The Information Available to a Moving Observer on Shape with Unknown, Isotropic BRDFs

Manmohan Chandraker

Abstract—Psychophysical studies show motion cues inform about shape even with unknown reflectance. Recent works in computer vision have considered shape recovery for an object of unknown BRDF using light source or object motions. This paper proposes a theory that addresses the remaining problem of determining shape from the (small or differential) motion of the camera, for unknown isotropic BRDFs. Our theory derives a differential stereo relation that relates camera motion to surface depth, which generalizes traditional Lambertian assumptions. Under orthographic projection, we show differential stereo may not determine shape for general BRDFs, but suffices to yield an invariant for several restricted (still unknown) BRDFs exhibited by common materials. For the perspective case, we show that differential stereo yields the surface depth for unknown isotropic BRDF and unknown directional lighting, while additional constraints are obtained with restrictions on the BRDF or lighting. The limits imposed by our theory are intrinsic to the shape recovery problem and independent of choice of reconstruction method. We also illustrate trends shared by theories on shape from differential motion of light source, object or camera, to relate the hardness of surface reconstruction to the complexity of imaging setup.

Index Terms—Surface reconstruction, general BRDF, multiview stereo, differential theory

1 Introduction

Image formation is an outcome of the interaction between shape, lighting and camera, governed by the reflectance of the underlying material. Motion of the object, light source or camera are important cues for recovering object shape from images. Each of those cues have been extensively studied in computer vision, under the umbrellas of optical flow for object motion [22], [25], photometric stereo for light source motion [40] and multiview stereo for motion of the camera [36]. Due to the complex and often unknown nature of the bidirectional reflectance distribution function (BRDF) that determines material behavior, simplifying assumptions like brightness constancy and Lambertian BRDF are often employed. Recent years have seen the rapid progress of robust and efficient algorithms for multiview stereo based on diffuse photoconsistency [17], [21], or optical flow based on brightness constancy [2]. However, even common materials such as metals, plastics or fabrics do not reflect light in a diffuse manner consistent with those assumptions [26], [31].

On the other hand, it is everyday experience and established by psychophysical studies that complex reflectance does not impede shape perception [5], [37]. Indeed, non-diffuse material behavior might aid perception of shape from disparity induced by motion [28], [32]. Recent studies also lend empirical support to the notion that perception of shape even with complex material and illumination is easier than suggested by prior computational frameworks [15]. Thus, it is natural for us to consider a theoretical question: *what does motion reveal about unknown shape and material?* Surprisingly, it can be shown that even from a computational viewpoint, shape recovery is possible even with complex, unknown material behavior.

Indeed, recent works in computer vision have shown that information about object shape may be obtained from differential motion of the light source [10] or the object [12], even with unknown BRDF. This paper develops the theoretical framework to solve the remaining problem of characterizing shape recovery for unknown BRDFs, using differential motion of the camera. In subsequent work [8], we show that material properties may also be recovered using motion cues.

The problem of shape recovery from camera motion is akin to a multiview stereo setting, where the goal is to determine dense correspondence across two or more views. Given differential motion between a pair of views, it is equivalent to determining the motion field induced in the image plane. In Section 3 of this paper, we derive a physically valid differential stereo relation that relates surface depth to camera motion, while accounting for general material behavior in the form of an isotropic BRDF. Diffuse photoconsistency of traditional Lambertian stereo follows as a special case. The linearity of differentiation ensures a natural decoupling between shape and BRDF terms in the differential stereo relation. Subsequently, it can be shown that considering a sequence of motions allows completely eliminating the BRDF dependence of differential stereo.

The mathematical basis for differential stereo is outwardly similar to differential flow for object motion [12]. However, while BRDFs are considered black-box functions in [12], our analysis explicitly considers the angular dependencies of isotropic BRDFs to derive additional insights. A particular benefit is to show that ambiguities exist for the case of camera motion, which render shape recovery

• *The author is with NEC Laboratories America, at Cupertino, CA 95014. E-mail: manu@nec-labs.com.*

Manuscript received 14 Nov. 2014; revised 2 Sept. 2015; accepted 9 Sept. 2015. Date of publication 22 Sept. 2015; date of current version 10 June 2016. Recommended for acceptance by A. Martinez, R. Basri, R. Vidal, and C. Fermuller.
For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org, and reference the Digital Object Identifier below.
Digital Object Identifier no. 10.1109/TPAMI.2015.2481415

TABLE 1
A Unified Look at Frameworks on General BRDF Shape Recovery from Differential Motions of Light Source, Object or Camera

Motion	Camera	Light	BRDF	#Motions	Surface Constraint	Shape Recovery	Theory
Light	Orth.	Known	Lambertian	2	Lin. eqns. on ∇z (Photo. stereo)	Gradient	[40]
Light	Orth.	Unknown	General, unknown	2	Linear PDE	Level cur. + $\|\nabla z\|$ isocontours	[10]
Object	Orth.	Colocated	Uni-angle, unknown	2	Homog. quasilin. PDE	Level curves	[12]
Object	Orth.	Unknown	Full, unknown	3	Inhomog. quasilin. PDE	Char. curves	[12]
Object	Persp.	Brightness constancy		1	Lin. eqn. (optical flow)	Depth	[22], [25]
Object	Persp.	Colocated	Uni-angle, unknown	3	Lin. eqn. + Homog. lin. PDE	Depth + Gradient	[12]
Object	Persp.	Unknown	Full, unknown	4	Lin. eqn. + Inhomog. lin. PDE	Depth + Gradient	[12]
Camera	Orth.	Unknown	General, unknown	-	Ambiguous for reconstruction		Prop. 3
Camera	Orth.	Unknown	View angle, unknown	2	Homog. quasilin. PDE	Level curves	Prop. 6
Camera	Orth.	Known	Half angle, unknown	2	Inhomog. quasilin. PDE	Char. curves	Prop. 4, 7
Camera	Orth.	Known	$\{\mathbf{s}, \mathbf{v}\}$ plane, unknown	2	Inhomog. quasilin. PDE	Char. curves	Prop. 5, 7
Camera	Persp.	Brightness constancy		1	Lin. eqn. (stereo)	Depth	[36]
Camera	Persp.	Unknown	General, unknown	3	Linear eqn.	Depth	Prop. 8
Camera	Persp.	Known	Half angle, unknown	3	Lin. eqn. + Inhomog. lin. PDE	Depth + Gradient	Prop. 9
Camera	Persp.	Known	$\{\mathbf{s}, \mathbf{v}\}$ plane, unknown	3	Lin. eqn. + Inhomog. lin. PDE	Depth + Gradient	Rem. 2

In each case, PDE invariants are derived which specify precise topological limits on shape recovery. More general BRDFs or imaging setups require greater number of motions to derive the invariant. For comparison, the traditional diffuse equivalents are also shown. Note that $k+1$ images are required for k motions[2].

more difficult. Consequently, for orthographic projection, Section 4 shows a negative result whereby constraints on the shape of a surface with general isotropic BRDF may not be derived using camera motion as a cue. But we show the existence of an invariant for several restricted isotropic BRDFs, exhibited by common materials like plastics, metals, some paints and fabrics. The invariant is characterized as a quasilinear partial differential equation (PDE), which specifies the topological class upto which reconstruction may be performed.

Under perspective projection, Section 5 shows that depth for a surface with unknown isotropic BRDF, under unknown directional lighting, may be obtained using differential stereo relations from a sequence of camera motions. Further, for the above restricted families of BRDFs, we show that an additional constraint on the surface gradient is available. These results substantially generalize Lambertian stereo, since depth information is obtained without assuming diffuse photoconsistency, while a weak assumption on material type yields even richer surface information. The main theoretical results of this paper are summarized in Table 1, within the portion on camera motion.

The limits on surface reconstruction derived in this paper are intrinsic to the shape recovery problem, regardless of choice of reconstruction method. Along with [10], [12], our theories for shape recovery from motion cues also suggest frameworks to address a question central to surface reconstruction: *how does imaging complexity relate to the hardness of shape recovery?* Qualitatively, we show that material or illumination complexity determines the topological class for surface reconstruction. Quantitatively, we show that they determine the minimum imaging budget required for shape recovery. Section 6 discusses relationships between theories on shape recovery with unknown material behavior, using differential motion of the object, light source or camera. We explore shared traits among those theories and their common trends on the hardness of surface reconstruction. This leads to a stratified framework for shape from motion, which is summarized in Table 1.

2 Related Work

Relating shape to intensity variations due to differential motion has a significant history in computer vision, dating to studies in optical flow [22], [25]. The limitations of the Lambertian assumption have been recognized by early works [29], [39].

Multiview stereo algorithms that rely on diffuse photoconsistency also have a long history in computer vision. We refer the reader to [36] for a comprehensive survey. In recent years, algorithms for multiview stereo have demonstrated reconstructions with impressive robustness [17], [21]. To handle drawbacks such as lack of texture or presence of specularities, several extensions have been proposed that include priors from Manhattan constraints [16], [18] or architectural schema [41]. In contrast to such approaches, we explicitly account for BRDF-dependence in image formation for shape recovery under differential motion of the camera.

Reconstruction methods have also been developed to handle non-Lambertian or unknown material behavior in the context of stereo. For instance, Zickler et al. use the Helmholtz reciprocity principle for reconstruction with arbitrary BRDFs [42]. An example-based stereo that uses reference shapes of known geometry and various material types is presented in [38]. Multiview reconstructions for specular surfaces using a voxel carving approach are studied by Bonfort and Sturm [6]. In contrast, this paper explores how an image sequence derived from camera motion informs about shape with unknown isotropic BRDFs, regardless of the reconstruction method.

A few early works in computer vision consider the theoretical extent of shape recovery from motion cues for mirror surfaces. A qualitative description of the pattern of specular highlights induced by motion is presented by Koenderink and van Doorn in [24]. A specular stereo framework for shape recovery is proposed by Blake and Brelstaff in [3], [4], while Zisserman et al. consider the information available to a moving observer in the case of mirror surfaces, by tracking specular highlights in [43]. In similar spirit, our theory precisely characterizes the extent to which shape may be

recovered from observer motion, for the more general problem of unknown, isotropic BRDFs.

In recent years, a few approaches have also been considered for joint recovery of shape and BRDF. Goldman et al. use a set of basis materials in photometric stereo to recover shape and reflectance [19]. Methods that alternatingly optimize over shape and material to recover both under light source motion [1] and natural illumination [33] have been proposed. In contrast, using cues from the relative motion between object and camera proposed by this paper, our subsequent work shows that shape and BRDF may be estimated simultaneously, while also delineating limits on the joint recovery problem [8].

Light source and object motions have also been used to recover shape with unknown BRDFs. A photometric stereo method based on small light source motions is presented by [14], while optical flow has been generalized to more general models by [20], [30]. An isometric relationship between changes in normals and radiance profiles under varying light is used by Sato et al. to recover shape with unknown reflectance [35]. Ray differentials [13] and first-order shading variations [34] have also been studied in computer graphics.

Closely related to this paper are our recent works that derive topological classes up to which reconstruction can be performed for unknown BRDFs, using differential motion of the light source [10] or the object [12]. This paper derives limits on shape recovery using the third cue, namely camera motion. The similarities and differences between the frameworks are explored throughout this paper and summarized in Section 6. A preliminary version of this paper appears as [9].

3 Differential Stereo: General BRDFs

In this section, we state our assumptions and derive the relationship between camera motion and surface depth, for unknown isotropic BRDFs. We also provide some intuition into that relationship, which will be used for subsequent shape recovery results. To facilitate presentation, we will occasionally point the reader to the appendices for details whose deferral does not impact understanding the theory.

3.1 Derivation of the Differential Stereo Relation

3.1.1 Assumptions and Setup

We assume static object and lighting, while the camera undergoes known rigid body motion. Our analysis equivalently considers a fixed camera, with the object and light source undergoing the inverse motion. Unless otherwise stated, all motions are completely general. The object BRDF is unknown, but assumed homogeneous (that is, the same function of incident and exitant angles at all points) and isotropic (that is, expressible as a function of the three angles between the surface normal, light source and viewing directions). We make an assumption that the camera direction is constant over the entire object. This assumption is exact for orthographic cameras, but only an approximation for perspective projection where viewing direction may vary over object dimensions. The approximation is reasonable in practical situations where the camera is not too close to the object, relative to object size. The illumination is assumed directional and distant, with negligible global effects.

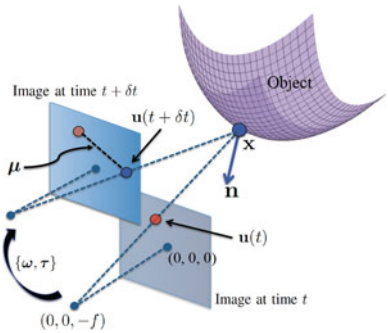


Fig. 1. Illustration of the imaging setup and notations.

Let the focal length of the calibrated camera be f. The camera model is perspective for finite values of f and approaches orthographic as $f \to \infty$. The principal point on the image plane is defined as the origin of the 3D coordinate system, with the camera center at $(0, 0, -f)^\top$. Denoting $\beta = \frac{1}{f}$, a 3D point $\mathbf{x} = (x, y, z)^\top$ is imaged at $\mathbf{u} = (u, v)^\top$, where

$$u = \frac{x}{1+\beta z}, \qquad v = \frac{y}{1+\beta z}. \tag{1}$$

Fig. 1 illustrates the imaging setup.

