What Camera Motion Reveals About Shape With Unknown BRDF

Manmohan Chandraker
NEC Labs America, Cupertino, CA

Abstract

Recent works 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 depth of a surface with unknown isotropic BRDF, which generalizes traditional Lambertian assumptions. Under orthographic projection, we show shape may not be constrained by differential stereo for general isotropic BRDFs, but two motions suffice to yield an invariant for several restricted (still unknown) BRDFs exhibited by common materials. For the perspective case, we show that three differential motions suffice to yield 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 outline with experiments how potential reconstruction methods may exploit our theory. We also illustrate trends shared by theories on shape from differential motion of light source, object or camera, to relate hardness of surface reconstruction to complexity of imaging setup.

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 [6, 8], photometric stereo for light source motion [17] and multiview stereo for motion of the camera [14]. 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. However, recent works have shown that differential motion of the light source [1] or the object [3] inform about shape even with unknown BRDFs. This paper solves the remaining problem of characterizing shape recovery for unknown BRDFs, using differential motion of the camera.

Section 3 of this paper proposes a physically valid differential stereo relation that relates 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. Surprisingly, it can be shown that considering a sequence of motions allows eliminating the BRDF dependence of differential stereo.

The mathematical basis for differential stereo is outwardly similar to differential flow for object motion [3]. However, while BRDFs are considered black-box functions in [3], 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 more difficult.

Consequently, for orthographic projection, Sec. 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, Sec. 5 shows that depth for a surface with unknown isotropic BRDF, under unknown directional lighting, may be obtained using differential stereo relations from three or more camera motions. Further, for the above restricted families of BRDFs, we show that an additional linear 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. Table 1 summarizes the main theoretical results of this paper.

Finally, Sec. 6 discusses relationships between theories on shape recovery from differential motion of the object, light source and camera. We explore shared traits among those theories that allow shape recovery and their common trends on the hardness of surface reconstruction. Those perspectives on shape from motion are summarized in Table 2.

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 [6, 8]. The limitations of the Lambertian assumption have been recognized by early works [10, 16].

Camera	Light	BRDF	#Motions	Surface Constraint	Shape Recovery	Theory
Persp.	Unknown	Lambertian	1	Linear eqn.	Depth	Stereo
Orth.	Unknown	General, unknown	-	Ambiguous for reconstruction		Prop. 3
Orth.	Unknown	View-angle, unknown	2	Homog. quasilinear PDE	Level curves	Rem. 1, Prop. 6
Orth.	Known	Half-angle, unknown	2	Inhomog. quasilinear PDE	Char. curves	Props. 4, 7
Persp.	Unknown	General, unknown	3	Linear eqn.	Depth	Prop. 8
Persp.	Unknown	View-angle, unknown	3	Lin. eqn. + Homog. lin. PDE	Depth + Gradient	Sec. 5.2
Persp.	Known	Half-angle, unknown	3	Lin. eqn. + Inhomog. lin. PDE	Depth + Gradient	Prop. 9

Table 1. Summary of the theoretical results of this paper. Note that k camera motions result in $k+1$ images. In general, more constrained BRDF or lighting yields richer reconstruction invariants.

In the context of stereo, several methods have been developed for non-Lambertian materials. Zickler et al. use the Helmholtz reciprocity principle for reconstruction with arbitrary BRDFs [18]. An example-based stereo that uses reference shapes of known geometry and various material types is presented in [15]. Stereo reconstructions for specular surfaces have been studied by Savarese et al. in [13]. In contrast, this paper explores how an image sequence derived from camera motion informs about shape with unknown isotropic BRDFs, regardless of reconstruction method.

Light source and object motions have also been used to understand shape with unknown BRDFs. A photometric stereo method based on small light source motions is presented by [4], while optical flow has been generalized to more general models by [5, 11]. 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 [12].

Closely related to this paper are the works of Chandraker et al. that derive topological classes up to which reconstruction can be performed for unknown BRDFs, using differential motion of the source [1] or object [3]. 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 Sec. 6.

3. Differential Stereo for 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

Assumptions and setup We assume static object and lighting, while the camera moves. For rigid body motion, our analysis equivalently considers a fixed camera, with the object and light source undergoing the inverse motion. The illumination is assumed directional and distant. The object BRDF is assumed to be homogeneous and isotropic, with unknown functional form. We make a technical assumption that the camera direction is constant over the entire object. [1] Global illumination effects are assumed negligible.

Let the focal length of the 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 $\beta = f^{-1}$, a 3D point $\mathbf{x} = (x, y, z)^\top$ is imaged at $\mathbf{u} = (u, v)^\top$, where

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

The camera is assumed to be geometrically calibrated.

Motion field 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 $\boldsymbol{\omega} = (\omega_1, \omega_2, \omega_3)^\top$.

The motion field, $\boldsymbol{\mu} = (\dot{u}, \dot{v})^\top$, is the differential motion of the image obtained by differentiating (1). We refer the reader to prior works like [10, 3] for a derivation and simply state here the motion field in a form similar to [3]:

$$\text{Perspective: } \boldsymbol{\mu} = \left(\alpha_1 + \frac{\alpha_2 + \omega_2 z}{1+\beta z}, \alpha_3 + \frac{\alpha_4 - \omega_1 z}{1+\beta z}\right)^\top, \tag{2}$$

$$\text{Orthographic: } \boldsymbol{\mu} = (\alpha_5 + \omega_2 z,\ \alpha_6 - \omega_1 z)^\top, \tag{3}$$

where α_i, $i = 1, \cdots, 6$ are known functions of $\boldsymbol{\omega}$, $\boldsymbol{\tau}$, $\mathbf{u}$ and β, whose algebraic forms are given by:

$$\alpha_1 = \omega_2 \beta u^2 - \omega_1 \beta u v - \omega_3 v, \qquad \alpha_2 = \tau_1 - \beta u \tau_3, \tag{4}$$

$$\alpha_3 = -\omega_1 \beta v^2 + \omega_2 \beta u v + \omega_3 u, \qquad \alpha_4 = \tau_2 - \beta v \tau_3, \tag{5}$$

$$\alpha_5 = \tau_1 - \omega_3 v, \qquad \alpha_6 = \tau_2 + \omega_3 u. \tag{6}$$

Differential stereo relation 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 surface, the unit surface normal is given by

[1] 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.

$$\mathbf{n} = (n_1, n_2, n_3)^\top = \frac{(z_x, z_y, -1)^\top}{\sqrt{z_x^2 + z_y^2 + 1}}, \quad (7)$$

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

$$I(\mathbf{u}, t) = \sigma(\mathbf{x})\rho(\mathbf{x}, \mathbf{n}, \mathbf{s}, \mathbf{v}), \quad (8)$$

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. We do not make any assumptions on the functional form of ρ, except smoothness.

Taking the *total* derivative on both sides of (8), we get

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

Since σ is intrinsically defined on the surface coordinates, its total derivative vanishes. For the rigid body motion we consider, $\dot{\mathbf{n}} = \boldsymbol{\omega} \times \mathbf{n}$ and $\dot{\mathbf{s}} = \boldsymbol{\omega} \times \mathbf{s}$, 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

$$(\nabla_{\mathbf{u}} I)^\top \boldsymbol{\mu} + I_t = \sigma \left[(\nabla_{\mathbf{x}}\rho)^\top \boldsymbol{\nu} + (\nabla_{\mathbf{n}}\rho)^\top (\boldsymbol{\omega} \times \mathbf{n}) + (\nabla_{\mathbf{s}}\rho)^\top (\boldsymbol{\omega} \times \mathbf{s}) \right] \quad (10)$$

where $\boldsymbol{\nu} = \dot{\mathbf{x}}$ is the linear velocity. While the above discussion gives intuition for the differential relation in (10), we refer the reader to Appendix A for a rigorous derivation.

