On Shape and Material Recovery from Motion

Manmohan Chandraker

NEC Labs America, Cupertino, USA

Abstract. We present a framework for the joint recovery of the shape and reflectance of an object with dichromatic BRDF, using motion cues. We show that four (small or differential) motions of the object, or three motions of the camera, suffice to yield a linear system that decouples shape and BRDF. The theoretical benefit is that precise limits on shape and reflectance recovery using motion cues may be derived. We show that shape may be recovered for unknown isotropic BRDF and light source. Simultaneous reflectance estimation is shown ambiguous for general isotropic BRDFs, but possible for restricted BRDFs representing commong materials like metals, plastics and paints. The practical benefit of the decoupling is that joint shape and BRDF recovery need not rely on alternating methods, or restrictive priors. Further, our theory yields conditions for the joint estimability of shape, albedo, BRDF and directional lighting using motion cues. Surprisingly, such problems are shown to be well-posed even for some non-Lambertian material types. Experiments on measured BRDFs from the MERL database validate our theory.

1 Introduction

Shape and lighting interact in complex ways through the bidirectional reflectance distribution function (BRDF) to produce the variety of images around us. Shape recovery with unknown BRDF and lighting is traditionally considered hard, while their joint recovery is deemed severely ill-posed. This paper presents a framework for understanding how cues from object or camera motion govern shape, BRDF and lighting recovery. Our theory leads to several surprising results – for instance, we show that a few (three or four) motions allow shape recovery with unknown isotropic BRDF and lighting, allow simultaneous shape and BRDF recovery for common materials like metals or plastics, or lead to a well-posed problem for joint recovery of shape, reflectance and directional lighting for such materials.

The appearance of many real-world materials is governed by a dichromatic model, which consists of a diffuse albedo and a non-diffuse reflectance that is a function of surface orientation, lighting and viewpoint. In Section 4, we show that change in image intensties for isotropic dichromatic materials, for both the cases of object and camera motion, may be linearly related to entities associated with shape, reflectance and lighting. We call these differential flow and stereo relations, respectively, following prior works for monochromatic materials [8,5].

A direct consequence of this linearity is that shape and reflectance terms are neatly decoupled by motion cues over an image sequence. In Sec. 5 and 6, we

D. Fleet et al. (Eds.): ECCV 2014, Part VII, LNCS 8695, pp. 202–217, 2014.
© Springer International Publishing Switzerland 2014

Table 1. Summary of the theoretical results of this paper. We establish conditions when shape, shape + BRDF and shape + BRDF + lighting estimation problems are well-posed for dichromatic BRDFs, using motion cues resulting from either object or camera motion. Green indicates well-posed estimation, pink indicates ill-posed, while gray indicates ill-posed but solvable under mild regularization. More constrained BRDF or restrictive input conditions yield better-posed estimations, so the top to bottom variation is largely green to gray to pink. For reference, BRDF estimation results from [7], with known shape and single image, are also shown ("?" denotes conjectured, but not proved). Motion cues can recover unknown shape, as well as estimate BRDF and determine BRDF estimability under a wider set of conditions.

BRDF	Knowns		Object Motion			Camera Motion			[7]
Type	Albedo	Light	Shape	+ BRDF	+ Light	Shape	+ BRDF	+ Light	BRDF
1-lobe	✓	✓	Prop. 3	Prop. 5	–	Prop. 4	Prop. 7	–	
1-lobe	✓	✗	Prop. 3	Sec. 7(c)		Prop. 4	Prop. 9		
1-lobe	✗	✓	Prop. 3	Prop. 5	–	Prop. 4	Prop. 7	–	?
1-lobe	✗	✗	Prop. 3	Sec. 7(a)		Prop. 4	Prop. 9		
2-lobe	✓	✓	Prop. 3	Cor. 2	–	Prop. 4	Prop. 7	–	
2-lobe	✓	✗	Prop. 3	Sec. 7(d)		Prop. 4	Sec. 7		
2-lobe	✗	✓	Prop. 3	Prop. 6	–	Prop. 4	Prop. 7	–	?
2-lobe	✗	✗	Prop. 3	Prop. 8		Prop. 4	Sec. 7		
K-lobe	✓	✓	Prop. 3	Cor. 3	–	Prop. 4	Sec. 6.2	–	?
K-lobe	✓	✗	Prop. 3	Sec. 7(e)		Prop. 4	Sec. 7		?
K-lobe	✗	✓	Prop. 3	Prop. 6	–	Prop. 4	Sec. 6.2	–	?
K-lobe	✗	✗	Prop. 3	Prop. 8		Prop. 4	Sec. 7		?

