.J. Cross and E.R. Hancock, "Graph Matching with a Dual-Step EM Algorithm," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 20, pp. 1236-1253, 1998.

[38] H. Chui and A. Rangarajan, "A New Point Matching Algorithm for Non-Rigid Registration," *Computer Vision and Image Understanding,* vol. 89, pp. 114-141, 2003.

[39] S. Gold, A. Rangarajan, C.P. Lu, S. Pappu, and E. Mjolsness, "New Algorithms for 2D and 3D Point Matching: Pose Estimation and Correspondence," *Pattern Recognition,* vol. 31, pp. 1019-1031, 1998.

[40] P.J. Besl and N.D. McKay, "A Method for Registration of 3D Shapes," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 14, pp. 239-256, 1992.

[41] R.J. Woodham, "Gradient and Curvature from Photometric Stereo Including Local Confidence Estimation," *Optical Soc. Am. J. A,* vol. 11, pp. 3050-3068, 1994.

[42] J.C. Lagarias, J.A. Reeds, M.H. Wright, and P.E. Wright, "Convergence Properties of the Nelder-Mead Simplex Method in Low Dimensions," *SIAM J. Optimization,* vol. 9, pp. 112-147, 1998.

[43] F.N. Fritsch and R.E. Carlson, "Monotone Piecewise Cubic Interpolation," *SIAM J. Numerical Analysis,* vol. 17, pp. 238-246, 1980.

Gary A. Atkinson received the undergraduate degree in physics from the University of Nottingham in 2003, with a final year project on noise reduction in audio signals. He completed the PhD degree at the University of York in 2007, under the supervision of Edwin Hancock. His research interests include shape recovery algorithms and reflectance function estimation for computer vision. Most of his work involves the exploitation of the polarizing properties of reflection from surfaces. He has published 10 papers in refereed journals and conference proceedings during his PhD research. He is a member of the British Machine Vision Association, the Institute of Physics, and he is a student member of the IEEE.

Edwin R. Hancock received the BSc degree in physics, with honors, in 1977 and the PhD degree in the area of high-energy physics in 1981 from the University of Durham. Following this, he worked for 10 years as researcher in the fields of high-energy nuclear physics and pattern recognition at the Rutherford-Appleton Laboratory (now the Central Research Laboratory of the Research Councils). During this period, he also held adjunct teaching posts at the University of Surrey and the Open University. In 1991, he moved to the University of York as a lecturer in the Department of Computer Science. He was promoted to senior lecturer in 1997 and to reader in 1998. In 1998, he was appointed to as chair of Computer Vision. He now leads a group of some 20 faculty, research staff, and PhD students working in the areas of computer vision and pattern recognition. His main research interests are in the use of optimization and probabilistic methods for high and intermediate level vision. He is also interested in the methodology of structural and statistical pattern recognition. He is currently working on graph matching, shape-from-X, image databases, and statistical learning theory. His work has found applications in areas such as radar terrain analysis, seismic section analysis, remote sensing, and medical imaging. He has published about 120 journal papers and 450 refereed conference publications. He was awarded the Pattern Recognition Society medal in 1991 and an outstanding paper award in 1997 by the journal *Pattern Recognition*. In 1998, he became a fellow of the International Association for Pattern Recognition. He is also a fellow of the Institute of Physics, the Institute of Engineering and Technology, and the British Computer Society. He has been a member of the editorial boards of the journals *IEEE Transactions on Pattern Analysis and Machine Intelligence* and *Pattern Recognition*. He has also been a guest editor for special editions of the journals *Image and Vision Computing* and *Pattern Recognition*. In 2006, he was appointed as the founding editor-in-chief of the *IET Computer Vision Journal*. He has been on the program committees for numerous national and international meetings. In 1997, with Marcello Pelillo, he established a new series of international meetings on energy minimization methods in computer vision and pattern recognition.

▷ **For more information on this or any other computing topic, please visit our Digital Library at** www.computer.org/publications/dlib.