For distant lighting and homogeneous reflectance in our setup, we may assume that $\nabla_{\mathbf{x}}\rho$ is negligible. Further, dividing the two sides of (10) with those of (8), we get

$$\boxed{(\nabla_{\mathbf{u}} E)^\top \boldsymbol{\mu} + E_t = (\mathbf{n} \times \nabla_{\mathbf{n}} \log \rho + \mathbf{s} \times \nabla_{\mathbf{s}} \log \rho)^\top \boldsymbol{\omega}} \quad (11)$$

where we use the notation $E = \log I$ and the identities $(\nabla_{\mathbf{a}}\rho)^\top (\boldsymbol{\omega} \times \mathbf{a}) = (\mathbf{a} \times \nabla_{\mathbf{a}}\rho)^\top \boldsymbol{\omega}$ and $\nabla_{\mathbf{a}} \log \rho = \rho^{-1}\nabla_{\mathbf{a}}\rho$, for some $\mathbf{a} \in \mathbb{R}^3$. We call (11) the *differential stereo relation*.

3.2. Understanding the Differential Stereo Relation

Generalization of Lambertian stereo Initial intuition into the differential stereo relation of (11) may be derived by noting how it generalizes traditional Lambertian stereo. For two images I^1 and I^2 related by a known motion, Lambertian stereo seeks the depth z that best satisfies brightness constancy: $I^1(\mathbf{u}) = I^2(\mathbf{u} + \boldsymbol{\mu}(z))$. Substituting a Lambertian reflectance $\rho(\mathbf{n}, \mathbf{s}, \mathbf{v}) = \mathbf{n}^\top \mathbf{s}$ in (11), we get

$$\begin{aligned}(\nabla_{\mathbf{u}} E)^\top \boldsymbol{\mu} + E_t &= \left(\mathbf{n} \times (\mathbf{n}^\top \mathbf{s})^{-1}\mathbf{s} + \mathbf{s} \times (\mathbf{n}^\top \mathbf{s})^{-1}\mathbf{n}\right)^\top \boldsymbol{\omega} \\ &= \mathbf{0}^\top \boldsymbol{\omega} = 0, \end{aligned} \quad (12)$$

which is precisely brightness constancy (total derivative of image intensity is zero). Thus, diffuse photoconsistency imposed by traditional stereo is a special case of our theory.

Relation to object motion For object motion, a differential flow relation is derived in [3]. The differential stereo relation of (11) has an additional dependence on BRDF derivatives with respect to the light source. This stands to reason, since both the object and lighting move relative to camera in the case of camera motion, while only the object moves in the case of object motion. Thus, intuitively, camera motion leads to a harder reconstruction problem. Indeed, as we will see next, the dependence of (11) on lighting leads to a somewhat surprising additional ambiguity.

Ambiguity for camera motion Now we derive some additional insights by making a crucial departure from the analysis of [3]. Namely, instead of treating isotropic BRDFs as entirely black box functions, we explicitly consider their physical property in the form of angular dependencies between the normal, light source and camera directions. We define

$$\boldsymbol{\pi} = \mathbf{n} \times \nabla_{\mathbf{n}} \log \rho + \mathbf{s} \times \nabla_{\mathbf{s}} \log \rho, \quad (13)$$

which allows rewriting the differential stereo relation (11) as:

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

The entity $\boldsymbol{\pi}$ is central to our theory, since it captures the dependence of differential stereo on BRDF. 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.

This definition of $\boldsymbol{\pi}$ leads to an observation intrinsic to shape recovery with isotropic BRDFs:

Proposition 1. *The BRDF dependence of differential stereo is captured by a 2D vector in the principal plane of the camera.*

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

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

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 (13) as

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

From the form of $\boldsymbol{\pi}$ in (16), it is evident that $\boldsymbol{\pi}^\top \mathbf{v} = 0$. For our choice of coordinate system, $\mathbf{v} = (0, 0, -1)^\top$. Thus,

$$\pi_3 = 0. \quad (17)$$

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 $\{z, \boldsymbol{\pi}\}$. 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 exists for camera motion too.

Under orthography, the motion field $\boldsymbol{\mu}$ is given by (3). Noting the 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 (3), 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 $i = 1, \cdots, m$. Let $E^0 = \log I^0$ be the logarithm of the base image, with E^i being 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 $\mathbf{u}$. We will simply denote it 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.

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 $\mathbf{q} = (q^1, \cdots, q^m)^\top$. 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. However, the result of Proposition 1 makes this 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 Prop. 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 [3] 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 (13) on BRDF-derivative with respect to lighting makes it indeterminate without further restrictions on BRDF or lighting.

While Prop. 3 is a negative result, its development provides valuable insight. An $m \times 3$ rank-deficient matrix $\tilde{\mathbf{A}}$ 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 Prop. 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 $\boldsymbol{\pi}$. Note that the functional form of the BRDF, $\rho(\cdot)$, remains unknown.

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 [9]. 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 (13):

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

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 (7) 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 that depends 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 (13) 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 $\boldsymbol{\pi}$ with respect to $\mathbf{n}$ and $\mathbf{s}$ means it is independent of ρ_θ. Now, 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 (7) 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

$$\lambda_1 = a_1(\gamma_3 - E_u \gamma_1) - b_1(\gamma_2 + E_v \gamma_1) \tag{40}$$
$$\lambda_2 = a_1 E_u + b_1 E_v \tag{41}$$
$$\lambda_3 = a_2(\gamma_3 - E_u \gamma_1) - b_2(\gamma_2 + E_v \gamma_1) \tag{42}$$
$$\lambda_4 = a_2 E_u + b_2 E_v \tag{43}$$
$$\lambda_5 = -a_3(\gamma_3 - E_u \gamma_1) + b_3(\gamma_2 + E_v \gamma_1). \tag{44}$$

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 Prop. 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, Prop. 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, $\boldsymbol{\pi}^\top \mathbf{v} = -\pi_3 = 0$ 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 [2]. In such cases, we may write

$$\log \rho(\mathbf{n}, \mathbf{s}, \mathbf{v}) = \bar{\rho}(\mathbf{n}^\top \mathbf{s}, \mathbf{n}^\top \mathbf{y}), \text{ with } \mathbf{y} = \frac{\mathbf{s} + \kappa \mathbf{v}}{\|\mathbf{s} + \kappa \mathbf{v}\|}, \quad (45)$$

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 [2]. 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 for complete details. We only note here that we obtain:

$$\boldsymbol{\pi} = \frac{\kappa \rho_{\mathbf{n}^\top \mathbf{y}}}{1 + \kappa^2 + 2\kappa\phi} \left[(1 + \kappa^2 + 2\kappa\phi)^{\frac{1}{2}} \mathbf{n} - (\mathbf{n}^\top \mathbf{y}) \mathbf{s} \right] \times \mathbf{v}, \quad (46)$$

thus, dependence on ρ may be eliminated similar to (34) using $\pi_3 = 0$ and considering 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:

$$(\lambda'_1 + \lambda'_2 z) z_x + (\lambda'_3 + \lambda'_4) z_y + \lambda'_5 = 0, \quad (47)$$

where λ'_i, for $i = 1, \cdots, 5$, are again known entities. □

We urge the reader to observe that the constraint (47) 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 [2]. Also, we again note that Proposition 1 is an intrinsic property of isotropic BRDFs and $\pi_3 = 0$ continues to hold even for the $\boldsymbol{\pi}$ in (46).