show that four differential object motions, or three camera motions, suffice to recover surface depth and in many cases, the unknown BRDF as well. This is surprising, since the BRDF can encode complex interactions between shape and lighting. The immediate practical benefit is that we may recover both shape and reflectance without resort to unstable alternating methods, iterative optimization, or restrictive priors on geometry and reflectance.

A theoretical benefit is that our analysis relates the precise extent of shape and BRDF recovery to the hardness of estimation conditions. In Sec. 6 and 7, we relate the well-posedness of shape and reflectance recovery to BRDF complexity, as well as to input conditions such as knowledge of lighting or uniform albedo. In the general isotropic case, we show that BRDF may not be estimated using motion cues alone, which justifies several works that impose priors for reflectance recovery. However, when the BRDF depends on one or more angles about the normal – for example, half-angle BRDFs for many metals, plastics or paints – we show that both shape and BRDF may be unambiguously recovered.

Finally, we analyze the well-posedness of joint recovery problems where shape, albedo, reflectance functions, lighting and reflection directions are all unknown. We show in Sec. 7 that the problem is well-posed even for some non-Lambertian materials (for example, half-angle BRDFs) under camera motion and only mildly ill-posed under object motion. This is contrary to conventional belief that such problems are severely ill-posed. Our theoretical results are summarized in Tab. 1.

2 Related Work

Motion cues for shape recovery have been extensively studied within the purviews of optical flow [11,13] and multiview stereo [21]. It is well-known from early works that a Lambertian reflectance has limitations [15,23]. Several approaches have been proposed for shape recovery with general BRDFs, such as Helmholtz reciprocity for stereo by Zickler et al. [25], intensity profiles for photometric stereo by Sato et al. [20] and specular flow for mirror surfaces by Canas et al. [4].

Our shape recovery results are closely related to prior works of Chandraker et al. for light source [6], object [8] and camera motions [5], which assume a monochromatic BRDF. The theory of this paper generalizes to dichromatic BRDFs and goes further to analyze the problem of BRDF estimation too.

For BRDF estimation, parametric models have a long history [3,22] and we refer the reader to [16] for an empirical comparison. Non-parametric [19,18] and data-driven [14] approaches are popular for their representation power, but require a large amount of data or rely on complex estimation whose properties are hard to characterize. Semiparametric models have also been proposed for BRDF editing [12] and estimation [7]. Our work extends such methods to unknown shape and characterizes how motion cues provide additional information.

Joint recovery of two or more elements among shape, BRDF and illumination have also attracted significant interest. Shape and illumination have been estimated under the Lambertian assumption by imposing priors [2]. Goldman et al. use a set of basis materials in photometric stereo to recover shape and reflectance [10]. Alldrin et al. alternatingly optimize over shape and material to recover both under light source motion [1], as do Zhang et al. for shape, motion and lighting for Lambertian reflectance [24]. An alternating method to estimate shape and isotropic BRDF under natural illumination is proposed in [17]. This paper shows that motion cues decouple shape and reflectance, so they may be estimated simultaneously, rather than in an alternating fashion.