3.1.2 Motion Field

Using the projection equations in (1), the motion field is given by the differential motion of an image pixel:

$$\boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} = \begin{bmatrix} \dot{u} \\ \dot{v} \end{bmatrix} = \frac{1}{1+\beta z} \begin{bmatrix} -\beta u \dot{z} \\ \dot{y} - \beta v \dot{z} \end{bmatrix}. \tag{2}$$

Suppose the camera undergoes rotation $\widetilde{\mathbf{R}}$ and translation $\widetilde{\boldsymbol{\tau}}$. For rigid body motion, we equivalently assume that the object and light source undergo a rotation $\mathbf{R} = \widetilde{\mathbf{R}}^{-1}$ and translation $\boldsymbol{\tau} = \widetilde{\mathbf{R}}^{-1}\widetilde{\boldsymbol{\tau}}$, with a fixed camera. For differential motion, we approximate $\mathbf{R} \approx \mathbf{I} + [\boldsymbol{\omega}]_\times$, where $\boldsymbol{\omega} = (\omega_1, \omega_2, \omega_3)^\top$. Then, for point $\mathbf{x}$ on the object, $\dot{\mathbf{x}} = \boldsymbol{\omega} \times \mathbf{x} + \boldsymbol{\tau}$, that is,

$$\begin{aligned} \dot{x} &= \omega_2 z - \omega_3 y + \tau_1 = \omega_2 z - \omega_3(1+\beta z)v + \tau_1, \\ \dot{y} &= \omega_3 x - \omega_1 z + \tau_2 = -\omega_1 z + \omega_3(1+\beta z)u + \tau_2, \\ \dot{z} &= \omega_1 y - \omega_2 x + \tau_3 = (1+\beta z)(\omega_1 v - \omega_2 u) + \tau_3. \end{aligned} \tag{3}$$

In the perspective case, the motion field is given by,

$$\boldsymbol{\mu} = \left(\alpha_1 + \frac{\alpha_2 + \omega_2 z}{1+\beta z},\ \alpha_3 + \frac{\alpha_4 - \omega_1 z}{1+\beta z} \right)^\top, \tag{4}$$

where $\alpha_1 = \omega_2 \beta u^2 - \omega_1 \beta u v - \omega_3 v, \alpha_2 = \tau_1 - \beta u \tau_3,\ \alpha_3 = -\omega_1 \beta v^2 + \omega_2 \beta u v + \omega_3 u$ and $\alpha_4 = \tau_2 - \beta v \tau_3$. Under orthography, $\beta \to 0$, thus, the motion field is

$$\boldsymbol{\mu} = (\alpha_5 + \omega_2 z,\ \alpha_6 - \omega_1 z)^\top, \tag{5}$$

where $\alpha_5 = \tau_1 - \omega_3 v$ and $\alpha_6 = \tau_2 + \omega_3 u$.

3.1.3 Image Formation

Let $\mathbf{s}$ be the unit vector denoting light source direction and $\mathbf{v} = (0, 0, -1)^\top$ be the camera direction. For a 3D point $\mathbf{x} = (x, y, z(x, y))^\top$ on the object, the unit surface normal is

$$\mathbf{n} = (n_1, n_2, n_3)^\top = \frac{(z_x, z_y, -1)^\top}{\sqrt{z_x^2 + z_y^2 + 1}}, \tag{6}$$

where $\nabla z = (z_x, z_y)^\top$ is the surface gradient. In the monochromatic case, for a homogeneous isotropic BRDF ρ, with distant light source, the image intensity observed at pixel $\mathbf{u}$ corresponding to a 3D point $\mathbf{x}$ is

$$I(\mathbf{u}, t) = \sigma(\mathbf{x})\rho(\mathbf{x}, \mathbf{n}, \mathbf{s}, \mathbf{v}), \tag{7}$$

where σ is the albedo and the cosine fall-off is absorbed in ρ. This is a reasonable imaging model that subsumes traditional ones like Lambertian and allows general isotropic BRDFs modulated by spatially varying albedo. While the geometry variables above are expressed in a global coordinate system, we note that they may be equivalently interpreted in intrinsic surface coordinates, as Appendix A. We do not make any assumptions on the functional form of ρ, except smoothness. In the dichromatic case, the diffuse component of reflection is made explicit, leading to an image formation model given by

$$I(\mathbf{u}, t) = \sigma(\mathbf{x})\mathbf{n}^\top\mathbf{s} + \rho(\mathbf{x}, \mathbf{n}, \mathbf{s}, \mathbf{v}). \tag{8}$$

The theory of this paper is applicable to both the above models.

3.1.4 Differential Stereo Relation

We will now provide an intuition into how change in image intensities may be related to shape. Consider the monochromatic case. Taking the *total* derivative on both sides of (7), we get

$$I_u\dot{u} + I_v\dot{v} + I_t = \sigma\frac{d}{dt}\rho(\mathbf{x}, \mathbf{n}, \mathbf{s}, \mathbf{v}) + \rho\frac{d\sigma}{dt}. \tag{9}$$

Intuitively, since σ is intrinsically defined on the surface coordinates, its total derivative vanishes. For the rigid body motion we consider, the object and lighting undergo a differential motion given by $\{\boldsymbol{\omega}, \boldsymbol{\tau}\}$ relative to the camera. The changes in surface normal and distant light directions is given by $\dot{\mathbf{n}} = \boldsymbol{\omega} \times \mathbf{n}$ and $\dot{\mathbf{s}} = \boldsymbol{\omega} \times \mathbf{s}$, respectively, while the camera direction remains unchanged. Using chain rule differentiation and noting that $\boldsymbol{\mu} = (\dot{u}, \dot{v})^\top$ is the motion field, we have from (9):

$$\begin{aligned}(\nabla_{\mathbf{u}} I)^\top\boldsymbol{\mu} + I_t = \sigma[&(\nabla_{\mathbf{x}}\rho)^\top\boldsymbol{\nu} + (\nabla_{\mathbf{n}}\rho)^\top(\boldsymbol{\omega} \times \mathbf{n}) \\ &+ (\nabla_{\mathbf{s}}\rho)^\top(\boldsymbol{\omega} \times \mathbf{s})],\end{aligned} \tag{10}$$

where $\boldsymbol{\nu} = \dot{\mathbf{x}}$ is the linear velocity. We refer the reader to Appendix A for a rigorous derivation.

Dividing the two sides of (10) with those of (7) and using the notation $E = \log I$ and $\rho' = \log\rho$, we get

$$\boxed{(\nabla_{\mathbf{u}} E)^\top\boldsymbol{\mu} + E_t = \boldsymbol{\pi}^\top\boldsymbol{\omega} + (\nabla_{\mathbf{x}}\rho')^\top\boldsymbol{\nu},} \tag{11}$$

where we use the identities $(\nabla_{\mathbf{a}}\rho)^\top(\boldsymbol{\omega} \times \mathbf{a}) = (\mathbf{a} \times \nabla_{\mathbf{a}}\rho)^\top\boldsymbol{\omega}$ and $\nabla_{\mathbf{a}}\log\rho = \frac{1}{\rho}\nabla_{\mathbf{a}}\rho$, for $\mathbf{a} \in \mathbb{R}^3$, having defined

$$\boldsymbol{\pi} = \mathbf{n} \times \nabla_{\mathbf{n}}\log\rho + \mathbf{s} \times \nabla_{\mathbf{s}}\log\rho. \tag{12}$$

We call (11) the *differential stereo relation*.

For the dichromatic case, similar steps may again be followed. Taking the total derivative of both sides of (8), we obtain

$$I_u\dot{u} + I_v\dot{v} + I_t = \sigma\frac{d}{dt}\mathbf{n}^\top\mathbf{s} + \mathbf{n}^\top\mathbf{s}\frac{d\sigma}{dt} + \frac{d}{dt}\rho(\mathbf{x}, \mathbf{n}, \mathbf{s}, \mathbf{v}). \tag{13}$$

We observe that in the case of camera motion, the relative orientation between the surface normal and light source remains unchanged (that is, $\mathbf{n}^\top\mathbf{s}$ is constant), while the total derivative of the intrinsic property albedo again vanishes. Again, from the chain rule of differentiation, we obtain a differential stereo relation of the exact same form as (11), but without the need to take logarithms. Our approach will be to completely eliminate the BRDF terms from a sequence of differential stereo relations, thus, our theory is equally applicable for both the monochromatic and dichromatic image formation models.

The main result of this work is that differential stereo relations of the form (11) obtained from motion cues yield information on surface depth. To provide an intuitive description, we will work with an approximate version of the differential stereo relation in the next few sections. In particular, we consider distant lighting and homogeneous reflectance under the assumption that the viewpoint does not vary significantly over object dimensions. Note that the entity $\nabla_{\mathbf{x}}\rho$ encodes change in BRDF with respect to object dimensions, so will be larger for higher surface curvature and greater sharpness of the BRDF. For surface geometries and BRDFs whose curvatures are not high, with the object dimension small compared to its distance, it is reasonable to assume that $\nabla_{\mathbf{x}}\rho$ is small. Thus, we have an *approximate differential stereo relation*:

$$(\nabla_{\mathbf{u}} E)^\top\boldsymbol{\mu} + E_t = \boldsymbol{\pi}^\top\boldsymbol{\omega}. \tag{14}$$

Shape recovery results derived for the approximate relation are also available in similar forms for the exact relation without neglecting $\nabla_{\mathbf{x}}\rho$ in Appendix D, with more motions required corresponding to its more complex form. In the following, we will drop the term approximate unless required for clarity.

3.2 A Property of the BRDF-Dependent Entity $\boldsymbol{\pi}$

For object motion, a differential flow relation of similar form as (14) is derived in [12]. Now we derive additional insights by making a crucial departure from the analysis of [12]. Namely, instead of treating isotropic BRDFs as black-box functions, we consider angular dependencies between the normal, light source and camera directions.

The entity $\boldsymbol{\pi}$ defined in (12) is central to our theory. Its practical significance is that any shape recovery method that seeks invariance to material behavior must either accurately model $\boldsymbol{\pi}$, or eliminate it. Our work adopts the latter approach. For isotropic BRDFs, we make the following observation:

Proposition 1. *The BRDF-dependent entity $\boldsymbol{\pi}$ is a 2D vector in the principal plane of the camera.*

Proof. An isotropic BRDF depends on the three angles between normal, camera and light directions, so we may write

$$\tilde{\rho}(\mathbf{n}^\top\mathbf{s}, \mathbf{s}^\top\mathbf{v}, \mathbf{n}^\top\mathbf{v}) = \log\rho(\mathbf{n}, \mathbf{s}, \mathbf{v}). \tag{15}$$

Denote $\theta = \mathbf{n}^\top \mathbf{s}$, $\phi = \mathbf{s}^\top \mathbf{v}$ and $\psi = \mathbf{n}^\top \mathbf{v}$. Then, applying chain-rule differentiation, we may write (12) as

$$\begin{aligned}\boldsymbol{\pi} &= \mathbf{n} \times \nabla_{\mathbf{n}} \tilde{\rho} + \mathbf{s} \times \nabla_{\mathbf{s}} \tilde{\rho} \\ &= \tilde{\rho}_\theta (\mathbf{n} \times \mathbf{s}) + \tilde{\rho}_\psi (\mathbf{n} \times \mathbf{v}) + \tilde{\rho}_\theta (\mathbf{s} \times \mathbf{n}) + \tilde{\rho}_\phi (\mathbf{s} \times \mathbf{v}) \\ &= \tilde{\rho}_\psi (\mathbf{n} \times \mathbf{v}) + \tilde{\rho}_\phi (\mathbf{s} \times \mathbf{v}).\end{aligned} \tag{16}$$

From the form of $\boldsymbol{\pi}$ in (16), it is evident that

$$\boldsymbol{\pi}^\top \mathbf{v} = 0. \tag{17}$$

For our assumed setup and coordinate system, $\mathbf{v} = (0, 0, -1)^\top$. Thus, $\pi_3 = 0$ and it follows that the BRDF-dependent entity $\boldsymbol{\pi} = (\pi_1, \pi_2, 0)^\top$ lies on the principal plane of the camera. □

This is an important result that limits the extent to which shape may be recovered from differential stereo. The following sections explore the precise nature of those limits.

4 Orthographic Projection

Estimating the motion field, $\boldsymbol{\mu}$, is equivalent to determining dense correspondence and thereby, object shape. We now consider shape recovery with unknown BRDF under orthographic projection, using a sequence of differential motions.

4.1 Rank Deficiency and Depth Ambiguities

It is clear that just one differential motion of the camera is insufficient to extract depth, since (14) is a linear relation in the multiple unknowns $\{\boldsymbol{\mu}, \boldsymbol{\pi}\}$. Consequently, we consider a sequence of differential motions. We start by observing that, in the case of orthography, a rank deficiency similar to the case of object motion derived in [12] exists for camera motion too.