4.3. Results on Surface Estimation

Recall that Sec. 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, Sec. 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* [7]. For a surface level curve $z(x, y) = z_0$, the solution to (28) from PDE theory is:

$$z = z_0, \quad \frac{dy}{dx} = \frac{(\gamma_3 - E_u \gamma_1) + E_u z_0}{(\gamma_2 + E_v \gamma_1) - E_v z_0}. \quad (48)$$

That is, given the depth at a point, the ODE (48) 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 [7]. Here, we again state the result while referring the reader to Appendix C.2 for a proof:

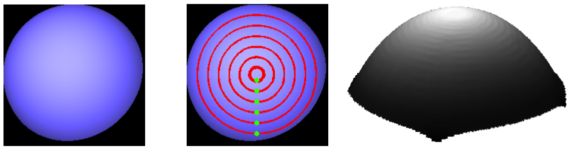
Proposition 7. *Under orthography, for a surface with half-angle BRDF, two or more differential camera motions yield characteristic surface curves $\mathcal{C}(x(s), y(s), z(s))$ defined by*

$$\frac{1}{\lambda_1 + \lambda_2 z} \frac{dx}{ds} = \frac{1}{\lambda_3 + \lambda_4 z} \frac{dy}{ds} = \frac{-1}{\lambda_5} \frac{dz}{ds} \quad (49)$$

corresponding to depths at some (possibly isolated) points.

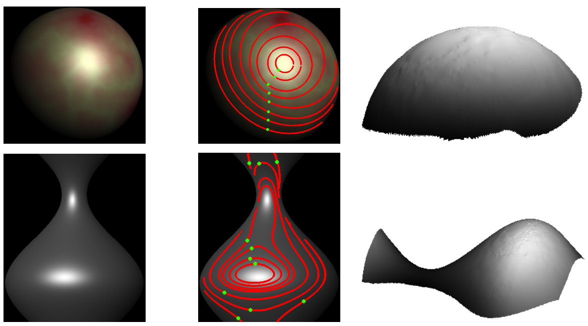
Finally, for a BRDF dependent on an arbitrary angle in the source-view plane given by (45), the invariant (47) is also an inhomogeneous quasilinear PDE. So, as a direct generalization of Prop. 7, differential stereo also yields characteristic curves for this case (with λ'_i instead of λ_i).

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. 1 for a BRDF that depends on source and view angles (unknown lighting) and in Fig. 2 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.



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

Figure 1. (a) One of three simulated orthographic images (two motions), with unknown lighting and arbitrary non-Lambertian BRDF dependent on source and view angles. (b) Level curves estimated using Prop. 6. (c) Surface reconstructed after interpolation.



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

Figure 2. (a) One of three synthetic images (two motions), with unknown half-angle BRDF and known lighting, under orthographic projection. (b) Characteristic curves estimated using Prop. 7. (c) Surface reconstructed after interpolation.

5. Perspective Projection

In this section, we consider shape recovery from differential stereo under perspective projection. In particular, we show that unlike the orthographic case, depth may be unambiguously recovered in the perspective case, even when both the BRDF and lighting are unknown.

5.1. Depth from Differential Stereo

In the perspective case, the motion field $\boldsymbol{\mu}$ is given by (2). Substituting for $\boldsymbol{\mu}$ in the differential stereo relation (11), we obtain its expression for the perspective case:

$$p'\left(\frac{z}{1+\beta z}\right) + r'\left(\frac{1}{1+\beta z}\right) + q' \;=\; \boldsymbol{\omega}^\top \boldsymbol{\pi}, \qquad (50)$$

where $\boldsymbol{\pi}$ is defined as before by (13) and $p' = E_u\omega_2 - E_v\omega_1$, $q' = \alpha_1 E_u + \alpha_3 E_v + E_t$ and $r' = \alpha_2 E_u + \alpha_4 E_v$ are known entities. 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 $E_1, \cdots, E_m$ be related to a base image E_0 by known differential motions $\{\boldsymbol{\omega}^i, \boldsymbol{\tau}^i\}$. Then, we obtain from (50) a sequence of differential stereo relations:

$$(p'^i + \beta q'^i)z - ((1+\beta z)\boldsymbol{\pi})^\top \boldsymbol{\omega}^i + (q'^i + r'^i) = 0, \quad (51)$$

for $i = 1, \cdots, m$. Let $\tilde{\mathbf{C}} = \left[\mathbf{c}^1, \cdots, \mathbf{c}^m\right]^\top$ 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 $\mathbf{q}' = [q'^1, \cdots, q'^m]^\top$ and $\mathbf{r}' = [r'^1, \cdots, r'^m]^\top$. Then, the system of differential stereo relations (51) may be written as

$$\tilde{\mathbf{C}} \begin{bmatrix} z \\ (1+\beta z)\pi_1 \\ (1+\beta z)\pi_2 \end{bmatrix} = \mathbf{q}' + \mathbf{r}', \qquad (52)$$

since $\pi_3 = 0$, from Prop. 1. With $\tilde{\mathbf{C}}^+$ as 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, (52) has the solution:

$$[z,\, (1+\beta z)\pi_1,\, (1+\beta z)\pi_2]^\top \;=\; \boldsymbol{\epsilon} \qquad (53)$$

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

Thus, Prop. 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 [3]. With object motion, depth and an additional constraint on the gradient are recovered with unknown BRDF and lighting. While camera motion also recovers depth, following Prop. 1, it is likely that further information on the gradient is recoverable only with additional constraints on the BRDF or lighting.

5.2. Additional Constraints for Certain Materials

As indicated above, now we 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.*

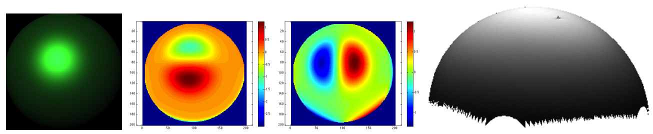
Proof. Prop. 8 already shows depth recovery from $m \geq 3$ differential motions of the camera. Further, from the form of $\boldsymbol{\pi}$ for half-angle BRDFs in (33) and using (53), we have

$$\frac{\pi_1}{\pi_2} = \frac{\mathbf{a}^\top \mathbf{n}}{\mathbf{b}^\top \mathbf{n}} = \frac{\epsilon_2}{\epsilon_3}, \qquad (54)$$

where $\mathbf{a}$ and $\mathbf{b}$ are known entities dependent on light, defined by (34). Finally, using (7) yields a linear PDE in depth:

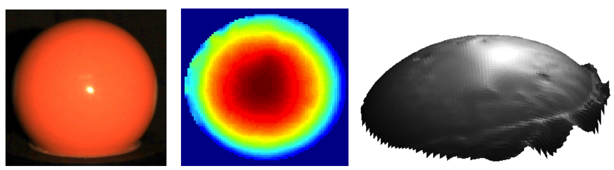
$$l_1 z_x + l_2 z_y + l_3 = 0, \qquad (55)$$

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. □



(a) Image (b) π_1 (c) π_2 (d) Depth

Figure 3. (a) One of four synthetic images (three motions), with arbitrary non-Lambertian BRDF and unknown lighting, under perspective projection. (b,c) π_1 and π_2 recovered using (57), as discussed in Sec. 6. (d) Depth estimated using Prop. 8.