Our focus is on establishing limits to shape and reflectance recovery using motion cues, regardless of estimation method. Some works like [7] derive conditions on well-posedness of BRDF estimation using a single image, with known shape. As discussed in Sec. 7, our theory not only supports the conclusions of [7], but also generalizes it both to unknown shape and to show how motion cues may sometimes enable BRDF estimation that is ill-posed for single images.

3 Preliminaries

Assumptions and Setup. We assume that the lighting is directional and distant, while the BRDF is isotropic and homogeneous (or having slow spatial variation). Global illumination effects like interreflections and shadows are assumed negligible. The origin of 3D coordinates is defined as the principal point on the image plane. So, the camera center is $(0, 0, -f)^\top$, where f is the focal length. The image of a 3D point $\mathbf{x} = (x, y, z)^\top$ is given by a point $\mathbf{u} = (u, v)^\top$ on the image plane, with

$$(1 + \beta z)u \;=\; x, \qquad (1 + \beta z)v \;=\; y, \quad \text{where } \beta = f^{-1}. \tag{1}$$

Motion Field. In the case of object motion, we assume the object undergoes rotation $\mathbf{R}$ and translation $\boldsymbol{\tau}$ relative to camera. For a camera motion $\{\mathbf{R}^\top, -\mathbf{R}^\top\boldsymbol{\tau}\}$, the object and lighting are equivalently assumed to undergo a relative motion of $\{\mathbf{R}, \boldsymbol{\tau}\}$. In either case, for differential motion, we approximate $\mathbf{R} \approx \mathbf{I} + [\boldsymbol{\omega}]_\times$, where $\boldsymbol{\omega} = (\omega_1, \omega_2, \omega_3)^\top$ and $[\cdot]_\times$ denotes the cross-product operator.

The motion field $\boldsymbol{\mu}$ is the image velocity, that is, $\boldsymbol{\mu} = (\dot{u}, \dot{v})^\top$. Substituting from (1), with α_i known functions having forms shown in [9], we obtain

$$\boldsymbol{\mu} = (1+\beta z)^{-1}\left[\alpha_1(1+\beta z) + (\alpha_2 + \omega_2 z),\ \alpha_3(1+\beta z) + (\alpha_4 - \omega_1 z)\right]^\top. \quad (2)$$

Image Formation. For surface normal $\mathbf{n}$, light source $\mathbf{s}$ and viewing direction $\mathbf{v}$, the dichromatic imaging model at a surface point $\mathbf{x}$ is

$$I(\mathbf{u}, t) = \sigma(\mathbf{x})\mathbf{n}^\top\mathbf{s} + \rho(\mathbf{x}, \mathbf{n}, \mathbf{s}, \mathbf{v}), \quad (3)$$

where σ is the diffuse albedo and ρ is the BRDF. Such models closely approximate real-world materials [16]. Parametric models like Torrance-Sparrow are often used to model ρ, but this work considers the form of ρ unknown.

4 Differential Relations for Dichromatic BRDFs

We now derive differential relations between shape and reflectance, induced by motion. We present intuitions here and refer the reader to [9] for details.

4.1 Object Motion

Consider the setup where the camera and lighting are fixed, while the object moves relative to the camera. Since the light position $\mathbf{s}$ does not change with time, we may write the BRDF of a point as a function of its position and normal, that is, $\rho(\mathbf{x}, \mathbf{n})$. Taking the total derivative on both sides of (3), we get

$$I_u\dot{u} + I_v\dot{v} + I_t = \sigma\frac{d}{dt}(\mathbf{n}^\top\mathbf{s}) + (\mathbf{n}^\top\mathbf{s})\frac{d\sigma}{dt} + \frac{d}{dt}\rho(\mathbf{x}, \mathbf{n}). \quad (4)$$