Under orthography, the motion field $\boldsymbol{\mu}$ is given by (5). Noting that μ_1 and μ_2 are linear in z, we observe that the differential stereo relation of (14) reduces to:

$$pz + q = \boldsymbol{\omega}^\top \boldsymbol{\pi}, \tag{18}$$

where, using (5), p and q are known entities given by

$$p = \omega_2 E_u - \omega_1 E_v, \tag{19}$$

$$q = \alpha_5 E_u + \alpha_6 E_v + E_t. \tag{20}$$

Consider $m > 0$ differential motions of the camera about a base position, given by $\{\boldsymbol{\omega}^i, \boldsymbol{\tau}^i\}$, for $i = 1, \ldots, m$. To avoid degeneracies, we assume that the motions are general, with rotations $\boldsymbol{\omega}^i$ that are independent when $m < 3$ and span $\mathbb{R}^3$ when $m \geq 3$. Let $E^0 = \log I^0$ be the logarithm of the base image, with E^i the log-image for each motion $\{\boldsymbol{\omega}^i, \boldsymbol{\tau}^i\}$. Note that the spatial gradient of the image is independent of motion and corresponds to derivative of E_0 with respect to $\mathbf{u}$, which we denote as $\nabla_{\mathbf{u}} E = (E_u, E_v)^\top$. The temporal derivative, E_t, as well as p and q, depend on the motion. We again note that all subsequent analysis holds for the case of dichromatic image formation too, but with $E^i = I^i$ (no logarithms).

To recover the unknown depth z, an initial approach may consider $m \geq 3$ relations of the form (14) as a linear system:

$$\tilde{\mathbf{A}} \begin{bmatrix} z \\ \pi_1 \\ \pi_2 \end{bmatrix} = \mathbf{q}, \text{ with } \tilde{\mathbf{A}} = \begin{bmatrix} -p^1 & \omega_1^1 & \omega_2^1 \\ \vdots & & \vdots \\ -p^m & \omega_1^m & \omega_2^m \end{bmatrix}, \tag{21}$$

where $\mathbf{q} = (q^1, \ldots, q^m)^\top$. Note that Proposition 1 allows us to drop any dependence on ω_3, since $\pi_3 = 0$ is not an unknown. But observe the form of $p^i = \omega_2^i E_u - \omega_1^i E_v$ from (19), which makes $\tilde{\mathbf{A}}$ rank-deficient. Thus, we have shown:

Proposition 2. *Under orthographic projection, surface depth under unknown BRDF may not be unambiguously recovered using solely camera motion as the cue.*

While depth cannot be directly recovered from differential stereo under orthography, a natural next step is to consider the possibility of any constraints on the depth. Indeed, considering a sequence of differential flow relations for the case of object motion in [12], a quasilinear PDE constraint on surface depth is obtained. However, Proposition 1 makes it challenging for the case of camera motion. To see this, we note that the null vector of the rank-deficient matrix $\tilde{\mathbf{A}}$ is $(1, -E_v, E_u)^\top$. With $\tilde{\mathbf{A}}^+$ denoting the Moore-Penrose pseudoinverse of $\tilde{\mathbf{A}}$, for any $k \neq 0$, we have a parameterized solution to (21):

$$\begin{bmatrix} z \\ \pi_1 \\ \pi_2 \end{bmatrix} = \boldsymbol{\gamma} + k \begin{bmatrix} 1 \\ -E_v \\ E_u \end{bmatrix}, \tag{22}$$

where $\boldsymbol{\gamma} = (\gamma_1, \gamma_2, \gamma_3)^\top = \tilde{\mathbf{A}}^+ \mathbf{q}$. From the first equation in the above system, we have $k = z - \gamma_1$. Thereby, we get the following two relations between z and $\boldsymbol{\pi}$:

$$\pi_1 = (\gamma_2 + E_v \gamma_1) - E_v z, \tag{23}$$

$$\pi_2 = (\gamma_3 - E_u \gamma_1) + E_u z. \tag{24}$$

Now, any relationship between π_1 and π_2 gives a constraint on the depth, z. But from Proposition 1, $\boldsymbol{\pi}$ is an arbitrary vector in the principal plane. That is, from (16), π_1 and π_2 depend on two unknown BRDF derivatives $\tilde{\rho}_\psi$ and $\tilde{\rho}_\phi$. It follows that without imposing any external constraint on $\tilde{\rho}_\psi$ and $\tilde{\rho}_\phi$, one may not derive a constraint on surface depth. Thus, we state:

Proposition 3. *Under orthographic projection, for unknown isotropic BRDF, an unambiguous constraint on surface depth may not be derived using solely camera motion as the cue.*

The above is an example of the comparative limits on shape recovery with object or camera motion. For object motion, only the relative motion between camera and object must be accounted. Thus, the definition of $\boldsymbol{\pi} = \mathbf{n} \times \nabla_{\mathbf{n}} \log \rho$ in [12] depends only on the BRDF-derivative with respect to surface normal. For camera motion, both object and lighting move relative to the camera. Additional dependence of $\boldsymbol{\pi}$ in (12) on BRDF-derivative with respect to lighting makes it indeterminate without further restrictions on BRDF or lighting.

While Proposition 3 is a negative result, its development provides valuable insight. An $m \times 3$ rank-deficient matrix $\tilde{\mathbf{A}}$

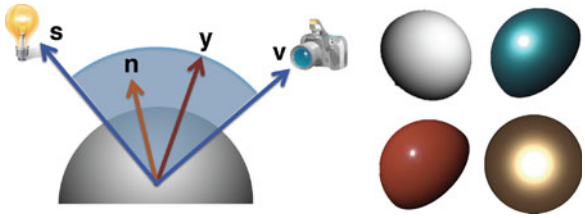


Fig. 2. (a) Constraints on surface estimation in the orthographic case may be derived for BRDFs that depend on a reflection angle $\mathbf{y}$ in the plane of the light source and the camera. (b) BRDF for common material types, such as Lambertian (top left), metals (top right), plastics (bottom left) and any isotropic BRDF under colocated illumination (bottom right) satisfy this property.

constructed using m differential motions, for any $m \geq 2$, determines $\boldsymbol{\gamma} = \tilde{\mathbf{A}}^{+}\mathbf{q}$. From (17), (23) and (24), it follows:

Corollary 1. *Under orthography, two differential motions of the camera suffice to constrain $\boldsymbol{\pi}$ to a linear relation in z.*

This fact will now be used, along with some restrictions on the BRDF, to derive BRDF-invariant constraints on depth.

4.2 BRDF-Invariance for Certain Material Types

The result of Proposition 3 immediately suggests a possibility to constrain z: consider a BRDF whose dependence on ψ and ϕ is restricted in a way that introduces a constraint on $\boldsymbol{\pi}$. In this section, we show that it is possible to do so for an important class of restricted BRDFs—which depend on the angle between the surface normal and an arbitrary direction, $\mathbf{y}$, in the plane of the light source and the camera (see Fig. 2). Many common materials, such as those with diffuse reflectance, or metals and plastics that predominantly reflect light about the half-angle, can be represented by this class of BRDFs. Note that even with this restriction on reflection direction, we continue to assume that the functional form of the BRDF, $\rho(\cdot)$, remains unknown.

We first consider special cases where the direction $\mathbf{y}$ coincides with the camera direction or the half-angle.

4.2.1 BRDFs Dependent on View Angle

Some reflectances depend on the angles subtended by the normal on the source and view directions. Such BRDFs can explain the darkening near image edges for materials like fabrics. Another well-known example is the Minnaert BRDF for lunar reflectance [27]. In such cases, we may define

$$\bar{\rho}(\mathbf{n}^{\top}\mathbf{s}, \mathbf{n}^{\top}\mathbf{v}) = \log \rho(\mathbf{n}, \mathbf{s}, \mathbf{v}). \tag{25}$$

Again denoting $\theta = \mathbf{n}^{\top}\mathbf{s}$ and $\psi = \mathbf{n}^{\top}\mathbf{v}$, we get from (12):

$$\begin{aligned} \boldsymbol{\pi} &= \mathbf{n} \times \nabla_{\mathbf{n}}\bar{\rho} + \mathbf{s} \times \nabla_{\mathbf{s}}\bar{\rho} \\ &= \mathbf{n} \times (\bar{\rho}_{\theta}\mathbf{s} + \bar{\rho}_{\psi}\mathbf{v}) + \mathbf{s} \times \bar{\rho}_{\theta}\mathbf{n} = \bar{\rho}_{\psi}\mathbf{n} \times \mathbf{v}. \end{aligned} \tag{26}$$

Noting that $\mathbf{n} \times \mathbf{v} = (-n_2, n_1, 0)^{\top}$, one may eliminate the BRDF-dependent term $\bar{\rho}_{\psi}$ using (23) and (24) to obtain a relationship between depths and normals:

$$\frac{\pi_1}{\pi_2} = \frac{-n_2}{n_1} = \frac{(\gamma_2 + E_v\gamma_1) - E_v z}{(\gamma_3 - E_u\gamma_1) + E_u z}. \tag{27}$$

Using (6) to relate the normal to the gradient, this reduces to

$$[(\gamma_2 + E_v\gamma_1) - E_v z]z_x + [(\gamma_3 - E_u\gamma_1) + E_u z]z_y = 0, \tag{28}$$

which is a constraint on surface depth and gradient that is independent of both the BRDF and lighting. We note that $m \geq 2$ differential motions of the camera suffice to determine $\boldsymbol{\gamma}$ from (22) and yield the constraint in (28). Thus, we state:

Remark 1. Under orthography, for a BRDF of unknown functional form that depends on light and view directions, two differential motions of the camera suffice to yield a constraint on surface depth independent of BRDF and lighting.

4.2.2 BRDFs Dependent on Half-Angle

For many common materials like metals or plastics, it is reasonable to assume that reflectance depends on the angle between the surface normal and the half-angle between the source and view directions. For a surface of such material type, we can show that a sequence of differential stereo relations yields a BRDF-invariant constraint on surface depth. For this case, we assume a known light source direction.

Proposition 4. *Under orthographic projection, for a BRDF of unknown functional form that depends on known light and half-angle directions, two differential motions of the camera suffice to yield a BRDF-invariant constraint on surface depth.*

Proof. For a BRDF dependent on half-angle $\mathbf{h}$, we define

$$\bar{\rho}(\mathbf{n}^{\top}\mathbf{s}, \mathbf{n}^{\top}\mathbf{h}) = \log \rho(\mathbf{n}, \mathbf{s}, \mathbf{v}), \text{ where } \mathbf{h} = \frac{\mathbf{s} + \mathbf{v}}{\|\mathbf{s} + \mathbf{v}\|}. \tag{29}$$

The definition of $\boldsymbol{\pi}$ in (12) may now be rewritten as

$$\boldsymbol{\pi} = \mathbf{n} \times \nabla_{\mathbf{n}}\bar{\rho} + \mathbf{s} \times \nabla_{\mathbf{s}}\bar{\rho}. \tag{30}$$

We again denote $\theta = \mathbf{n}^{\top}\mathbf{s}$, $\phi = \mathbf{s}^{\top}\mathbf{v}$ and define $\eta = \mathbf{n}^{\top}\mathbf{h}$. Then, using the definition of $\mathbf{h}$ in (29), we apply chain-rule differentiation to obtain:

$$\mathbf{n} \times \nabla_{\mathbf{n}}\bar{\rho} = \bar{\rho}_{\theta}(\mathbf{n} \times \mathbf{s}) + \bar{\rho}_{\eta}\frac{\mathbf{n} \times \mathbf{s}}{\|\mathbf{s} + \mathbf{v}\|} + \bar{\rho}_{\eta}\frac{\mathbf{n} \times \mathbf{v}}{\|\mathbf{s} + \mathbf{v}\|}, \tag{31}$$

$$\mathbf{s} \times \nabla_{\mathbf{s}}\bar{\rho} = \bar{\rho}_{\theta}(\mathbf{s} \times \mathbf{n}) + \bar{\rho}_{\eta}\frac{\mathbf{s} \times \mathbf{n}}{\|\mathbf{s} + \mathbf{v}\|} - \bar{\rho}_{\eta}\frac{(\mathbf{n}^{\top}\mathbf{h})\mathbf{s} \times \mathbf{v}}{\|\mathbf{s} + \mathbf{v}\|^2}. \tag{32}$$

Adding (31) and (32), the first two terms in each cancel out and we obtain using (30):

$$\boldsymbol{\pi} = \bar{\rho}_{\eta}\left[\frac{\mathbf{n} \times \mathbf{v}}{\|\mathbf{s} + \mathbf{v}\|} - (\mathbf{n}^{\top}\mathbf{h})\frac{\mathbf{s} \times \mathbf{v}}{\|\mathbf{s} + \mathbf{v}\|^2}\right]. \tag{33}$$