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

Figure 4. Reconstruction for a non-Lambertian ball with an unknown BRDF, under unknown light source, using Proposition 8.

It is evident from the above result 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, since their ratio is a measurable quantity from (54). 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.*

5.3. Shape Recovery

From Prop. 8, depth is directly available in the perspective case for unknown lighting and unknown BRDF. Fig. 3 illustrates this for synthetic data, where the object is imaged under an unknown light source. The object has diameter 20cm and is placed 1.5m away from the camera of focal length 10cm. Three random motions, of approximately $2°$ rotation and 5mm translation are imparted to the camera. The recovered shape using the theory of Sec. 5.1 is shown in Fig. 3. No prior knowledge of the BRDF, lighting or depth are required.

For evaluation with real data, images of the plastic sphere in Fig. 4(a) are obtained against a textured background under unknown directional lighting. Note the clear non-Lambertian effects. The camera focal length is 55mm and the object is approximately 80cm away. Five small motions (and several large ones for stable pose optimization) are imparted to the camera and estimated using bundle adjustment. The surface depth obtained from Prop. 8 is shown in Fig. 4(b) and (c).

While we do not explore the direction in this paper, note that with known lighting, one may use the constraint (55) to solve a joint depth and gradient optimization:

$$\min_z \, (z - \epsilon_1)^2 \; + \; \lambda(l_1 z_x + l_2 z_y + l_3)^2, \qquad (56)$$

where λ is a relative weight. With standard differencing, the above is a highly sparse linear system in z which may be solved efficiently.

6. Perspectives on Shape from Motion Theories

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 [1] and object [3] motions. This paper has studied camera motion to complete the differential analysis framework for shape recovery with unknown BRDFs.

In Table 2, 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 colocated and 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

Motion	Camera	Light	BRDF	#Motions	Surface Constraint	Shape Recovery	Theory
Light	Orth.	Known	Lambertian	2	Lin. eqns. on ∇z (Photo. stereo)	Gradient	[17]
Light	Orth.	Unknown	General, unknown	2	Linear PDE	Level cur. + $\|\nabla z\|$ isocontours	[1]
Object	Orth.	Colocated	Uni-angle, unknown	2	Homog. quasilin. PDE	Level curves	[3]
Object	Orth.	Unknown	Full, unknown	3	Inhomog. quasilin. PDE	Char. curves	[3]
Object	Persp.	Brightness constancy		1	Lin. eqn. (optical flow)	Depth	[6, 8]
Object	Persp.	Colocated	Uni-angle, unknown	3	Lin. eqn. + Homog. lin. PDE	Depth + Gradient	[3]
Object	Persp.	Unknown	Full, unknown	4	Lin. eqn. + Inhomog. lin. PDE	Depth + Gradient	[3]
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	[14]
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

Table 2. A unified look at frameworks on general BRDF shape recovery from differential motions of light source, object or camera. 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 reconstruction invariant. For comparison, the traditional diffuse equivalents are also shown.

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.

Future work Our theory focuses on shape recovery, but some BRDF information may also be recovered as a by-product. For instance, with perspective projection, Props. 8 and 1 completely define $\boldsymbol{\pi}$. That is, from (53),

$$\boldsymbol{\pi} = \frac{1}{1+\beta\epsilon_1}\,(\epsilon_2, \epsilon_3, 0)^\top. \tag{57}$$

Fig. 3 shows an example recovery of $\boldsymbol{\pi}$. In turn, this places constraints on the derivative of BRDF. An avenue for future work is to characterize the extents to which BRDF may be recovered using motions of the light source, object or camera.

Further, while [1, 3] and this paper together 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. Such a framework is the subject of our ongoing work.

A. Derivation of the Differential Stereo Relation

Section 3 provides an intuitive development of the differential stereo relation of (11). 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 roation $\mathbf{R} = \widetilde{\mathbf{R}}^{-1}$ 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 [10, 16] 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 $\mathbf{u} = (x, y)^\top$ 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

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

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

$$\delta\mathbf{x} = \boldsymbol{\nu}_t(a - \delta a, b - \delta b)\delta t. \tag{59}$$

Similarly, the unit surface normals corresponding to the image point $\mathbf{u} = (x, y)^\top$ at times t and $t + \delta t$ are related by

$$\mathbf{n}_{t+\delta t}(a - \delta a, b - \delta b) = \mathbf{n}_t(a - \delta a, b - \delta b) + \delta\mathbf{n}. \tag{60}$$

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,

$$\delta\mathbf{n} = \boldsymbol{\omega} \times \mathbf{n}_t(a - \delta a, b - \delta b)\delta t. \tag{61}$$

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 $\mathbf{u} = (x, y)^\top$ at times t and $t + \delta t$ are related by

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

Thus, we have defined the following differential relations:

$$\mathbf{x}_{t+\delta t}(a - \delta a, b - \delta b) = \mathbf{x}_t(a - \delta a, b - \delta b) + \boldsymbol{\nu}_t \delta t, \tag{63}$$

$$\mathbf{n}_{t+\delta t}(a - \delta a, b - \delta b) = \mathbf{n}_t(a - \delta a, b - \delta b) + \boldsymbol{\omega} \times \mathbf{n}_t \delta t, \tag{64}$$

$$\mathbf{s}_{t+\delta t}(a - \delta a, b - \delta b) = \mathbf{s}_t(a - \delta a, b - \delta b) + \boldsymbol{\omega} \times \mathbf{s}_t \delta t. \tag{65}$$

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:

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

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:

$$I(\mathbf{u}, t + \delta t) = \sigma(a - \delta a, b - \delta b)\rho\left(\mathbf{x}_{t+\delta t}, \mathbf{n}_{t+\delta t}, \mathbf{s}_{t+\delta t}, \mathbf{v}\right), \tag{67}$$

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:

$$I(\mathbf{u} - \delta \mathbf{u}, t) = \sigma(a - \delta a, b - \delta b)\rho\left(\mathbf{x}_t, \mathbf{n}_t, \mathbf{s}_t, \mathbf{v}\right). \tag{68}$$

with all entities in ρ again evaluated at $(a - \delta a, b - \delta b)$. Subtracting (68) from (67), we have

$$I(\mathbf{u}, t + \delta t) - I(\mathbf{u} - \delta \mathbf{u}, t) = \sigma\left[\ \rho\left(\mathbf{x}_{t+\delta t}, \mathbf{n}_{t+\delta t}, \mathbf{s}_{t+\delta t}, \mathbf{v}\right) - \rho\left(\mathbf{x}_t, \mathbf{n}_t, \mathbf{s}_t, \mathbf{v}\right)\ \right] \tag{69}$$