Since albedo is intrinsically defined on surface coordinates, its total derivative in 3D coordinates vanishes. For rigid body motion, change in normal is given by $\dot{\mathbf{n}} = \boldsymbol{\omega} \times \mathbf{n}$, while change in position is the linear velocity, $\dot{\mathbf{x}} = \boldsymbol{\nu}$. Using chain rule differentiation and recognizing $\boldsymbol{\mu} = (\dot{u}, \dot{v})^\top$ as the motion field, we have

$$(\nabla I)^\top\boldsymbol{\mu} + I_t = (\sigma\mathbf{s} + \nabla_{\mathbf{n}}\rho)^\top(\boldsymbol{\omega} \times \mathbf{n}) + (\nabla_{\mathbf{x}}\rho)^\top\boldsymbol{\nu}. \quad (5)$$

In our setup, the BRDF is homogeneous and lighting is distant, thus, $\nabla_{\mathbf{x}}\rho$ is negligible. Thus, we obtain the following relation:

$$\boxed{(\nabla_{\mathbf{u}} I)^\top\boldsymbol{\mu} + I_t = \left[\mathbf{n} \times (\sigma\mathbf{s} + \nabla_{\mathbf{n}}\rho)\right]^\top\boldsymbol{\omega}.} \quad (6)$$

Similar to [8], we call this the *differential flow relation*. However, note that the above is a relation for dichromatic BRDFs given by (3), while [8] assumes a monochromatic model. For now, we make an observation which will be used later:

Proposition 1. *For an object with dichromatic BRDF undergoing differential motion, a differential flow relation exists that is linear in entities that depend on shape (motion field and surface normals), reflectance and lighting.*

4.2 Camera Motion

Next, a similar analysis for the case of camera motion shows the existence of a *differential stereo relation* (see [9] for a derivation from first principles):

$$\boxed{(\nabla_{\mathbf{u}} I)^\top \boldsymbol{\mu} + I_t = (\mathbf{n} \times \nabla_{\mathbf{n}} \rho + \mathbf{s} \times \nabla_{\mathbf{s}} \rho)^\top \boldsymbol{\omega}.} \tag{7}$$

We again observe a similarity to the monochromatic case of [5], while noting:

Proposition 2. *For an object with dichromatic BRDF observed by a camera undergoing differential motion, a differential stereo relation exists that is linear in entities that depend on shape, reflectance and lighting.*

The above linearities ensconced within the differential flow and stereo relations play a key role in understanding the limits of both shape and reflectance recovery using motion cues. The following sections are devoted to exploring those limits.

5 Shape Recovery

In this section, we establish shape recovery from motion cues, with unknown dichromatic BRDF. Further, we may assume the lighting to also be unknown.

5.1 Object Motion

Substituting the motion field (2) into the differential flow relation (6), we obtain

$$(p + \beta q)z + (q + r) \;=\; (1 + \beta z)\boldsymbol{\omega}^\top \boldsymbol{\pi}, \tag{8}$$

where $p = I_u\omega_2 - I_v\omega_1$, $q = \alpha_1 I_u + \alpha_3 I_v + I_t$ and $r = \alpha_2 I_u + \alpha_4 I_v$ are known and

$$\boldsymbol{\pi} = \mathbf{n} \times (\sigma \mathbf{s} + \nabla_{\mathbf{n}} \rho). \tag{9}$$

We are now in a position to state the following:

Proposition 3. *Four or more differential motions of a surface with unknown dichromatic BRDF, under unknown light direction, suffice to yield surface depth.*

Proof. For $m \geq 4$, let known motions $\{\boldsymbol{\omega}^i, \boldsymbol{\tau}^i\}$, where $\boldsymbol{\omega}^i$ span $\mathbb{R}^3$, relate images $I_1, \cdots, I_m$ to I_0. From (8), we have a sequence of differential flow relations

$$(p^i + \beta q^i)z - (1 + \beta z)\boldsymbol{\pi}^\top \boldsymbol{\omega}^i + (q^i + r^i) = 0, \quad \text{for } i = 1, \cdots, m. \tag{10}$$