Intuitively, the symmetry of $\boldsymbol{\pi}$ in (30) with respect to $\mathbf{n}$ and $\mathbf{s}$ means it is independent of ρ_{θ}. Noting that $\mathbf{v} = (0, 0, -1)^{\top}$ and $\|\mathbf{s} + \mathbf{v}\| = \sqrt{2(1 + \phi)}$, we obtain from (33):

$$\frac{\pi_1}{\pi_2} = \frac{\mathbf{a}^{\top}\mathbf{n}}{\mathbf{b}^{\top}\mathbf{n}} = \frac{a_1 z_x + a_2 z_y - a_3}{b_1 z_x + b_2 z_y - b_3}, \tag{34}$$

where we have used the relationship between surface normal and gradient given by (6) and denoted

$$a_1 = s_2 h_1,\ a_2 = s_2 h_2 - \sqrt{2(1+\phi)},\ a_3 = s_2 h_3, \tag{35}$$

$$b_1 = \sqrt{2(1+\phi)} - s_1 h_1,\ b_2 = -s_1 h_2,\ b_3 = -s_1 h_3. \tag{36}$$

From Corollary 1, we also have that $m \geq 2$ differential motions of the camera suffice to restrict $\boldsymbol{\pi}$ to a linear relation in z. In particular, from (23) and (24), we have:

$$\frac{\pi_1}{\pi_2} = \frac{(\gamma_2 + E_v\gamma_1) - E_v z}{(\gamma_3 - E_u\gamma_1) + E_u z}, \tag{37}$$

where $\boldsymbol{\gamma} = (\gamma_1, \gamma_2, \gamma_3)^\top$ is defined by (22). Thus, from (34) and (37), we have obtained

$$\frac{a_1 z_x + a_2 z_y - a_3}{b_1 z_x + b_2 z_y - b_3} = \frac{(\gamma_2 + E_v\gamma_1) - E_v z}{(\gamma_3 - E_u\gamma_1) + E_u z}, \tag{38}$$

which may be rewritten as

$$(\lambda_1 + \lambda_2 z) z_x + (\lambda_3 + \lambda_4) z_y + \lambda_5 = 0, \tag{39}$$

where

$$\begin{aligned}
&\lambda_1 = a_1(\gamma_3 - E_u\gamma_1) - b_1(\gamma_2 + E_v\gamma_1),\ \lambda_2 = a_1 E_u + b_1 E_v,\\
&\lambda_3 = a_2(\gamma_3 - E_u\gamma_1) - b_2(\gamma_2 + E_v\gamma_1),\ \lambda_4 = a_2 E_u + b_2 E_v,\\
&\lambda_5 = -a_3(\gamma_3 - E_u\gamma_1) + b_3(\gamma_2 + E_v\gamma_1).
\end{aligned} \tag{40}$$

We note that the λ_i above are independent of $\bar{\rho}$, thus, (39) is a BRDF-invariant constraint on surface depth. □

Intuitively, the ambiguity of Proposition 3 exists since one cannot constrain π_1 and π_2 if they depend on two independent unknowns ρ_ϕ and ρ_ψ. Considering an appropriate BRDF, such as a half-angle one, allows expressing $\boldsymbol{\pi}$ in terms of a single unknown ρ_η. The linearity of differentiation eliminates ρ_η to yield a BRDF-invariant constraint on $\boldsymbol{\pi}$. Thus, Proposition 4 can derive an invariant purely in terms of depth and gradient.

Note that Proposition 1 is a basic property of isotropic BRDFs under camera motion. So, it may be verified that it holds true even in the restricted case of half-angle BRDFs, that is, $\boldsymbol{\pi}^\top \mathbf{v} = 0$ even for the $\boldsymbol{\pi}$ defined by (33).

4.2.3 Dependence on Arbitrary Angle in $\{\mathbf{s}, \mathbf{v}\}$-Plane

Recent works on empirical analysis of measured BRDFs show that reflectance functions often depend on the angles the surface normal makes with the light source and another direction in the plane defined by the source and camera directions [11]. In such cases, we may write

$$\log \rho(\mathbf{n}, \mathbf{s}, \mathbf{v}) = \bar{\rho}(\mathbf{n}^\top \mathbf{s}, \mathbf{n}^\top \mathbf{y}),\ \text{with}\ \mathbf{y} = \frac{\mathbf{s} + \kappa\mathbf{v}}{\|\mathbf{s} + \kappa\mathbf{v}\|}, \tag{41}$$

for some $\kappa \in \mathbb{R}$. Note that half-angle BRDFs for materials like plastics and metals considered in Section 4.2.2 are a special case with $\kappa = 1$. The BRDFs considered in Section 4.2.1 may also be considered a limit case with $\kappa \gg 1$. Empirical examples for BRDFs with finite $\kappa \neq 1$ are shown for materials like paints and fabrics in [11]. We may now state:

Proposition 5. *Under orthographic projection, for a BRDF of unknown functional form that depends on light source and an arbitrary direction in the source-view plane, two differential motions of the camera suffice to yield a BRDF-invariant constraint on surface depth.*

Proof. The proof directly generalizes the development in Proposition 4. We refer the reader to Appendix B.1 for the complete algebraic details. We only note here that we obtain:

$$\boldsymbol{\pi} = \frac{\kappa\rho_{\mathbf{n}^\top\mathbf{y}}}{1 + \kappa^2 + 2\kappa\phi}\left[(1 + \kappa^2 + 2\kappa\phi)^{\frac{1}{2}}\mathbf{n} - (\mathbf{n}^\top\mathbf{y})\mathbf{s}\right] \times \mathbf{v}, \tag{42}$$

thus, dependence on ρ may be eliminated similar to (34) using the ratio of π_1 and π_2. We then invoke Corollary 1 to constrain π_1 and π_2 to linear functions in z using $m \geq 2$ differential motions, yielding an invariant:

$$(\lambda'_1 + \lambda'_2 z) z_x + (\lambda'_3 + \lambda'_4) z_y + \lambda'_5 = 0, \tag{43}$$

where λ'_i, for $i = 1, \ldots, 5$, are again known entities whose forms are shown in Appendix B.1. □

We urge the reader to observe that the constraint (43) is invariant to the functional form of the BRDF, but is not independent of κ. However, note that κ can be estimated from image data without requiring a full BRDF estimation, for instance, using the methods proposed in [11]. Also, we again note that Proposition 1 is an intrinsic property of isotropic BRDFs and $\boldsymbol{\pi}^\top \mathbf{v} = 0$ continues to hold even for the $\boldsymbol{\pi}$ in (42).

It may be noted that at least two of the differential motions in the orthographic case must be linearly independent for the BRDF invariance results to hold.

4.3 Results on Surface Estimation

Recall that Section 4.1 establishes that, for general unknown isotropic BRDFs, one may neither estimate the surface depth, nor derive any BRDF-invariant constraints on it. However, Section 4.2 derives constraints on depth and gradient for restricted (but unknown) BRDFs. Here, we characterize the PDE defined by those constraints, which directly informs the extent to which shape may be recovered.

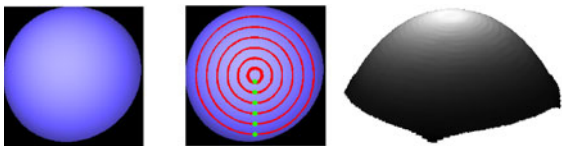
For a BRDF of the form (25), an invariant of the form (28) is obtained. We note that (28) is characterized as a *homogeneous first-order quasilinear PDE* [23]. For a surface level curve $z(x, y) = z_0$, the solution to (28) from PDE theory is:

$$z = z_0, \quad \frac{dy}{dx} = \frac{(\gamma_3 - E_u\gamma_1) + E_u z_0}{(\gamma_2 + E_v\gamma_1) - E_v z_0}. \tag{44}$$

That is, given the depth at a point, the ODE (44) defines a step in the tangent direction, thereby tracing out the level curve through that point. We refer the reader to Appendix C.1 for a formal proof, while only stating the result here:

Proposition 6. *Under orthography, two or more differential motions of the camera yield level curves of depth for a surface with BRDF dependent on light source and view angles.*

For half-angle BRDFs given by (29), a BRDF-invariant constraint on surface depth is obtained as (39). We note that it is characterized as an *inhomogeneous first-order quasilinear PDE*, whose solution is also available from PDE theory [23]. Here, we again state the result while referring the reader to Appendix C.2 for a proof:



(a) Input [1 of 3] (b) Level curves (c) Reconstruction

Fig. 3. (a) One of three simulated orthographic images (two motions), with unknown lighting and non-Lambertian BRDF dependent on source and view angles. (b) Level curves estimated using Proposition 6. (c) Surface reconstructed after interpolation.

Proposition 7. *Under orthography, for a surface with half-angle BRDF, two or more differential camera motions yield characteristic surface curves $\mathcal{C}(x(s), y(s), z(s))$ defined by*

$$\frac{1}{\lambda_1 + \lambda_2 z}\frac{dx}{ds} = \frac{1}{\lambda_3 + \lambda_4 z}\frac{dy}{ds} = \frac{-1}{\lambda_5}\frac{dz}{ds} \tag{45}$$

corresponding to depths at some (possibly isolated) points.

Finally, for a BRDF dependent on an arbitrary angle in the source-view plane given by (41), the invariant (43) is also an inhomogeneous quasilinear PDE. So, as a direct generalization of Proposition 7, differential stereo also yields characteristic curves for this case (with λ_i' instead of λ_i).

4.3.1 Shape Recovery

Given depths at a few points on a surface with unknown BRDF, the above propositions yield depths along certain characteristic curves. For a smooth surface, one may interpolate the depths between the curves, in order to recover depth for the whole surface. The procedure is shown for synthetic data in Fig. 3 for a BRDF that depends on source and view angles (unknown lighting) and in Fig. 4 for unknown half-angle BRDFs (known lighting). Depth is assumed known at the green points to yield characteristic curves shown in red. We note that reconstruction methods in practice may use tracked feature points as seeds.

5 Perspective Projection

In this section, we consider shape recovery from differential stereo under perspective projection. We show that depth may be unambiguously recovered in the perspective case, even when both the BRDF and lighting are unknown. For ease of presentation, we illustrate the theory through the approximate differential stereo relation, while noting that similar results are available for the exact relation in Appendix D.

5.1 Depth from Differential Stereo

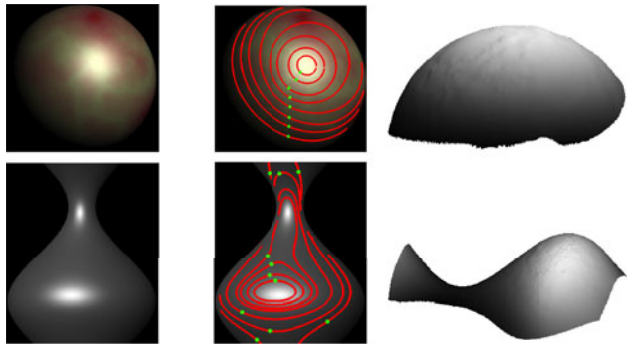
In the perspective case, the motion field $\boldsymbol{\mu}$ is given by (4). Substituting for $\boldsymbol{\mu}$ in the differential stereo relation (14), we obtain its expression for the perspective case:

$$p'\left(\frac{z}{1+\beta z}\right) + r'\left(\frac{1}{1+\beta z}\right) + q' \;=\; \boldsymbol{\omega}^\top \boldsymbol{\pi}, \tag{46}$$

where $\boldsymbol{\pi}$ is defined by (12) and we have known entities

$$p' = E_u\omega_2 - E_v\omega_1, \tag{47}$$

$$q' = \alpha_1 E_u + \alpha_3 E_v + E_t \tag{48}$$



(a) Input [1 of 3] (b) Characteristics (c) Reconstruction

Fig. 4. (a) One of three synthetic images (two motions), with unknown half-angle BRDF and known lighting, under orthographic projection. (b) Characteristic curves estimated using Proposition 7. (c) Surface reconstructed after interpolation.

$$\begin{aligned} &= \beta(u\omega_2 - v\omega_1)(uE_u + vE_v) + (uE_v - vE_u)\omega_3 + E_t, \\ r' &= \alpha_2 E_u + \alpha_4 E_v \\ &= E_u\tau_1 + E_v\tau_2 - \beta(uE_u + vE_v)\tau_3. \end{aligned} \tag{49}$$

Unlike the orthographic case, the differential stereo relation is not linear in $\{z, \boldsymbol{\pi}\}$ for perspective projection. We can now show that camera motion unambiguously determines surface depth under perspective projection:

Proposition 8. *Under perspective projection, three differential motions of the camera suffice to yield depth of a surface with unknown isotropic BRDF and unknown light source.*

Proof. For $m \geq 3$, let images $E_1, \ldots, E_m$ be related to a base image E_0 by known differential motions $\{\boldsymbol{\omega}^i, \boldsymbol{\tau}^i\}$. Then, we obtain from (46) a sequence of differential stereo relations:

$$(p'^i + \beta q'^i)z - ((1+\beta z)\boldsymbol{\pi})^\top \boldsymbol{\omega}^i + (q'^i + r'^i) = 0, \tag{50}$$

for $i = 1, \ldots, m$. Let $\tilde{\mathbf{C}} = [\mathbf{c}^1, \ldots, \mathbf{c}^m]^\top$ be the $m \times 3$ matrix with rows $\mathbf{c}^i = [-(p'^i + \beta q'^i), \omega_1^i, \omega_2^i]^\top$. Further, let $\mathbf{q}' = [q'^1, \ldots, q'^m]^\top$ and $\mathbf{r}' = [r'^1, \ldots, r'^m]^\top$. Then, the system of differential stereo relations (50) may be written as

$$\tilde{\mathbf{C}}\begin{bmatrix} z \\ (1+\beta z)\pi_1 \\ (1+\beta z)\pi_2 \end{bmatrix} = \mathbf{q}' + \mathbf{r}', \tag{51}$$

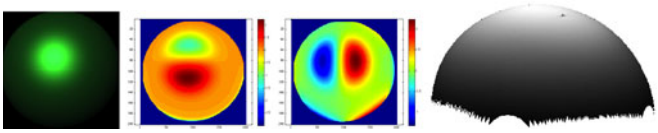
since $\pi_3 = 0$, from Proposition 1. With $\tilde{\mathbf{C}}^+$ the Moore-Penrose pseudoinverse of $\tilde{\mathbf{C}}$, let $\boldsymbol{\epsilon} = (\epsilon_1, \epsilon_2, \epsilon_3)^\top = \tilde{\mathbf{C}}^+(\mathbf{q}' + \mathbf{r}')$. Then, (51) has the solution:

$$[z,\ (1+\beta z)\pi_1,\ (1+\beta z)\pi_2]^\top \;=\; \boldsymbol{\epsilon}. \tag{52}$$

It follows that $z = \epsilon_1$ yields the surface depth. □

Thus, Proposition 8 yields surface depth from differential stereo when both the BRDF and lighting are unknown. Again, a comparison is merited to the case of object motion in [12]. With object motion, depth and an additional constraint on the gradient are recovered with unknown BRDF and lighting. While camera motion also recovers depth, further information on the gradient is recoverable with additional constraints on the BRDF or lighting, as explored in Section 5.2.