Applying chain-rule differentiation and using the differential entities defined in (63), (64) and (65), the above may be rewritten as:

$$I(\mathbf{u}, t + \delta t) - I(\mathbf{u} - \delta \mathbf{u}, t) = \sigma\left[\ (\nabla_{\mathbf{x}}\rho)^\top \boldsymbol{\nu}_t \delta t + (\nabla_{\mathbf{n}}\rho)^\top(\boldsymbol{\omega} \times \mathbf{n}_t)\delta t + (\nabla_{\mathbf{s}}\rho)^\top(\boldsymbol{\omega} \times \mathbf{s}_t)\delta t\ \right], \tag{70}$$

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 $I(\mathbf{u}, t)$:

$$(\nabla_{\mathbf{u}} I)^\top \delta \mathbf{u} = I(\mathbf{u}, t) - I(\mathbf{u} - \delta \mathbf{u}, t) \tag{71}$$

$$\frac{\partial I}{\partial t}\delta t = I(\mathbf{u}, t + \delta t) - I(\mathbf{u}, t). \tag{72}$$

Substituting both the above definitions into (70), we obtain

$$\frac{\partial I}{\partial t}\delta t = I(\mathbf{u}, t + \delta t) - I(\mathbf{u}, t) \tag{73}$$

$$= I(\mathbf{u} - \delta \mathbf{u}, t) + \sigma\left[\ (\nabla_{\mathbf{x}}\rho)^\top \boldsymbol{\nu}_t \delta t + (\nabla_{\mathbf{n}}\rho)^\top(\boldsymbol{\omega} \times \mathbf{n}_t)\delta t + (\nabla_{\mathbf{s}}\rho)^\top(\boldsymbol{\omega} \times \mathbf{s}_t)\delta t\ \right] - I(\mathbf{u}, t) \tag{74}$$

$$= -(\nabla_{\mathbf{u}} I)^\top \delta \mathbf{u} + \sigma\left[\ (\nabla_{\mathbf{x}}\rho)^\top \boldsymbol{\nu}_t \delta t + (\nabla_{\mathbf{n}}\rho)^\top(\boldsymbol{\omega} \times \mathbf{n}_t)\delta t + (\nabla_{\mathbf{s}}\rho)^\top(\boldsymbol{\omega} \times \mathbf{s}_t)\delta t\ \right]. \tag{75}$$

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

$$\boldsymbol{\mu} = \frac{\delta \mathbf{u}}{\delta t}. \tag{76}$$

Then, using (76), the relation in (75) can be written as

$$(\nabla_{\mathbf{u}} I)^\top \boldsymbol{\mu} + \frac{\partial I}{\partial t} = \sigma\left[\ (\nabla_{\mathbf{x}}\rho)^\top \boldsymbol{\nu} + (\nabla_{\mathbf{n}}\rho)^\top(\boldsymbol{\omega} \times \mathbf{n}) + (\nabla_{\mathbf{s}}\rho)^\top(\boldsymbol{\omega} \times \mathbf{s})\ \right]. \tag{77}$$

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 (11).

B. 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

$$\log \rho(\mathbf{n}, \mathbf{s}, \mathbf{v}) = \bar{\rho}(\mathbf{n}^\top \mathbf{s}, \mathbf{n}^\top \mathbf{y}), \text{ with } \mathbf{y} = \frac{\mathbf{s} + \kappa \mathbf{v}}{\|\mathbf{s} + \kappa \mathbf{v}\|}, \tag{78}$$

where $\kappa \in \mathbb{R}$. Recall the definition of $\boldsymbol{\pi}$ in (13), which may be rewritten as:

$$\boldsymbol{\pi} = \mathbf{n} \times \nabla_{\mathbf{n}}\bar{\rho} + \mathbf{s} \times \nabla_{\mathbf{s}}\bar{\rho}. \tag{79}$$

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 (78) and applying chain-rule differentiation, we obtain:

$$\mathbf{n} \times \nabla_{\mathbf{n}}\bar{\rho} = \bar{\rho}_\theta(\mathbf{n} \times \mathbf{s}) + \bar{\rho}_\eta \frac{\mathbf{n} \times \mathbf{s}}{\|\mathbf{s} + \kappa \mathbf{v}\|} + \kappa \bar{\rho}_\eta \frac{\mathbf{n} \times \mathbf{v}}{\|\mathbf{s} + \kappa \mathbf{v}\|}, \tag{80}$$

$$\mathbf{s} \times \nabla_{\mathbf{s}}\bar{\rho} = \bar{\rho}_\theta(\mathbf{s} \times \mathbf{n}) + \bar{\rho}_\eta \frac{\mathbf{s} \times \mathbf{n}}{\|\mathbf{s} + \kappa \mathbf{v}\|} - \kappa \bar{\rho}_\eta \frac{(\mathbf{n}^\top(\mathbf{s} + \kappa \mathbf{v}))\mathbf{s} \times \mathbf{v}}{\|\mathbf{s} + \kappa \mathbf{v}\|^3}. \tag{81}$$

Adding (80) and (81) and substituting in (79), we obtain:

$$\boldsymbol{\pi} = \frac{\kappa \bar{\rho}_\eta}{(1 + \kappa^2 + 2\kappa\phi)^{\frac{3}{2}}} \left[(1 + \kappa^2 + 2\kappa\phi)\mathbf{n} - (\mathbf{n}^\top \mathbf{s} + \kappa \mathbf{n}^\top \mathbf{v})\mathbf{s}\right] \times \mathbf{v}. \tag{82}$$

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

$$\frac{\pi_1}{\pi_2} = \frac{\mathbf{a}'^\top \mathbf{n}}{\mathbf{b}'^\top \mathbf{n}} = \frac{a'_1 z_x + a'_2 z_y - a'_3}{b'_1 z_x + b'_2 z_y - b'_3}, \tag{83}$$

where we have used the relationship between surface normal and gradient given by (7) and denoted

$$a'_1 = s_1 s_2, \qquad b'_1 = (1 + \kappa^2 + 2\kappa\phi) - s_1^2, \tag{84}$$

$$a'_2 = s_2^2 - (1 + \kappa^2 + 2\kappa\phi), \qquad b'_2 = -s_1 s_2, \tag{85}$$

$$a'_3 = s_2(s_3 - \kappa), \qquad b'_3 = -s_1(s_3 - \kappa). \tag{86}$$

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:

$$\frac{\pi_1}{\pi_2} = \frac{(\gamma_2 + E_v \gamma_1) - E_v z}{(\gamma_3 - E_u \gamma_1) + E_u z}, \tag{87}$$

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

$$\frac{a'_1 z_x + a'_2 z_y - a'_3}{b'_1 z_x + b'_2 z_y - b'_3} = \frac{(\gamma_2 + E_v \gamma_1) - E_v z}{(\gamma_3 - E_u \gamma_1) + E_u z}, \tag{88}$$

which may be rewritten as

$$(\lambda'_1 + \lambda'_2 z) z_x + (\lambda'_3 + \lambda'_4) z_y + \lambda'_5 = 0, \tag{89}$$

where

$$\lambda'_1 = a'_1(\gamma_3 - E_u \gamma_1) - b'_1(\gamma_2 + E_v \gamma_1) \tag{90}$$

$$\lambda'_2 = a'_1 E_u + b'_1 E_v \tag{91}$$

$$\lambda'_3 = a'_2(\gamma_3 - E_u \gamma_1) - b'_2(\gamma_2 + E_v \gamma_1) \tag{92}$$

$$\lambda'_4 = a'_2 E_u + b'_2 E_v \tag{93}$$

$$\lambda'_5 = -a'_3(\gamma_3 - E_u \gamma_1) + b'_3(\gamma_2 + E_v \gamma_1). \tag{94}$$

We have now obtained the constraint (47) that relates the surface depth z to image derivatives and is independent of the BRDF.