Let $\mathbf{c}^i = [p^i + \beta q^i, -\omega_1^i, -\omega_2^i, -\omega_3^i]^\top$ be rows of the $m \times 4$ matrix $\mathbf{C} = \left[\mathbf{c}^1, \cdots, \mathbf{c}^m\right]^\top$. Let $\mathbf{q} = [q^1, \cdots, q^m]^\top$ and $\mathbf{r} = [r^1, \cdots, r^m]^\top$. Define $\boldsymbol{\epsilon} = -\mathbf{C}^+(\mathbf{q} + \mathbf{r})$, where $\mathbf{C}^+$ is the Moore-Penrose pseudoinverse of $\mathbf{C}$ and let $\boldsymbol{\epsilon}' = (\epsilon_2, \epsilon_3, \epsilon_4)^\top$. Then, we have

$$z = \epsilon_1 \tag{11}$$

$$(1 + \beta z)\boldsymbol{\pi} = \boldsymbol{\epsilon}'. \tag{12}$$

Thus, from (11), we have obtained the surface depth. □

5.2 Camera Motion

We again start by observing that substituting the motion field (2) in the differential stereo relation (7) leads to an equation of the form (8). However, note that the definition of $\boldsymbol{\pi}$ is different for the case of camera motion. Indeed, an isotropic BRDF may be written as $\rho(\mathbf{n}, \mathbf{s}, \mathbf{v}) = \bar{\rho}(\mathbf{n}^\top \mathbf{s}, \mathbf{s}^\top \mathbf{v}, \mathbf{n}^\top \mathbf{v})$, whereby

$$\boldsymbol{\pi} \;=\; \mathbf{n} \times \nabla_{\mathbf{n}} \rho + \mathbf{s} \times \nabla_{\mathbf{s}} \rho \;=\; \bar{\rho}_{\mathbf{n}^\top \mathbf{v}}(\mathbf{n} \times \mathbf{v}) + \bar{\rho}_{\mathbf{s}^\top \mathbf{v}}(\mathbf{s} \times \mathbf{v}), \tag{13}$$

thus, $\boldsymbol{\pi}^\top \mathbf{v} = \pi_3 = 0$, as [5]. Using $m \geq 3$ differential motions $\{\boldsymbol{\omega}^i, \boldsymbol{\tau}^i\}$, one may define the $m \times 3$ matrix $\tilde{\mathbf{C}} = \left[\tilde{\mathbf{c}}^1, \cdots, \tilde{\mathbf{c}}^m\right]^\top$ with rows $\tilde{\mathbf{c}}^i = [-(p'^i + \beta q'^i),\, \omega_1^i,\, \omega_2^i]^\top$. Then, the system of m differential stereo relations (10) may be solved to obtain

$$[z,\, (1 + \beta z)\pi_1,\, (1 + \beta z)\pi_2]^\top \;=\; \tilde{\boldsymbol{\epsilon}}, \tag{14}$$

where $\tilde{\boldsymbol{\epsilon}} = (\tilde{\epsilon}_1, \tilde{\epsilon}_2, \tilde{\epsilon}_3)^\top = \tilde{\mathbf{C}}^+(\mathbf{q} + \mathbf{r})$, with $\mathbf{q}$ and $\mathbf{r}$ as defined previously. It follows that $z = \tilde{\epsilon}_1$ yields the surface depth. Thus, we have shown:

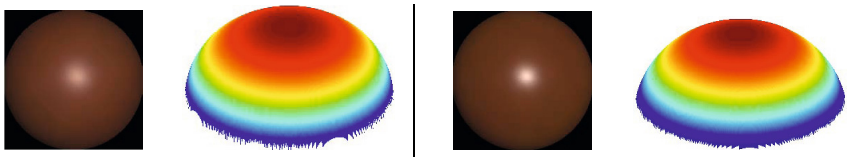
Proposition 4. *Three or more differential motions of the camera suffice to yield depth of a surface with unknown dichromatic BRDF and unknown light direction.*

We observe that even with the assumption of a dichromatic BRDF, the shape recovery results of Prop. 3 and 4 are similar to the monochromatic cases of [8,5]. Indeed, although the images are not logarithmic here and the definitions of $\boldsymbol{\pi}$ are different from [8,5], the overall forms of the differential flow and stereo relations exhibit similar linearities. Intuitively, this leads to similar shape recovery results.