We note that the rotations $\boldsymbol{\omega}^i$ for perspective shape recovery must span $\mathbb{R}^3$ to avoid degeneracy in (51). Further, a pure



(a) Image (b) π_1 (c) π_2 (d) Depth

Fig. 5. (a) One of four synthetic images (three motions), with unknown non-Lambertian BRDF and lighting, under perspective projection. (b, c) π_1 and π_2 recovered using (55), as discussed in Section 6. (d) Depth estimated using Proposition 8. Mean reconstruction error is 1.1 percent.

rotation through the optical center does not contain depth information. In our coordinate system with origin at the image center, note that a rotation $\boldsymbol{\omega}$ with $\boldsymbol{\tau} = \mathbf{0}$ displaces the optical center $(0, 0, -f)^\top$. A motion that maps the optical center to itself is characterized as $\{\boldsymbol{\omega}, \boldsymbol{\tau}\}$ with $\boldsymbol{\tau} = (f\omega_2, -f\omega_1, 0)^\top$. In that case, we observe that $r' = fp'$ and all the information on z in (46) is eliminated. Thus, no depth information is available from differential stereo in this case.

5.2 Additional Constraints for Certain Materials

As indicated above, we now show that additional constraints on the gradient are available for several material types.

Proposition 9. *Under perspective projection, three or more differential motions of the camera suffice to yield both depth and a linear constraint on the gradient for a surface with BRDF dependent on known light and half-angle directions.*[1]

Proof. Proposition 8 shows depth recovery from $m \geq 3$ differential motions of the camera. Further, from the form of $\boldsymbol{\pi}$ for half-angle BRDFs given by (33) and using (52), we have

$$\frac{\pi_1}{\pi_2} = \frac{\mathbf{a}^\top \mathbf{n}}{\mathbf{b}^\top \mathbf{n}} = \frac{\epsilon_2}{\epsilon_3}, \tag{53}$$

where $\mathbf{a}$ and $\mathbf{b}$ are known entities dependent on light, defined by (34). Finally, using (6) yields a linear PDE in depth:

$$l_1 z_x + l_2 z_y + l_3 = 0, \tag{54}$$

with $l_1 = \epsilon_2 b_1 - \epsilon_3 a_1$, $l_2 = \epsilon_2 b_2 - \epsilon_3 a_2$, $l_3 = \epsilon_3 a_3 - \epsilon_2 b_3$. Thus, we obtain a linear constraint on the gradient. □

We note that materials for which a relation between π_1 and π_2 may be derived independent of derivatives of the BRDF ρ, yield an additional constraint on the surface gradient. The form of this constraint will mirror the relationship between π_1 and π_2, whose ratio is obtainable from (53). In particular, it follows that a linear constraint may also be derived for the more general case considered in Section 4.2.3:

Remark 2. Under perspective projection, three differential motions of the camera suffice to yield both depth and a linear constraint on the gradient for a BRDF dependent on known light and an arbitrary direction in the source-view plane.

1. Appendix D proves analogous results for the exact differential stereo relation (11). Six motions instead of three are required for depth recovery and a quasilinear PDE in surface depth (instead of linear PDE) is obtained.

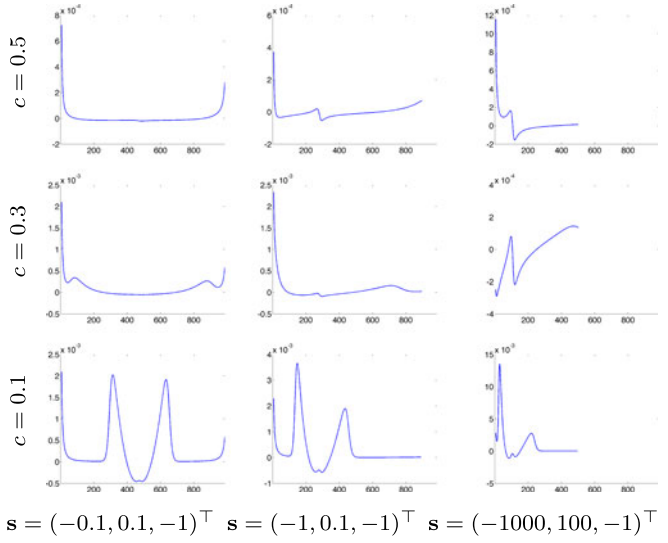


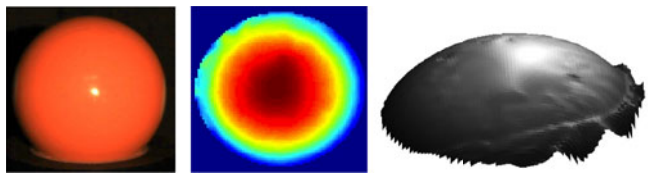
Fig. 6. Trends in the error made by neglecting $\nabla_{\mathbf{x}}\rho$. The units for the x-axis are pixels and those for the y-axis are meters. In each plot, we observe that the error in approximation is higher in the surface regions where the BRDF sharpness is high or where the surface normal changes more sharply. Along each column, we observe that the error in approximation is higher for a sharper BRDF.

5.3 Shape Recovery

From Proposition 8, depth is directly available in the perspective case for unknown lighting and unknown BRDF. Fig. 5 illustrates this for synthetic data, where a sphere of diameter 20 cm is imaged under an unknown light source. Three random motions, of approximately 2 degree rotation and 5 mm translation are imparted to the camera. The recovered shape using the theory of Section 5.1 is shown in Fig. 5. No prior knowledge of the BRDF, lighting or depth are required.

To illustrate the effect of neglecting $\nabla_{\mathbf{x}}\rho$, we use BRDFs of the form $\rho \sim \exp(-(\frac{\cos^{-1}(\mathbf{n}^\top \mathbf{h})}{c})^2)$. A smaller value of c corresponds to a sharper BRDF curve. The depth error across the horizontal cross-section of the sphere through the center are plotted in Fig. 6 for various values of c and light source positions. It is observed that the error increases for sharper changes in the BRDF or surface normal.

For evaluation with real data, images of the non-Lambertian object in Fig. 7a are obtained against a textured background under unknown directional lighting. The camera focal length is 55 mm and the object is approximately 80 cm away. Five small motions (and several large ones for stable pose optimization) are imparted to the camera and estimated using bundle adjustment. Since differentiation tends to amplify noise, spatial image derivatives are computed with a Savitzky-Golay smoothing filter. Temporal image derivatives are computed as differences of images



(a) Image [1 of 5] (b) Depth map (c) Recovered surface

Fig. 7. Reconstruction using Proposition 8 for a non-Lambertian ball with an unknown BRDF, under unknown light source.

smoothed by a Gaussian kernel. Surface depth obtained using Proposition 8 is shown in Figs. 7b and 7c.

A limitation of using differential motions for shape recovery is the sensitivity to noise. While the focus of this paper is to determine what information for shape recovery is available from camera motion, a practical surface reconstruction approach based on this theory must adopt techniques such as coarse-to-fine estimation to deal with large motions and noise. We defer the development of such methods to future work.

6 Perspectives on Shape from Motion

We now provide a unified perspective on shape from motion theories corresponding to light, object or camera motions. Despite the apparent complexity of material behavior, differential motion of light, object or camera allows shape recovery with unknown BRDF and often unknown lighting. More importantly, theoretical limits on shape from motion are also derivable in these frameworks. Prior works have presented theories for light [10] and object [12] motions. This paper has studied camera motion to complete the differential analysis framework for shape recovery with unknown BRDFs.

In Table 1, we summarize a few results from each theory. In each case, the theories generalize well-known special cases that assume Lambertian BRDF or brightness constancy. Specifically, the theory for light source motion generalizes photometric stereo, while those for object and camera motion generalize optical flow and multiview stereo, respectively. A few important traits are shared by these theories:

- They all rely on the linearity of chain rule differentiation to eliminate the BRDF from a system of equations.
- The invariant in each case can be characterized as a PDE amenable to solution through standard analysis tools.
- The involved PDEs provide intrinsic limits on the topological class upto which shape may be recovered from each motion cue, regardless of reconstruction method.
- More general imaging necessitates greater number of motions for shape recovery. For instance, general lighting requires more motions than a colocated one, or perspective projection requires more motions than orthographic.
- Constraining the BRDF either reduces the minimum requirement on number of motions (compare BRDFs under colocated illumination with general BRDFs for object motion), or provides richer shape information (compare half-angle and general BRDFs for camera motion).

The cases of object and camera motion are more closely related, but with important differences due to additional ambiguities entailed by camera motion. Qualitatively, this leads to a harder problem for the case of camera motion. The practical manifestation of this hardness is requiring a more restricted BRDF (although still unknown) to obtain the same shape information. For instance, a half-angle BRDF yields depth and a gradient constraint with camera motion, while the same can be obtained with general BRDFs for object motion. Allowing a general BRDF means only depth may be obtained for camera motion, while object motion yields an additional gradient constraint. Throughout this paper, we have highlighted such distinctions and their impact on shape recovery.

6.1 Future Work

A unified framework for understanding the information available from motion cues can have several applications, which form the subject of our ongoing and future work. The following paragraphs illustrate a few such avenues for further research.

6.1.1 Shape Recovery

An[2] interesting aspect of our theory is that diffuse photoconsistency of conventional multiview stereo follows as a special case, just as brightness constancy of optical flow is a special case for [12]. Our differential flow and stereo relations enjoy a similarity in form to optical flow, for which robust frameworks have been designed in recent years to handle large motions [7]. Such estimation techniques will also be crucial for overcoming the sensitivity to noise inherent in the use of differential cues. We anticipate the development of robust estimation methods that exploit our theory to extend traditional implementations of optical flow and multiview stereo to handle general BRDFs. Further, while [10], [12] and this paper provide limits on shape recovery from motions corresponding to each imaging element, an interesting problem is to achieve similar limits when lighting, object and camera all undergo simultaneous motion.

6.1.2 Joint Shape and BRDF Estimation

Our theory focuses on shape recovery, but some BRDF information may also be recovered as a by-product. For instance, with perspective projection, we obtain from (52)

$$\boldsymbol{\pi} = \frac{1}{1+\beta\epsilon_1}\,(\epsilon_2, \epsilon_3, 0)^\top. \tag{55}$$

Fig. 5 shows an example recovery of $\boldsymbol{\pi}$. In turn, this places constraints on the derivative of BRDF. Our subsequent work in [8] characterizes the extent to which BRDF may be recovered using motions of the object or the camera. Again, a stratification may be obtained that relates the well-posedness of joint shape and material recovery to imaging or material complexity. Surprisingly, in many cases, problems that have traditionally been assumed underconstrained are shown to be well-posed. This allows the development of joint shape and BRDF estimation methods that do not need to impose restrictive priors or perform alternating optimization with uncertain convergence, which is a subject of our ongoing work.

6.1.3 Shape and Material Perception

We believe our results also motivate research in perception of shape and material. It is well-known that specular highlights aid the perception of shape. Blake and Bülthoff determine in [5] that human perception of shape exploits a physical model for specular reflection based on ray optics

2. The number of camera motions is listed for the approximate differential stereo relation. Three additional motions are required with the exact relation.

and differential geometry. Their conclusions are also supported by works in computer vision, such as [43]. Our works on shape recovery from motion of the light source [10], object [12] or camera (this paper) also exploit differential geometry, but are more general since they span the entire spectrum of gloss from diffuse to specular. Moreover, we provide a stratification that establishes the relative "hardness" of surface reconstruction, with respect to complexity of material type or imaging setup. It will be an interesting avenue of future work to determine whether visual perception follows similar notions of hardness.