C. Surface Estimation Under Orthography

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

C.1. Proof of Proposition 6

Proposition 6 shows that for a surface with unknown BRDF dependent on light 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 source and view angles, Remark 1 stipulates a BRDF-invariant constraint that relates surface shape to image derivatives. In particular, we have from (28):

$$\frac{z_x}{z_y} = -\frac{(\gamma_3 - E_u \gamma_1) + E_u z}{(\gamma_2 + E_v \gamma_1) - E_v z}. \tag{95}$$

Note that (95) 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

$$\mathbf{a} = ((\gamma_2 + E_v \gamma_1) - E_v z, (\gamma_3 - E_u \gamma_1) + E_u z)^\top. \tag{96}$$

Then, from (95), we have that

$$\mathbf{a}^\top \nabla z = 0. \tag{97}$$

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

$$\frac{dz}{ds} = \frac{\partial z}{\partial x}\frac{dx}{ds} + \frac{\partial z}{\partial y}\frac{dy}{ds}. \tag{98}$$

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 $\mathbf{b} = \left(\frac{dx}{ds}, \frac{dy}{ds}\right)$. Then, we also have

$$\mathbf{b}^\top \nabla z = 0. \tag{99}$$

From (97) and (99), 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 (95):

$$\frac{dy}{dx} = \frac{(\gamma_3 - E_u \gamma_1) + E_u z}{(\gamma_2 + E_v \gamma_1) - E_v z}. \tag{100}$$

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

$$z = z_0 \tag{101}$$

$$\frac{dy}{dx} = \frac{(\gamma_3 - E_u \gamma_1) + E_u z_0}{(\gamma_2 + E_v \gamma_1) - E_v z_0}. \tag{102}$$

Now, given the value of z at any point on the surface, we can use the ODE in (102) to determine all other points on the surface with the same value of z. Thus, (95) allows reconstruction of level curves of the surface, with unknown BRDF and unknown light source.

C.2. Proof of Proposition 7

Proposition 7 states that for a BRDF of unknown functional form that depend on the half-angle, two or more differential camera motions yield characteristic surface curves given by (49).

Consider the PDE in (39), given by

$$(\lambda_1 + \lambda_2 u)u_x + (\lambda_3 + \lambda_4 u)u_y + \lambda_5 = 0, \qquad (103)$$

where λ_i, for $i = 1, \cdots, 5$, are known functions of (x, y) given by (40)-(44). We note that (103) is a first-order, inhomogeneous, quasilinear PDE, which may again be solved using standard constructs from PDE theory.

We established in Section 4.2.2 that our surface of interest is the integral surface of PDE (103), denoted as $\mathcal{S} : z = u(x, y)$. It is also shown in Section 4.2.2 that the coefficient functions λ_i, for $i = 1, \cdots, 5$, can be obtained from two or more differential motions of the camera. We now rewrite (103) in the form

$$\mathbf{a}^\top \begin{bmatrix} \nabla u \\ -1 \end{bmatrix} = 0 \qquad (104)$$

where $\mathbf{a} = (\lambda_1 + \lambda_2 u,\ \lambda_3 + \lambda_4 u,\ -\lambda_5)^\top$. 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 (49), represented with a parameter $s \in \mathbf{I} \subset \mathbb{R}$:

$$\mathcal{C} :\ x = x(s),\ y = y(s),\ z = z(s), \qquad (105)$$

where

$$\frac{dx}{ds} = \lambda_1 + \lambda_2 z = a_1(x, y, z) \qquad (106)$$

$$\frac{dy}{ds} = \lambda_3 + \lambda_4 z = a_2(x, y, z) \qquad (107)$$

$$\frac{dz}{ds} = -\lambda_5 = a_3(x, y, z)\,, \qquad (108)$$

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

$$x_0 = x(s_0),\ y_0 = y(s_0),\ z_0 = z(s_0) = u(x_0, y_0). \qquad (109)$$

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

$$w = w(s) = z(s) - u(x(s), y(s)). \qquad (110)$$

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

$$\frac{dw}{ds} = u_x a_1(x, y, w + u) + u_y a_2(x, y, w + u) + a_3(x, y, w + u) \qquad (111)$$

$$w(s_0) = 0. \qquad (112)$$

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 (103). Also, the solution to (111) must be unique. Thus, we have $z(s) = u(x(s), y(s))$, which establishes that $\mathcal{C} \subset \mathcal{S}$. This completes the proof that the characteristic curves $\mathcal{C}$, given by (49), reside on the surface $\mathcal{S}$.

D. Full Derivation Including $\nabla_{\mathbf{x}}\rho$

To derive the differential stereo relation in (11), we make an assumption that $\nabla_{\mathbf{x}}\rho$ is negligible. While the assumption is reasonable for our setup, we note that it is not a necessary requirement for the shape recovery results to hold. We now show that shape may be recovered using motion for a surface with unknown BRDF, even if $\nabla_{\mathbf{x}}\rho$ is not assumed to be negligible.

Differential Stereo Relation We start with the relation in (10), at which point no approximations have been made:

$$(\nabla_{\mathbf{u}}E)^\top \boldsymbol{\mu} + E_t = (\nabla_{\mathbf{x}}\rho)^\top \boldsymbol{\nu} + (\nabla_{\mathbf{n}}\rho)^\top (\boldsymbol{\omega} \times \mathbf{n}) + (\nabla_{\mathbf{s}}\rho)^\top (\boldsymbol{\omega} \times \mathbf{s}), \qquad (113)$$

where $\boldsymbol{\nu} = \dot{\mathbf{x}} = \boldsymbol{\omega} \times \mathbf{x} + \boldsymbol{\tau}$ is the linear velocity. Then, we rewrite the above as

$$(\nabla_{\mathbf{u}}E)^\top \boldsymbol{\mu} + E_t = (\mathbf{n} \times \nabla_{\mathbf{n}}\rho + \mathbf{s} \times \nabla_{\mathbf{s}}\rho + \mathbf{x} \times \nabla_{\mathbf{x}}\rho)^\top \boldsymbol{\omega} + (\nabla_{\mathbf{x}}\rho)^\top \boldsymbol{\tau}. \qquad (114)$$

Let us denote

$$\boldsymbol{\pi}' = \mathbf{n} \times \nabla_{\mathbf{n}}\rho + \mathbf{s} \times \nabla_{\mathbf{s}}\rho + \mathbf{x} \times \nabla_{\mathbf{x}}\rho \qquad (115)$$

$$= \boldsymbol{\pi} + \mathbf{x} \times \nabla_{\mathbf{x}}\rho, \qquad (116)$$

where $\boldsymbol{\pi} = \mathbf{n} \times \nabla_{\mathbf{n}}\rho + \mathbf{s} \times \nabla_{\mathbf{s}}\rho$ is the same as defined in (13). Then, the basic relation is of the form:

$$(\nabla_{\mathbf{u}}E)^\top \boldsymbol{\mu} + E_t = \boldsymbol{\pi}'^\top \boldsymbol{\omega} + (\nabla_{\mathbf{x}}\rho)^\top \boldsymbol{\tau}. \qquad (117)$$

Note that we have made no approximations to obtain the above relation, it is just a rewriting of (10).