But more importantly, we note an additional benefit of the linear relationship between shape and BRDF in the differential flow and stereo relations. Namely, in (12) and (14), we also obtain information about the BRDF in the form of $\boldsymbol{\pi}$. Our focus for the remainder of the paper will be on how the differential flow and stereo relations aid understanding of reflectance recovery.

5.3 Experimental Validation

We use real measured BRDFs from the MERL database [14] to illustrate shape recovery in Fig. 1. For the `colonial-maple-223` material, images such as Fig. 1(a) are observed under five differential motions of the object. Image derivatives are computed using a smoothing filter to account for noise in the measured BRDF data. The shape recovered using Prop. 3 is shown in Fig. 1(b). Similarly, for the `natural-209` material, images are observed under five differential motions of the camera (Fig. 1(c)) and the shape recovered using Prop. 4 is shown in Fig. 1(d).



(a) Input [1 of 6] (Object motion) (b) Recovered shape (c) Input [1 of 6] (Camera motion) (d) Recovered shape

Fig. 1. (a) One of six images using five object motions, with `colonial-maple-223` material (unknown BRDF) and unknown lighting. (b) Shape recovered using Proposition 3. (c) One of six images using five camera motions, with `natural-209` material (unknown BRDF) and unknown lighting. (d) Shape recovered using Proposition 4.

6 Shape and Reflectance Recovery

We now consider the problem of simultaneous shape and reflectance recovery. For both the cases of object and camera motion, in addition to the shape, we have obtained information about the reflectance in (12) and (14):

$$\text{Object: } \boldsymbol{\pi} = \frac{1}{1+\beta\epsilon_1}(\epsilon_2, \epsilon_3, \epsilon_4)^\top, \qquad \text{Camera: } \boldsymbol{\pi} = \frac{1}{1+\beta\tilde{\epsilon}_1}(\tilde{\epsilon}_2, \tilde{\epsilon}_3, 0)^\top. \quad (15)$$

It is interesting that shape and reflectance may be decoupled using motion cues, despite the complex interactions enabled by an unknown dichromatic BRDF. We now show how the linearity of differential flow and stereo allows us to impose limits on the extent to which BRDF may be recovered using motion cues. In this section, we will assume a known light source direction.

6.1 Object Motion

Using $m \geq 4$ motions of an object, we may always obtain the shape using Proposition 3. We will now explore the extent to which BRDF may be recovered.

General Isotropic BRDF. For an isotropic BRDF, image formation depends on the three angles between surface normal, camera and lighting directions:

$$I = \sigma \mathbf{n}^\top \mathbf{s} + \rho(\theta, \phi, \psi), \text{ where } \theta = \mathbf{n}^\top \mathbf{s},\ \phi = \mathbf{s}^\top \mathbf{v} \text{ and } \psi = \mathbf{n}^\top \mathbf{v}. \quad (16)$$

Using (9) to define $\boldsymbol{\pi}$ and substituting in (12), we have the following relation:

$$(1+\beta z)\mathbf{n} \times [(\sigma + \rho_\theta)\mathbf{s} + \rho_\psi \mathbf{v}] = \boldsymbol{\epsilon}', \quad (17)$$

where $\rho_\phi = 0$ since ϕ remains unchanged for object motion. Further, the albedo and BRDF-derivative along the θ direction, ρ_θ, cannot be disambiguated. This can also be intuitively understood since ρ is an arbitrary function and may ambiguously incorporate any information about θ that is included in the diffuse term. Thus, only BRDF variation along ψ is captured by object motion.