Appendix A
Differential Stereo Relation

Section 3 outlines intuition into the differential stereo relation of (14). Here, we provide a rigorous derivation from first principles.

We wish to relate change in image intensities to rigid-body motion of the camera, given by rotation $\widetilde{\mathbf{R}}$ and translation $\widetilde{\boldsymbol{\tau}}$, while the scene (object and light source) remain static. Recall that for the purposes of analysis, this is equivalent to assuming that the camera is fixed, while the object and source undergo the inverse motion given by rotation $\mathbf{R} = \widetilde{\mathbf{R}}^{-1}$ and translation $\boldsymbol{\tau} = \widetilde{\mathbf{R}}^{-1}\widetilde{\boldsymbol{\tau}}$. For differential motion, we may approximate $\mathbf{R} \approx \mathbf{I} + [\boldsymbol{\omega}]_\times$, where $\boldsymbol{\omega} = (\omega_1, \omega_2, \omega_3)^\top$.

Differential Entities

We define the position vector $\mathbf{x}_t(a,b)$ which encodes the 3D coordinates of the point $(a,b)^\top$ on the surface at time t. Similarly, $\mathbf{n}_t(a,b)$ is the corresponding unit surface normal. We will follow optical flow studies like [29], [39] to distinguish between intrinsic coordinates (a,b) for entities on the surface (such as albedo), as opposed to 3D coordinates (for entities like the camera).

Consider a point $\mathbf{u}$ on the image. At time t, it is the image of the point $\mathbf{p} = \mathbf{x}_t(a,b)$. At time $t+\delta t$, it is the image of a different point $\mathbf{q} = \mathbf{x}_{t+\delta t}(a-\delta a, b-\delta b)$. The displacement of the point $(a-\delta a, b-\delta b)^\top$ between times t and $t+\delta t$ is given by

$$\mathbf{x}_{t+\delta t}(a-\delta a, b-\delta b) = \mathbf{x}_t(a-\delta a, b-\delta b) + \delta\mathbf{x}. \tag{56}$$

We have suppressed the $(a-\delta a, b-\delta b)$ argument of $\delta\mathbf{x}$. Denoting $\boldsymbol{\nu}_t(a,b)$ as the linear velocity of $(a,b)^\top$ at time t, we have

$$\delta\mathbf{x} = \boldsymbol{\nu}_t(a-\delta a, b-\delta b)\delta t. \tag{57}$$

Similarly, the unit surface normals corresponding to the image point $\mathbf{u}$ at times t and $t+\delta t$ are related by

$$\mathbf{n}_{t+\delta t}(a-\delta a, b-\delta b) = \mathbf{n}_t(a-\delta a, b-\delta b) + \delta\mathbf{n}. \tag{58}$$

For the translational component of the rigid body motion, $\delta\mathbf{n} = 0$. For the rotational component, the change in surface normal is determined by the angular velocity. Thus,

$$\delta\mathbf{n} = \boldsymbol{\omega} \times \mathbf{n}_t(a-\delta a, b-\delta b)\delta t. \tag{59}$$

In general, the light source must be considered in the 3D world coordinate system. However, in our particular setup of camera motion with a fixed object and directional distant lighting, the relative position of the lighting does not change with respect to the surface. Thus, the lighting may also be considered in intrinsic surface coordinates. Consequently, the light source directions corresponding to the image point $\mathbf{u}$ at times t and $t+\delta t$ are related by

$$\delta\mathbf{s} = \boldsymbol{\omega} \times \mathbf{s}_t(a-\delta a, b-\delta b)\delta t. \tag{60}$$

Thus, we have defined the following differential relations:

$$\mathbf{x}_{t+\delta t}(a-\delta a, b-\delta b) = \mathbf{x}_t(a-\delta a, b-\delta b) + \boldsymbol{\nu}_t\delta t, \tag{61}$$

$$\mathbf{n}_{t+\delta t}(a-\delta a, b-\delta b) = \mathbf{n}_t(a-\delta a, b-\delta b) + \boldsymbol{\omega} \times \mathbf{n}_t\delta t, \tag{62}$$

$$\mathbf{s}_{t+\delta t}(a-\delta a, b-\delta b) = \mathbf{s}_t(a-\delta a, b-\delta b) + \boldsymbol{\omega} \times \mathbf{s}_t\delta t. \tag{63}$$

Differential Stereo

The BRDF ρ at a point is a function of its position, normal, light source and camera directions. Let the albedo, which is an intrinsic surface property, be $\sigma(a,b)$. Then, at time t, suppose a 3D point $\mathbf{p} = \mathbf{x}_t(a,b)$ is imaged at pixel $\mathbf{u}$. The image formation may be written as:

$$I(\mathbf{u},t) = \sigma(a,b)\ \rho(\mathbf{x}_t(a,b), \mathbf{n}_t(a,b), \mathbf{s}_t(a,b), \mathbf{v}). \tag{64}$$

At time $t+\delta t$, the image at the same pixel $\mathbf{u}$ will correspond to a different 3D point $\mathbf{q} = \mathbf{x}_{t+\delta t}(a-\delta a, b-\delta b)$, since the object has moved relative to the camera. Thus, image formation is given by:

$$I(\mathbf{u}, t+\delta t) = \sigma(a-\delta a, b-\delta b)\rho(\mathbf{x}_{t+\delta t}, \mathbf{n}_{t+\delta t}, \mathbf{s}_{t+\delta t}, \mathbf{v}), \tag{65}$$

where all entities in ρ are evaluated at $(a-\delta a, b-\delta b)$. The image of this 3D point $\mathbf{q}$ at time t must have been formed at a different 2D location on the image plane, $\mathbf{u} - \delta\mathbf{u}$. Thus, the image formation for 3D point $\mathbf{q}$ at time t is given by:

$$I(\mathbf{u} - \delta\mathbf{u}, t) = \sigma(a-\delta a, b-\delta b)\rho(\mathbf{x}_t, \mathbf{n}_t, \mathbf{s}_t, \mathbf{v}). \tag{66}$$

with all entities in ρ again evaluated at $(a-\delta a, b-\delta b)$. Subtracting (66) from (65), we have

$$\begin{aligned} I(\mathbf{u}, t+\delta t) - I(\mathbf{u} - \delta\mathbf{u}, t) = \sigma[&\rho(\mathbf{x}_{t+\delta t}, \mathbf{n}_{t+\delta t}, \mathbf{s}_{t+\delta t}, \mathbf{v}) \\ &- \rho(\mathbf{x}_t, \mathbf{n}_t, \mathbf{s}_t, \mathbf{v})]. \end{aligned} \tag{67}$$

Applying chain-rule differentiation and using the differential entities defined in (61), (62) and (63), the above may be rewritten as:

$$\begin{aligned} I(\mathbf{u}, t+\delta t) - I(\mathbf{u} - \delta\mathbf{u}, t) = \sigma[&(\nabla_{\mathbf{x}}\rho)^\top \boldsymbol{\nu}_t\delta t \\ &+ (\nabla_{\mathbf{n}}\rho)^\top(\boldsymbol{\omega} \times \mathbf{n}_t)\delta t + (\nabla_{\mathbf{s}}\rho)^\top(\boldsymbol{\omega} \times \mathbf{s}_t)\delta t], \end{aligned} \tag{68}$$

where complete arguments for above variables are $\boldsymbol{\nu}_t(a-\delta a, b-\delta b)$, $\mathbf{n}_t(a-\delta a, b-\delta b)$ and $\mathbf{s}_t(a-\delta a, b-\delta b)$. The BRDF-derivatives $\nabla_{\mathbf{x}}\rho$, $\nabla_{\mathbf{n}}\rho$ and $\nabla_{\mathbf{s}}\rho$ are also evaluated at $(a-\delta a, b-\delta b)$, at time t. We now note definitions for spatial and temporal partial derivatives of $I(\mathbf{u},t)$:

$$(\nabla_{\mathbf{u}} I)^\top \delta\mathbf{u} = I(\mathbf{u},t) - I(\mathbf{u} - \delta\mathbf{u}, t) \tag{69}$$

$$\frac{\partial I}{\partial t}\delta t = I(\mathbf{u}, t+\delta t) - I(\mathbf{u}, t). \tag{70}$$

Substituting both the above definitions into (68), we obtain

$$\begin{aligned}\frac{\partial I}{\partial t}\delta t &= I(\mathbf{u}, t+\delta t) - I(\mathbf{u}, t)\\ &= I(\mathbf{u}-\delta\mathbf{u}, t) + \sigma[(\nabla_{\mathbf{x}}\rho)^\top \boldsymbol{\nu}_t \delta t + (\nabla_{\mathbf{n}}\rho)^\top(\boldsymbol{\omega}\times\mathbf{n}_t)\delta t\\ &\quad + (\nabla_{\mathbf{s}}\rho)^\top(\boldsymbol{\omega}\times\mathbf{s}_t)\delta t] - I(\mathbf{u}, t)\\ &= -(\nabla_{\mathbf{u}} I)^\top \delta\mathbf{u} + \sigma[(\nabla_{\mathbf{x}}\rho)^\top \boldsymbol{\nu}_t \delta t + (\nabla_{\mathbf{n}}\rho)^\top(\boldsymbol{\omega}\times\mathbf{n}_t)\delta t\\ &\quad + (\nabla_{\mathbf{s}}\rho)^\top(\boldsymbol{\omega}\times\mathbf{s}_t)\delta t].\end{aligned} \tag{71}$$

Recall the definition of motion field, $\boldsymbol{\mu}$, as the velocity of the image pixel $\mathbf{u}$:

$$\boldsymbol{\mu} = \frac{\delta\mathbf{u}}{\delta t}. \tag{72}$$

Then, using (72), the relation in (71) can be written as

$$\begin{aligned}(\nabla_{\mathbf{u}} I)^\top \boldsymbol{\mu} + \frac{\partial I}{\partial t} = \sigma[(\nabla_{\mathbf{x}}\rho)^\top \boldsymbol{\nu} + (\nabla_{\mathbf{n}}\rho)^\top(\boldsymbol{\omega}\times\mathbf{n})\\ + (\nabla_{\mathbf{s}}\rho)^\top(\boldsymbol{\omega}\times\mathbf{s})].\end{aligned} \tag{73}$$

Thus, we have derived (10) from first principles. Following the subsequent steps as described in Section 3.1 leads to the differential stereo relation in (14).

Appendix B
BRDF-Invariance under Orthography

We now prove the shape recovery results under orthographic projection stated in Section 4.

Proof of Proposition 5

Proposition 5 generalizes Proposition 4 to an arbitrary angle in the $\{\mathbf{s}, \mathbf{v}\}$-plane. Its proof follows the same constructs and the algebraic details are listed below.

For a BRDF that depends on an arbitrary angle in the $\{\mathbf{s}, \mathbf{v}\}$-plane, we may define

$$\log\rho(\mathbf{n}, \mathbf{s}, \mathbf{v}) = \bar{\rho}(\mathbf{n}^\top\mathbf{s}, \mathbf{n}^\top\mathbf{y}), \text{ with } \mathbf{y} = \frac{\mathbf{s}+\kappa\mathbf{v}}{\|\mathbf{s}+\kappa\mathbf{v}\|}, \tag{74}$$

where $\kappa \in \mathbb{R}$. Recall the definition of $\boldsymbol{\pi}$ in (12), which may be rewritten as:

$$\boldsymbol{\pi} = \mathbf{n}\times\nabla_{\mathbf{n}}\bar{\rho} + \mathbf{s}\times\nabla_{\mathbf{s}}\bar{\rho}. \tag{75}$$

We denote $\theta = \mathbf{n}^\top\mathbf{s}$, $\phi = \mathbf{s}^\top\mathbf{v}$, $\psi = \mathbf{n}^\top\mathbf{v}$ and $\eta = \mathbf{n}^\top\mathbf{y}$. Using the definition of $\mathbf{y}$ in (74) and applying chain-rule differentiation, we obtain:

$$\mathbf{n}\times\nabla_{\mathbf{n}}\bar{\rho} = \bar{\rho}_\theta(\mathbf{n}\times\mathbf{s}) + \bar{\rho}_\eta\frac{\mathbf{n}\times\mathbf{s}}{\|\mathbf{s}+\kappa\mathbf{v}\|} + \kappa\bar{\rho}_\eta\frac{\mathbf{n}\times\mathbf{v}}{\|\mathbf{s}+\kappa\mathbf{v}\|}, \tag{76}$$

$$\mathbf{s}\times\nabla_{\mathbf{s}}\bar{\rho} = \bar{\rho}_\theta(\mathbf{s}\times\mathbf{n}) + \bar{\rho}_\eta\frac{\mathbf{s}\times\mathbf{n}}{\|\mathbf{s}+\kappa\mathbf{v}\|} - \kappa\bar{\rho}_\eta\frac{(\mathbf{n}^\top(\mathbf{s}+\kappa\mathbf{v}))\mathbf{s}\times\mathbf{v}}{\|\mathbf{s}+\kappa\mathbf{v}\|^3}. \tag{77}$$

Adding (76) and (77) and substituting in (75), we obtain:

$$\boldsymbol{\pi} = \frac{\kappa\rho_\eta}{(1+\kappa^2+2\kappa\phi)^{\frac{3}{2}}}\left[(1+\kappa^2+2\kappa\phi)\mathbf{n} - (\mathbf{n}^\top\mathbf{s} + \kappa\mathbf{n}^\top\mathbf{v})\mathbf{s}\right]\times\mathbf{v}. \tag{78}$$