Constraints from an Image Sequence Let us now substitute for the motion field in (117) using (2), to obtain:

$$p'\left(\frac{z}{1 + \beta z}\right) + r'\left(\frac{1}{1 + \beta z}\right) + q' = \boldsymbol{\omega}^\top \boldsymbol{\pi}' + \boldsymbol{\tau}^\top \nabla_{\mathbf{x}}\rho, \qquad (118)$$

which is the counterpart of (50) without considering $\nabla_{\mathbf{x}}\rho$ as negligible. For ease of reference, we list the forms of p', q' and r':

$$p' = E_u\omega_2 - E_v\omega_1, \qquad (119)$$

$$q' = \alpha_1 E_u + \alpha_3 E_v + E_t$$
$$= \beta(u\omega_2 - v\omega_1)(uE_u + vE_v) + (uE_v - vE_u)\omega_3 + E_t, \qquad (120)$$

$$r' = \alpha_2 E_u + \alpha_4 E_v$$
$$= E_u\tau_1 + E_v\tau_2 - \beta(uE_u + vE_v)\tau_3. \qquad (121)$$

Given observations from $m \geq 6$ motions, we arrange relations of the form (118) into a linear system:

$$\begin{bmatrix} p'^1 & r'^1 & -\boldsymbol{\omega}^{1\top} & -\boldsymbol{\tau}^{1\top} \\ & \vdots & & \\ p'^m & r'^m & -\boldsymbol{\omega}^{m\top} & -\boldsymbol{\tau}^{m\top} \end{bmatrix} \begin{bmatrix} \frac{z}{1+\beta z} \\ \frac{1}{1+\beta z} \\ \boldsymbol{\pi}' \\ \nabla_{\mathbf{x}}\rho \end{bmatrix} = -\begin{bmatrix} q'^1 \\ \vdots \\ q'^m \end{bmatrix}. \qquad (122)$$

Let us denote the above $m \times 8$ matrix as $\mathbf{B}$ and define $\mathbf{q} = ({q'}^1, \cdots, {q'}^m)^\top$. We assume that the rotations and translations span $\mathbb{R}^3$. Then, it can be observed from the forms of p' and r' that $\text{rank}(\mathbf{B}) = 6$. 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 above linear system has solutions of the form:

$$\begin{bmatrix} \frac{z}{1+\beta z} \\ \frac{1}{1+\beta z} \\ \boldsymbol{\pi}' \\ \nabla_{\mathbf{x}}\rho \end{bmatrix} = \boldsymbol{\gamma} + \lambda_1 \begin{bmatrix} 1 \\ 0 \\ \mathbf{g}_1 \\ \mathbf{0}_{3\times1} \end{bmatrix} + \lambda_2 \begin{bmatrix} 0 \\ 1 \\ \mathbf{0}_{3\times1} \\ \mathbf{g}_2 \end{bmatrix}, \quad (123)$$

where

$$\mathbf{g}_1 = \begin{bmatrix} g_{11} \\ g_{12} \\ g_{13} \end{bmatrix} = \begin{bmatrix} -E_v \\ E_u \\ 0 \end{bmatrix}, \quad (124)$$

$$\mathbf{g}_2 = \begin{bmatrix} g_{21} \\ g_{22} \\ g_{23} \end{bmatrix} = \begin{bmatrix} E_u \\ E_v \\ -\beta(uE_u + vE_v) \end{bmatrix}. \quad (125)$$

Surface Depth Recovery Now, we can derive several relations from the above solution. First, we observe from (123) that

$$\lambda_1 = \frac{z}{1+\beta z} - \gamma_1, \quad (126)$$

$$\lambda_2 = \frac{1}{1+\beta z} - \gamma_2. \quad (127)$$

Next, using (123) and (127), we note that

$$\nabla_{\mathbf{x}}\rho = \begin{bmatrix} \gamma_6 \\ \gamma_7 \\ \gamma_8 \end{bmatrix} + \lambda_2 \mathbf{g}_2$$

$$= \begin{bmatrix} \gamma_6 \\ \gamma_7 \\ \gamma_8 \end{bmatrix} + \left(\frac{1}{1+\beta z} - \gamma_2\right)\mathbf{g}_2 = \begin{bmatrix} a_1 + \frac{g_{21}}{1+\beta z} \\ a_2 + \frac{g_{22}}{1+\beta z} \\ a_3 + \frac{g_{23}}{1+\beta z} \end{bmatrix}, \quad (128)$$

where

$$\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} \gamma_6 - \gamma_2 g_{21} \\ \gamma_7 - \gamma_2 g_{22} \\ \gamma_8 - \gamma_2 g_{23} \end{bmatrix} = \begin{bmatrix} \gamma_6 - \gamma_2 E_u \\ \gamma_7 - \gamma_2 E_v \\ \gamma_8 + \beta\gamma_2(uE_u + vE_v) \end{bmatrix}. \quad (129)$$

From the projection equations in (1), we have

$$\mathbf{x} = (x, y, z)^\top = (u(1+\beta z), v(1+\beta z), z)^\top. \quad (130)$$

Then, using (128) and (130), we may write

$$\mathbf{x} \times \nabla_{\mathbf{x}}\rho = \begin{bmatrix} v(a_3 + g_{23}) + va_3\beta z - (a_2 + g_{22} + a_2\beta z)\frac{z}{1+\beta z} \\ -u(a_3 + g_{23}) - ua_3\beta z + (a_1 + g_{21} + a_1\beta z)\frac{z}{1+\beta z} \\ (ua_2 - va_1)(1+\beta z) + (ug_{22} - vg_{21}) \end{bmatrix}. \quad (131)$$

From the definition of $\boldsymbol{\pi}'$ in (116), we have

$$\pi'_3 = \pi_3 + (\mathbf{x} \times \nabla_{\mathbf{x}}\rho)_3 = (\mathbf{x} \times \nabla_{\mathbf{x}}\rho)_3, \quad (132)$$

since $\boldsymbol{\pi}^\top\mathbf{v} = \pi_3 = 0$ for an isotropic BRDF, from Proposition 1. From (123), we have

$$\pi'_3 = \gamma_5 + \lambda_1 g_{13} = \gamma_5, \quad (133)$$

since $g_{13} = 0$ by definition of $\mathbf{g}_1$. Substituting into the above from (132) and (131), we obtain

$$\pi'_3 = (\mathbf{x} \times \nabla_{\mathbf{x}}\rho)_3 = (ua_2 - va_1)(1+\beta z) + (ug_{22} - vg_{21}) = \gamma_5. \quad (134)$$

Thus, one may recover surface depth from differential stereo as:

$$z = \frac{1}{\beta}\left[\frac{\gamma_5 - (ug_{22} - vg_{21})}{ua_2 - va_1} - 1\right]$$
$$= \frac{1}{\beta}\left[\frac{\gamma_5 - uE_v + vE_u}{u(\gamma_7 - \gamma_2 E_v) - v(\gamma_6 - \gamma_2 E_u)} - 1\right]. \quad (135)$$