Then, we may eliminate dependence on the BRDF by considering the ratio:

$$\frac{\pi_1}{\pi_2} = \frac{\mathbf{a}'^\top\mathbf{n}}{\mathbf{b}'^\top\mathbf{n}} = \frac{a'_1 z_x + a'_2 z_y - a'_3}{b'_1 z_x + b'_2 z_y - b'_3}, \tag{79}$$

where we have used the relationship between surface normal and gradient given by (6) and denoted

$$\begin{aligned} a'_1 &= s_1 s_2, & b'_1 &= (1+\kappa^2+2\kappa\phi) - s_1^2,\\ a'_2 &= s_2^2 - (1+\kappa^2+2\kappa\phi), & b'_2 &= -s_1 s_2,\\ a'_3 &= s_2(s_3-\kappa), & b'_3 &= -s_1(s_3-\kappa).\end{aligned} \tag{80}$$

Now we invoke Corollary 1, which stipulates that $m \geq 2$ differential motions of the camera suffice to restrict $\boldsymbol{\pi}$ to a linear relation in z. In particular, from (23) and (24), we have:

$$\frac{\pi_1}{\pi_2} = \frac{(\gamma_2 + E_v\gamma_1) - E_v z}{(\gamma_3 - E_u\gamma_1) + E_u z}, \tag{81}$$

where $\boldsymbol{\gamma} = (\gamma_1, \gamma_2, \gamma_3)^\top$ is defined by (22). Thus, from (34) and (37), we have obtained

$$\frac{a'_1 z_x + a'_2 z_y - a'_3}{b'_1 z_x + b'_2 z_y - b'_3} = \frac{(\gamma_2 + E_v\gamma_1) - E_v z}{(\gamma_3 - E_u\gamma_1) + E_u z}, \tag{82}$$

which may be rewritten as

$$(\lambda'_1 + \lambda'_2 z)z_x + (\lambda'_3 + \lambda'_4)z_y + \lambda'_5 = 0, \tag{83}$$

where

$$\begin{aligned} \lambda'_1 &= a'_1(\gamma_3 - E_u\gamma_1) - b'_1(\gamma_2 + E_v\gamma_1), & \lambda'_2 &= a'_1 E_u + b'_1 E_v,\\ \lambda'_3 &= a'_2(\gamma_3 - E_u\gamma_1) - b'_2(\gamma_2 + E_v\gamma_1), & \lambda'_4 &= a'_2 E_u + b'_2 E_v,\\ \lambda'_5 &= -a'_3(\gamma_3 - E_u\gamma_1) + b'_3(\gamma_2 + E_v\gamma_1).\end{aligned} \tag{84}$$

We have now obtained the constraint (43) that relates the surface depth z to image derivatives and is independent of the BRDF.

Appendix C
Surface Estimation under Orthography

We now prove the shape recovery results under orthographic projection stated as Propositions 6 and 7.

Proof of Proposition 6

Proposition 6 shows that for a surface with unknown BRDF dependent on light source and view angles, observed under unknown light source with orthographic projection, two differential motions of the camera suffice to recover level curves of surface depth corresponding to depths at some (possibly isolated) points.

For a BRDF that depends on light source and view angles, Remark 1 stipulates a BRDF-invariant constraint that relates depth to image derivatives. In particular, we have from (28):

$$\frac{z_x}{z_y} = -\frac{(\gamma_3 - E_u\gamma_1) + E_u z}{(\gamma_2 + E_v\gamma_1) - E_v z}. \tag{85}$$

Note that (85) represents a first-order, homogeneous, quasi-linear PDE. This immediately suggests a method of characteristics to solve it, using standard constructs from PDE theory. Specifically, we define

$$\mathbf{a} = ((\gamma_2 + E_v\gamma_1) - E_v z,\ (\gamma_3 - E_u\gamma_1) + E_u z)^\top. \tag{86}$$

Then, from (85), we have that

$$\mathbf{a}^\top \nabla z = 0. \tag{87}$$

From differential geometry, we know that the gradient ∇z is orthogonal to the level curves of surface z. Thus, the tangent space to the level curves of z is defined by $\mathbf{a}$. Consider a rectifiable curve, $\mathcal{C}(x(s), y(s))$, parameterized by the arc length parameter s. The derivative of z along $\mathcal{C}$ is given by

$$\frac{dz}{ds} = \frac{\partial z}{\partial x}\frac{dx}{ds} + \frac{\partial z}{\partial y}\frac{dy}{ds}. \tag{88}$$

If $\mathcal{C}$ is a level curve of $z(x, y)$, then the value of z is constant, thus, $\frac{dz}{ds} = 0$ on $\mathcal{C}$. Define $\mathbf{b} = (\frac{dx}{ds}, \frac{dy}{ds})$. Then, we also have

$$\mathbf{b}^\top \nabla z = 0. \tag{89}$$

From (87) and (89), it follows that $\mathbf{a}$ and $\mathbf{b}$ are parallel. Thus, $\frac{b_2}{b_1} = \frac{a_2}{a_1}$, whereby we get from (85):

$$\frac{dy}{dx} = \frac{(\gamma_3 - E_u\gamma_1) + E_u z}{(\gamma_2 + E_v\gamma_1) - E_v z}. \tag{90}$$

Along a level curve $z(x, y) = z_0$, the solution is given by

$$z = z_0 \tag{91}$$

$$\frac{dy}{dx} = \frac{(\gamma_3 - E_u\gamma_1) + E_u z_0}{(\gamma_2 + E_v\gamma_1) - E_v z_0}. \tag{92}$$

Now, given the value of z at any point on the surface, we can use the ODE in (92) to determine all other points on the surface with the same value of z. Thus, (85) allows reconstruction of level curves of the surface, with unknown BRDF and unknown light source.

Proof of Proposition 7

Proposition 7 states that for a BRDF of unknown functional form that depends on the half-angle, two or more differential motions of the camera suffice to yield characteristic surface curves given by (45).

Consider the PDE in (39), given by

$$(\lambda_1 + \lambda_2 u)u_x + (\lambda_3 + \lambda_4 u)u_y + \lambda_5 = 0, \tag{93}$$

where λ_i, for $i = 1, \ldots, 5$, are known functions of (x, y) from (40).

We established in Section 4.2.2 that our surface of interest is the integral surface of PDE (93), denoted as $\mathcal{S} : z = u(x, y)$. It is also shown in Section 4.2.2 that the coefficient functions λ_i, for $i = 1, \ldots, 5$, can be obtained from two or more differential motions of the camera. We now rewrite (93) in the form

$$\mathbf{a}^\top \begin{bmatrix} \nabla u \\ -1 \end{bmatrix} = 0 \tag{94}$$

with $\mathbf{a} = (\lambda_1 + \lambda_2 u, \lambda_3 + \lambda_4 u, -\lambda_5)^\top$. Then, we note that the integral surface $\mathcal{S} : z = u(x, y)$ is tangent everywhere to the vector field $\mathbf{a}$. Consider the curve $\mathcal{C}$ of (45), with a parameter $s \in \mathbf{I} \subset \mathbb{R}$:

$$\mathcal{C} :\ x = x(s), y = y(s), z = z(s), \tag{95}$$

where

$$\begin{aligned} &\frac{dx}{ds} = \lambda_1 + \lambda_2 z = a_1(x, y, z),\ \frac{dy}{ds} = \lambda_3 + \lambda_4 z = a_2(x, y, z), \\ &\frac{dz}{ds} = -\lambda_5 = a_3(x, y, z), \end{aligned} \tag{96}$$

where $\mathbf{a} = (a_1, a_2, a_3)^\top$. We note that the curves $\mathcal{C}$, if they exist, have $\mathbf{a}$ as tangent directions. Next, we derive the relationship between $\mathcal{C}$ and $\mathcal{S}$, in particular, we show that if a point $\mathbf{p} = (x_0, y_0, z_0)^\top \in \mathcal{C}$ lies on the integral surface $\mathcal{S}$, then $\mathcal{C} \subset \mathcal{S}$.

Suppose there exists $\mathbf{p} = (x_0, y_0, z_0)^\top \in \mathcal{C}$, such that $\mathbf{p} \in \mathcal{S}$, that is,

$$x_0 = x(s_0), y_0 = y(s_0), z_0 = z(s_0) = u(x_0, y_0), \tag{97}$$

for some parameter value $s = s_0 \in \mathbf{I}$. Next, we define

$$w = w(s) = z(s) - u(x(s), y(s)). \tag{98}$$

Then, it is clear that $w(s)$ is the solution to the initial value problem

$$\frac{dw}{ds} = (u_x, u_y, 1)^\top \mathbf{a}(x, y, w + u) \tag{99}$$

$$w(s_0) = 0. \tag{100}$$

Further, we note that $w = 0$ is a particular solution of the above ordinary differential equation, since $z = u(x, y)$ is a solution to (93). Also, the solution to (99) is unique, thus, $z(s) = u(x(s), y(s))$, which establishes that $\mathcal{C} \subset \mathcal{S}$. This completes the proof that the characteristic curves $\mathcal{C}$, given by (45), reside on the surface $\mathcal{S}$.

Appendix D
Derivation for Exact Relation with $\nabla_{\mathbf{x}}\rho$

The approximate differential stereo relation in (14) assumes that $\nabla_{\mathbf{x}}\rho$ is negligible. While the assumption is reasonable for our setup and allows us to illustrate the developments, we note that it is not a necessary requirement. We now show that depth information may be recovered using motion cues for a surface with unknown BRDF, even if $\nabla_{\mathbf{x}}\rho$ is not assumed to be negligible.

We follow similar steps as illustrated previously. We start with the exact form of the differential stereo relation. Then we use a sequence of motions to derive constraints on surface depth. Next, we show that shape information can be recovered using motion cues.

Differential Stereo Relation

We start with the relation in (10), with no approximations:

$$(\nabla_{\mathbf{u}} E)^\top \boldsymbol{\mu} + E_t = \boldsymbol{\pi}^\top \boldsymbol{\omega} + (\nabla_{\mathbf{x}} \rho)^\top \boldsymbol{\nu}, \tag{101}$$

where $\boldsymbol{\nu} = \dot{\mathbf{x}} = \boldsymbol{\omega} \times \mathbf{x} + \boldsymbol{\tau}$ is the linear velocity and $\boldsymbol{\pi}$ is defined by (12). Then, we rewrite the above as

$$(\nabla_{\mathbf{u}} E)^\top \boldsymbol{\mu} + E_t = \boldsymbol{\pi}'^\top \boldsymbol{\omega} + (\nabla_{\mathbf{x}} \rho)^\top \boldsymbol{\tau}. \tag{102}$$

where we have defined

$$\boldsymbol{\pi}' = \boldsymbol{\pi} + \mathbf{x} \times \nabla_{\mathbf{x}} \rho. \tag{103}$$

Constraints from an Image Sequence

We now substitute for the motion field in (102) using (4), to obtain the counterpart of (46) without considering $\nabla_{\mathbf{x}} \rho$ as negligible:

$$p'\left(\frac{z}{1+\beta z}\right) + r'\left(\frac{1}{1+\beta z}\right) + q' \;=\; \boldsymbol{\omega}^\top \boldsymbol{\pi}' + \boldsymbol{\tau}^\top \nabla_{\mathbf{x}} \rho, \tag{104}$$

where p', q' and r' are defined by (47), (48) and (49). Given observations from $m \geq 6$ motions, we arrange relations of the form (104) into a linear system:

$$\begin{bmatrix} p'^1 & r'^1 & -\boldsymbol{\omega}^{1\top} & -\boldsymbol{\tau}^{1\top} \\ & \vdots & \vdots & \\ p'^m & r'^m & -\boldsymbol{\omega}^{m\top} & -\boldsymbol{\tau}^{m\top} \end{bmatrix} \begin{bmatrix} \frac{z}{1+\beta z} \\ \frac{1}{1+\beta z} \\ \boldsymbol{\pi}' \\ \nabla_{\mathbf{x}} \rho \end{bmatrix} = - \begin{bmatrix} q'^1 \\ \vdots \\ q'^m \end{bmatrix}. \tag{105}$$

We denote the above $m \times 8$ matrix $\mathbf{B}$ and define $\mathbf{q} = (q'^1, \ldots, q'^m)^\top$. As before, we assume that all motions are general, with rotations and translations that span $\mathbb{R}^3$, with the combined $(\boldsymbol{\omega}^{i\top}, \boldsymbol{\tau}^{i\top})^\top$, for $i = 1, \ldots, m$, spanning $\mathbb{R}^6$. Then, it can be observed from the forms of p' and r' that $\mathrm{rank}(\mathbf{B}) = 6$. Let $\boldsymbol{\gamma} = -\mathbf{B}^+ \mathbf{q}$, where $\mathbf{B}^+$ is the Moore-Penrose pseudoinverse of $\mathbf{B}$. Then, for arbitrary $\lambda_1, \lambda_2 \in \mathbb{R}$, the linear system in (105) has solutions of the form:

$$\begin{bmatrix} \frac{z}{1+\beta z} \\ \frac{1}{1+\beta z} \\ \boldsymbol{\pi}' \\ \nabla_{\mathbf{x}} \rho \end{bmatrix} = \boldsymbol{\gamma} + \lambda_1 \begin{bmatrix} 1 \\ 0 \\ \mathbf{g}_1 \\ \mathbf{0}_{3\times 1} \end{bmatrix} + \lambda_2 \begin{bmatrix} 0 \\ 1 \\ \mathbf{0}_{3\times 1} \\ \mathbf{g}_2 \end{bmatrix}, \tag{106}$$

where $\mathbf{g}_1$ and $\mathbf{g}_2$ are known entities given by

$$\mathbf{g}_1 = \begin{bmatrix} g_{11} \\ g_{12} \\ g_{13} \end{bmatrix} = \begin{bmatrix} -E_v \\ E_u \\ 0 \end{bmatrix}, \; \mathbf{g}_2 = \begin{bmatrix} g_{21} \\ g_{22} \\ g_{23} \end{bmatrix} = \begin{bmatrix} E_u \\ E_v \\ -\beta(uE_u + vE_v) \end{bmatrix}. \tag{107}$$

Shape Recovery

Now, we can derive several relations from the above solution. First, we observe from the solution in (106) that

$$\lambda_1 = \frac{z}{1+\beta z} - \gamma_1, \; \lambda_2 = \frac{1}{1+\beta z} - \gamma_2. \tag{108}$$

Next, substituting from (108) into the solution in (106), we obtain

$$\boldsymbol{\pi}' = \boldsymbol{\gamma}' + \left(\frac{z}{1+\beta z} - \gamma_1\right)\mathbf{g}_1, \; \nabla_{\mathbf{x}} \rho = \tilde{\boldsymbol{\gamma}}' + \left(\frac{1}{1+\beta z} - \gamma_2\right)\mathbf{g}_2, \tag{109}$$

where, we have defined $\boldsymbol{\gamma}' = (\gamma_3, \gamma_4, \gamma_5)^\top$ and $\tilde{\boldsymbol{\gamma}}' = (\gamma_6, \gamma_7, \gamma_8)^\top$. From the projection equations in (1), we have

$$\mathbf{x} = (x, y, z)^\top = (u(1+\beta z), v(1+\beta z), z)^\top. \tag{110}$$

Then, using the definition of $\boldsymbol{\pi}'$ along with (109) and (110), we obtain

$$\boldsymbol{\pi} = \boldsymbol{\pi}' - (\mathbf{x} \times \nabla_{\mathbf{x}} \rho) = (c_1 z + c_2, c_3 z + c_4, c_5 z + c_6)^\top, \tag{111}$$

where

$$c_1 = -(\beta^2 v(uE_u + vE_v) + E_v)\gamma_2 + \gamma_7 - \beta v \gamma_8, \tag{112}$$

$$c_2 = \beta(1-\gamma_2)v(uE_u + vE_v) + (E_v \gamma_1 + \gamma_3 - v\gamma_8), \tag{113}$$

$$c_3 = (\beta^2 u(uE_u + vE_v) + E_u)\gamma_2 - \gamma_6 + \beta u \gamma_8, \tag{114}$$

$$c_4 = -\beta(1-\gamma_2)u(uE_u + vE_v) - (E_u \gamma_1 - \gamma_4 - u\gamma_8), \tag{115}$$

$$c_5 = \beta(\gamma_2(uE_v - vE_u) + (v\gamma_6 - u\gamma_7)), \tag{116}$$

$$c_6 = -(1-\gamma_2)(uE_v - vE_u) + (\gamma_5 + v\gamma_6 - u\gamma_7). \tag{117}$$

Similar to the derivations in Section 4.2, for restricted BRDF types such as those dependent on the half-angle, we may now write

$$\frac{\pi_1}{\pi_2} = \frac{\mathbf{e}_1^\top \mathbf{n}}{\mathbf{e}_2^\top \mathbf{n}} = \frac{c_1 z + c_2}{c_3 z + c_4}, \tag{118}$$

where it may be observed from, for instance (33), that $\mathbf{e}_1$ and $\mathbf{e}_2$ depend only on known entities $\mathbf{s}$ and $\mathbf{v}$. Thus, the relation (118) is a quasilinear PDE, which yields characteristics curves of surface depth. Further, with the assumption of Proposition 1, one may obtain the surface depth using $\pi_3 = 0$ in (111), as

$$z = \frac{-c_6}{c_5} = \frac{1}{\beta}\left[\frac{\gamma_5 - uE_v + vE_u}{u(\gamma_7 - \gamma_2 E_v) - v(\gamma_6 - \gamma_2 E_u)} - 1\right]. \tag{119}$$

Thus, six or more differential motions of the observer yield information on surface shape with unknown BRDF.

Acknowledgments

The authors thank Ravi Ramamoorthi for several helpful discussions, as well as Jiamin Bai, Dikpal Reddy and Yizhou Wang for collaborations on preceding works [10], [12]. We also thank the anonymous reviewers of [9] and this paper for their valuable feedback.

References

[1] N. Alldrin, T. Zickler, and D. Kriegman, "Photometric stereo with non-parametric and spatially-varying reflectance," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, 2008, pp. 1–8.

[2] S. Baker, D. Scharstein, J. P. Lewis, S. Roth, M. J. Black, and R. Szeliski, "A database and evaluation methodology for optical flow," *Int. J. Computer Vision*, vol. 92, no. 1, pp. 1–31, 2011.

[3] A. Blake, "Specular stereo," in *Proc. Int. Joint Conf. Artif. Intell.*, 1985, pp. 973–976.
[4] A. Blake and G. Brelstaff, "Geometry from specularities," in *Proc. IEEE Int. Conf. Comput. Vis.*, 1988, pp. 394–403.
[5] A. Blake and H. Bülthoff, "Does the brain know the physics of specular reflection?" *Nature*, vol. 343, no. 6254, pp. 165–168, 1990.
[6] T. Bonfort and P. Sturm, "Voxel carving for specular surfaces," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2003, pp. 591–596.
[7] T. Brox, A. Bruhn, N. Papenberg, and J. Weickert, "High accuracy optical flow estimation based on a theory for warping," in *Proc. Eur. Conf. Comput. Vis.*, 2004, pp. 25–36.
[8] M. Chandraker, "On shape and material recovery from motion," in *Proc. Eur. Conf. Comput. Vis.*, 2014, pp. 202–217.
[9] M. Chandraker, "What camera motion reveals about shape with unknown BRDF," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, 2014, pp. 2179–2186.
[10] M. Chandraker, J. Bai, and R. Ramamoorthi, "A theory of differential photometric stereo for unknown isotropic BRDFs," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, pp. 2505–2512, 2011.
[11] M. Chandraker and R. Ramamoorthi, "What an image reveals about material reflectance," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2011, pp. 1076–1083.
[12] M. Chandraker, D. Reddy, Y. Wang, and R. Ramamoorthi, "What object motion reveals about shape with unknown BRDF and lighting," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, 2013, pp. 2523–2530.
[13] M. Chen and J. Arvo, "Theory and application of specular path perturbation," *ACM Trans. Graph.*, vol. 19, no. 4, pp. 246–278, 2000.
[14] J. Clark, "Active photometric stereo," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, 1992, pp. 29–34.
[15] R. W. Fleming, A. Torralba, and E. H. Adelson, "Specular reflections and the perception of shape," *J. Vis.*, vol. 4, no. 9, pp. 798–820, 2004.
[16] Y. Furukawa, B. Curless, S. Seitz, and R. Szeliski, "Manhattan-world stereo," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, 2009, pp. 1422–1429.
[17] Y. Furukawa and J. Ponce, "Accurate, dense and robust multiview stereopsis," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 8, pp. 1362–1376, Aug. 2010.
[18] D. Gallup, J.-M. Frahm, and M. Pollefeys, "Piecewise planar and non-planar stereo for urban scene reconstruction," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, 2010, pp. 1418–1425.
[19] D. B. Goldman, B. Curless, A. Hertzmann, and S. M. Seitz, "Shape and spatially-varying BRDFs from photometric stereo," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 6, pp. 1060–1071, Jun. 2010.
[20] H. W. Haussecker and D. J. Fleet, "Computing optical flow with physical models of brightness variation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 6, pp. 661–673, Jun. 2001.
[21] C. Hernández and G. Vogiatzis, "Shape from photographs: A multi-view stereo pipeline," in *Computer Vision*, Studies in Computational Intelligence. Berlin, Germany: Springer, 2010.
[22] B. Horn and B. Schunck, "Determining optical flow," *Artificial Intell.*, vol. 17, pp. 185–203, 1981.
[23] F. John, *Partial Differential Equations*, App. Math. Sci. New York, NY, USA: Springer, 1981.
[24] J. Koenderink and A. van Doorn, "Photometric invariants related to solid shape," *Optica Apta*, vol. 27, pp. 981–996, 1980.
[25] B. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," in *Proc. Image Understanding Workshop*, 1981, pp. 121–130.
[26] W. Matusik, H. Pfister, M. Brand, and L. McMillan, "A data-driven reflectance model," *ACM Trans. Graph.*, vol. 22, no. 3, pp. 759–769, 2003.
[27] M. Minnaert, "The reciprocity principle in lunar photometry," *Astrophysical J.*, vol. 3, pp. 403–410, 1941.
[28] A. A. Muryy, A. E. Welchman, A. Blake, and R. W. Fleming, "Specular reflections and the estimation of shape from binocular disparity," *Proc. Nat. Acad. Sci. USA*, vol. 110, no. 6, pp. 2413–2418, 2013.
[29] H.-H. Nagel, "On a constraint equation for the estimation of displacement rates in image sequences," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11, no. 1, pp. 13–30, Jan. 1989.
[30] S. Negahdaripour, "Revised definition of optical flow: Integration of radiometric and geometric cues for dynamic scene analysis," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 9, pp. 961–979, Sep. 1998.
[31] A. Ngan, F. Durand, and W. Matusik, "Experimental analysis of BRDF models," in *Proc. Eurograph. Symp. Rendering*, 2005, pp. 117–126.
[32] J. Norman, J. Todd, and G. Orban, "Perception of three-dimensional shape from specular highlights, deformations of shading and other types of visual information," *Psychological Sci.*, vol. 15, no. 8, pp. 565–570, 2004.
[33] G. Oxholm and K. Nishino, "Shape and reflectance from natural illumination," in *Proc. Eur. Conf. Comput. Vis.*, 2012, pp. 528–541.
[34] R. Ramamoorthi, D. Mahajan, and P. Belhumeur, "A first order analysis of lighting, shading and shadows," *ACM Trans. Graph.*, vol. 26, no. 1, pp. 2:1–2:21, 2007.
[35] I. Sato, T. Okabe, Q. Yu, and Y. Sato, "Shape reconstruction based on similarity in radiance changes under varying illumination," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2007, pp. 1–8.
[36] S. Seitz, B. Curless, J. Diebel, D. Scharstein, and R. Szeliski, "A comparison and evaluation of multiview stereo reconstruction algorithms," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, 2006, pp. 519–526.
[37] J. T. Todd, J. F. Norman, J. J. Koenderink, and A. M. L. Kappers, "Effects of texture, illumination and surface reflectance on stereoscopic shape perception," *Perception*, vol. 26, no. 7, pp. 807–822, 1997.
[38] A. Treuille, A. Hertzmann, and S. Seitz, "Example-based stereo with general BRDFs," in *Proc. Eur. Conf. Comput. Vis.*, 2004, pp. 457–469.
[39] A. Verri and T. Poggio, "Motion field and optical flow: Qualitative properties," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11, no. 5, pp. 490–498, May 1989.
[40] P. Woodham, "Photometric method for determining surface orientation from multiple images," *Opt. Eng.*, vol. 19, no. 1, pp. 139–144, 1980.
[41] C. Wu, S. Agarwal, B. Curless, and S. Seitz, "Schematic surface reconstruction," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, 2012, pp. 1498–1505.
[42] T. Zickler, P. Belhumeur, and D. Kriegman, "Helmholtz stereopsis: Exploiting reciprocity for surface reconstruction," *Int. J. Comput. Vis.*, vol. 49, no. 2/3, pp. 1215–1227, 2002.
[43] A. Zisserman, P. Giblin, and A. Blake, "The information available to a moving observer from specularities," *Image Vis. Comput.*, vol. 7, no. 1, pp. 38–42, 1989.

Manmohan Chandraker received the BTech degree in electrical engineering from the Indian Institute of Technology, Bombay, and the PhD degree in computer science from the University of California, San Diego. Following a postdoctoral scholarship at the University of California, Berkeley, he joined NEC Labs America in Cupertino, where he conducts research in computer vision. His principal research interests include modern optimization methods for geometric 3D reconstruction and theoretical analysis of shape recovery in the presence of complex illumination effects and material behavior. His work on provably optimal algorithms for structure and motion estimation has received the Marr Prize Honorable Mention for Best Paper at IEEE International Conference on Computer Vision and the CSE Dissertation Award for Best Thesis at UC San Diego. His work on shape recovery with complex materials has received the Best Paper Award at CVPR 2014.