We may now state the analogue of Proposition 8, without the approximation made by neglecting $\nabla_{\mathbf{x}}\rho$:

Remark 3. *Under perspective projection, six differential motions of the camera suffice to yield depth of a surface with unknown isotropic BRDF and unknown light source.*

Additional Constraint on Surface Gradient One may also obtain additional constraints on the surface gradient for restricted BRDF types. From the definition of $\boldsymbol{\pi}'$ in (116), we have

$$\boldsymbol{\pi} = \mathbf{n} \times \nabla_{\mathbf{n}}\rho + \mathbf{s} \times \nabla_{\mathbf{s}}\rho = \boldsymbol{\pi}' - (\mathbf{x} \times \nabla_{\mathbf{x}}\rho). \quad (136)$$

Substituting from the solution for $\boldsymbol{\pi}'$ in (123) and the defintion of λ_1 in (126), we have

$$\boldsymbol{\pi} = \begin{bmatrix} \gamma_3 \\ \gamma_4 \\ \gamma_5 \end{bmatrix} + \lambda_1 \mathbf{g}_1 - (\mathbf{x} \times \nabla_{\mathbf{x}}\rho)$$
$$= \begin{bmatrix} \gamma_3 \\ \gamma_4 \\ \gamma_5 \end{bmatrix} + \left(\frac{z}{1+\beta z} - \gamma_1\right)\mathbf{g}_1 - (\mathbf{x} \times \nabla_{\mathbf{x}}\rho). \quad (137)$$

Then, using the relation in (131) and noting that $g_{13} = 0$, we obtain

$$\boldsymbol{\pi} = (f_1(z),\ f_2(z),\ f_3(z))^\top, \quad (138)$$

where

$$f_1(z) = \gamma_3 - g_{11}\gamma_1 - v(a_3 + g_{23}) - va_3\beta z + (a_2 + g_{22} + g_{11})\frac{z}{1+\beta z} + a_2\beta\frac{z^2}{1+\beta z}, \quad (139)$$

$$f_2(z) = \gamma_4 - g_{12}\gamma_1 + u(a_3 + g_{23}) + ua_3\beta z - (a_1 + g_{21} - g_{12})\frac{z}{1+\beta z} - a_1\beta\frac{z^2}{1+\beta z}, \quad (140)$$

$$f_3(z) = \gamma_5 - (ua_2 - va_1)(1+\beta z) - (ug_{22} - vg_{21}), \quad (141)$$

with $\mathbf{g}_1$, $\mathbf{g}_2$ and $\mathbf{a}$ determined by image derivatives, camera parameters and motion, as defined by (124), (125) and (129). Similar to the derivations in Section 4.2, for restricted BRDF types such as those dependent on the half-angle, we may now write

$$\frac{\pi_1}{\pi_2} = \frac{\mathbf{e}_1^\top \mathbf{n}}{\mathbf{e}_2^\top \mathbf{n}} = \frac{f_1(z)}{f_2(z)}, \quad (142)$$

where $f_1(z)$ and $f_2(z)$ are known functions of z given by (139) and (140), while $\mathbf{e}_1$ and $\mathbf{e}_2$ are entities such as $\mathbf{a}$ and $\mathbf{b}$ in (34) for the case of a half-angle BRDF, or $\mathbf{a}'$ and $\mathbf{b}'$ in (83) for the more general case of a BRDF dependent on an arbitrary angle in the $\{\mathbf{s}, \mathbf{v}\}$-plane. Using the relationship between the surface normal and the gradient given by (7), we note that the above may be writen in the form

$$h_1(z)z_x + h_2(z)z_y + h_3(z) = 0, \quad (143)$$

for known functions h_1, h_2 and h_3 that depend on $\mathbf{e}_1$, $\mathbf{e}_2$, f_1, f_2 and f_3. Thus, for restricted isotropic BRDFs, one may obtain a quasilinear relation in surface depth.

We note that the special form of $\boldsymbol{\pi}$ for isotropic BRDFs and the dependence of $\boldsymbol{\pi}'$ on $\nabla_{\mathbf{x}}\rho$ allow us to derive both (135) and (142) from the linear system in (122). We may now state the analogue of Proposition 9:

Remark 4. *Under perspective projection, six or more differential motions of the camera suffice to yield both depth and a quasilinear PDE constraint on the depth for a surface with BRDF dependent on known light and half-angle directions.*

It is clear that an analogue for Remark 2 may also be stated following a similar derivation as above.

References

[1] M. Chandraker, J. Bai, and R. Ramamoorthi. A theory of differential photometric stereo for unknown isotropic BRDFs. In *CVPR*, pages 2505–2512, 2011.

[2] M. Chandraker and R. Ramamoorthi. What an image reveals about material reflectance. In *ICCV*, pages 1076–1083, 2011.

[3] M. Chandraker, D. Reddy, Y. Wang, and R. Ramamoorthi. What object motion reveals about shape with unknown BRDF and lighting. In *CVPR*, pages 2523–2530, 2013.

[4] J. Clark. Active photometric stereo. In *CVPR*, pages 29–34, 1992.

[5] H. W. Haussecker and D. J. Fleet. Computing optical flow with physical models of brightness variation. *PAMI*, 23(6):661–673, 2001.

[6] B. Horn and B. Schunck. Determining optical flow. *Artificial Intelligence*, 17:185–203, 1981.

[7] F. John. *Partial Differential Equations*. App. Math. Sci. Springer, 1981.

[8] B. Lucas and T. Kanade. An iterative image registration technique with an application to stereo vision. In *Image Understanding Workshop*, pages 121–130, 1981.

[9] M. Minnaert. The reciprocity principle in lunar photometry. *Astrophysical Journal*, 3:403–410, 1941.

[10] H.-H. Nagel. On a constraint equation for the estimation of displacement rates in image sequences. *PAMI*, 11(1):13 –30, 1989.

[11] S. Negahdaripour. Revised definition of optical flow: Integration of radiometric and geometric cues for dynamic scene analysis. *PAMI*, 20(9):961–979, 1998.

[12] I. Sato, T. Okabe, Q. Yu, and Y. Sato. Shape reconstruction based on similarity in radiance changes under varying illumination. In *ICCV*, pages 1–8, 2007.

[13] S. Savarese, M. Chen, and P. Perona. Local shape from mirror reflections. *IJCV*, 64(1):31–67, 2005.

[14] S. Seitz, B. Curless, J. Diebel, D. Scharstein, and R. Szeliski. A comparison and evaluation of multiview stereo reconstruction algorithms. In *CVPR*, pages 519–526, 2006.

[15] A. Treuille, A. Hertzmann, and S. Seitz. Example-based stereo with general BRDFs. In *ECCV*, pages 457–469, 2004.

[16] A. Verri and T. Poggio. Motion field and optical flow: Qualitative properties. *PAMI*, 11(5):490–498, 1989.

[17] P. Woodham. Photometric method for determining surface orientation from multiple images. *Opt. Engg.*, 19(1):139–144, 1980.

[18] T. Zickler, P. Belhumeur, and D. Kriegman. Helmholtz stereopsis: Exploiting reciprocity for surface reconstruction. *IJCV*, 49(2/3):1215–1227, 2002.