Even though estimation of a dichromatic BRDF from object motion is ambiguous in the fully general case, we show that it is unique for more restricted BRDFs exhibited by several real-world materials.

Single-Lobe Dichromatic Reflectance. For many materials, the reflectance depends predominantly on the angle between the surface normals and a single reflection direction, $\mathbf{r}$. Most commonly, such as with metals, plastics and many paints, the reflection direction is aligned with the half-angle between the source and viewing directions. This observation has also been used to propose parametric models like Blinn-Phong [3] and (simplified) Torrance-Sparrow [22]. For many materials in the MERL dataset, empirical studies have found a single lobe BRDF to be sufficiently descriptive [16,7]. For such materials, we show:

Proposition 5. *Four or more differential motions of an object with single-lobe dichromatic BRDF suffice to uniquely determine its shape, albedo and reflectance.*

Proof. The image formation for an object with single-lobe BRDF is given by $I = \sigma \mathbf{n}^\top \mathbf{s} + \rho(\eta)$, where $\eta = \mathbf{n}^\top \mathbf{r}$. Substituting in (9), we obtain

$$\boldsymbol{\pi} = \mathbf{n} \times (\sigma \mathbf{s} + \nabla_{\mathbf{n}} \rho) = \mathbf{n} \times (\sigma \mathbf{s} + \rho_\eta \mathbf{r}). \tag{18}$$

Given images under four or more differential motions, Proposition 3 and (15) guarantee the existence of a relation between depth and reflectance:

$$(1 + \beta\epsilon_1)\,[\mathbf{n}(\epsilon_1) \times (\sigma \mathbf{s} + \rho_\eta \mathbf{r})] = \boldsymbol{\epsilon}', \tag{19}$$

where the normals $\mathbf{n}(\epsilon_1)$ are obtained from the derivatives of surface depth estimated in (11). Thus, the above is a rank 2 system of three linear equations in the two unknowns σ and ρ_η, which may both be recovered. Finally, we note that for most materials, reflection vanishes around grazing angles (indeed, the non-diffuse component of half-angle BRDFs is often super-linear). Thus, $\rho(0) = 0$, whereby ρ_η may be integrated to recover the BRDF function ρ. □

Thus, we have shown that for a large class of dichromatic materials, motion cues alone can determine all of shape, albedo and BRDF. Intuitively, the linear separability of shape and reflectance established by Proposition 1 allows us to determine conditions when BRDF is recoverable. Further, it also allows us to determine when BRDF estimation is ambiguous, as discussed next.

Degeneracy. The result of Proposition 5 relies on the direction $\mathbf{r}$ being distinct from the light source $\mathbf{s}$, otherwise (19) reduces to: $(1 + \beta\epsilon_1)\,[\mathbf{n}(\epsilon_1) \times (\sigma + \rho_\eta)\mathbf{s}] = \boldsymbol{\epsilon}'$. Clearly, in this case, one may not independently recover both albedo σ and the BRDF-derivative ρ_η. For most materials, it is indeed the case that $\mathbf{r} \neq \mathbf{s}$ (for instance, $\mathbf{r}$ is often the half-angle). However, there are two important exceptions. First, an object with arbitrary isotropic BRDF observed under colocated illumination follows an image formation model given by $I = \sigma \mathbf{n}^\top \mathbf{s} + \bar{\rho}(\mathbf{n}^\top \mathbf{s})$ (since $\mathbf{s} = \mathbf{v}$ and $\|\mathbf{s}\| = 1$, there exists a function $\bar{\rho}$ such that $\rho(\mathbf{n}^\top \mathbf{s}, \mathbf{s}^\top \mathbf{v}, \mathbf{n}^\top \mathbf{v}) = \bar{\rho}(\mathbf{n}^\top \mathbf{s})$). Second, retroreflective materials such as those used to enhance visibility of road signs reflect light back towards the source direction. Thus, we may state:

Corollary 1. *Albedo and reflectance cannot be disambiguated using motion cues for an object with retroreflective BRDF or one observed under colocated lighting.*

Multi-lobe Reflectance. For some materials, the image may be explained by reflection along two or more angles with respect to the surface normal. That is,

$$I = \sigma \mathbf{n}^\top \mathbf{s} + \rho(\eta_1, \cdots, \eta_K), \text{ where } \eta_i = \mathbf{n}^\top \mathbf{r}_i, \text{ for } i = 1, \cdots, K, \tag{20}$$

where $K \geq 2$. Empirical studies like [16,7] show that accounting for BRDF dependence on a second direction besides the half-angle leads to a better approximation for materials like veneer paints and fabrics. We will refer to directions η_i as lobes.

Unknown Albedo. Given four or more differential motions, shape may be recovered for such BRDFs using Proposition 3. Substituting from (20) into the expression for $\boldsymbol{\pi}$ in (9) and using (15), we obtain a relation between depth and reflectance:

$$(1 + \beta\epsilon_1)\, \mathbf{n}(\epsilon_1) \times (\sigma \mathbf{s} + \sum_{i=1}^{K} \rho_{\eta_i} \mathbf{r}_i) = \boldsymbol{\epsilon}', \tag{21}$$

which is a system of three linear equations in $K+1$ unknowns $\{\sigma, \rho_{\eta_1}, \cdots, \rho_{\eta_K}\}$. For $K > 2$, clearly the system (21) is underdetermined and no unique solution is possible. For $K = 2$, the above is a system of three linear equations in three unknowns σ, ρ_{η_1} and ρ_{η_2}. However, note that the 3×3 matrix associated with the system in (21), $\mathbf{A} = (\mathbf{n} \times \mathbf{s}, \mathbf{n} \times \mathbf{r}_1, \mathbf{n} \times \mathbf{r}_2)$, is rank-deficient. Thus, we state:

Proposition 6. *A K-lobe BRDF may not be recovered using object motion alone for an object with unknown albedo when $K \geq 2$ (although shape may be recovered).*

It is interesting that the above ambiguity also affects important classes of parametric BRDFs. An example is the Torrance-Sparrow model ignoring geometric attenuation and Fresnel terms, for which image formation may be expressed as

$$I = \sigma \mathbf{n}^\top \mathbf{s} + \rho(\mathbf{n}^\top \mathbf{h}, \mathbf{n}^\top \mathbf{v}), \text{ with } \rho \sim (\mathbf{n}^\top \mathbf{v})^{-1} \exp\left(-\lambda^2 (\cos^{-1} \mathbf{n}^\top \mathbf{h})^2\right), \tag{22}$$

where λ is a surface roughness parameter.

Known Albedo. We now consider the important case of known albedo. Note that uniform albedo, which is a common assumption in BRDF acquisition and estimation settings like [14,16], reduces to known albedo when the non-diffuse components of a dichromatic BRDF are super-linear and rapidly diminish away from the lobe directions, as is true for most materials. Since the matrix $\mathbf{A}$ defined above is rank 2, the remaining unknowns ρ_{η_1} and ρ_{η_2} may still be recovered when the albedo is known. Thus, we have:

Corollary 2. *With known albedo, both shape and a BRDF with up to two lobes may be recovered using four or more differential motions of the object.*

Finally, we note that with $K \geq 3$ lobes, even with known albedo, the above rank 2 system of equations is underdetermined, so we state:

Corollary 3. *Object motion cannot disambiguate the estimation of a BRDF with $K \geq 3$ lobes, even with known albedo (although shape may still be recovered).*

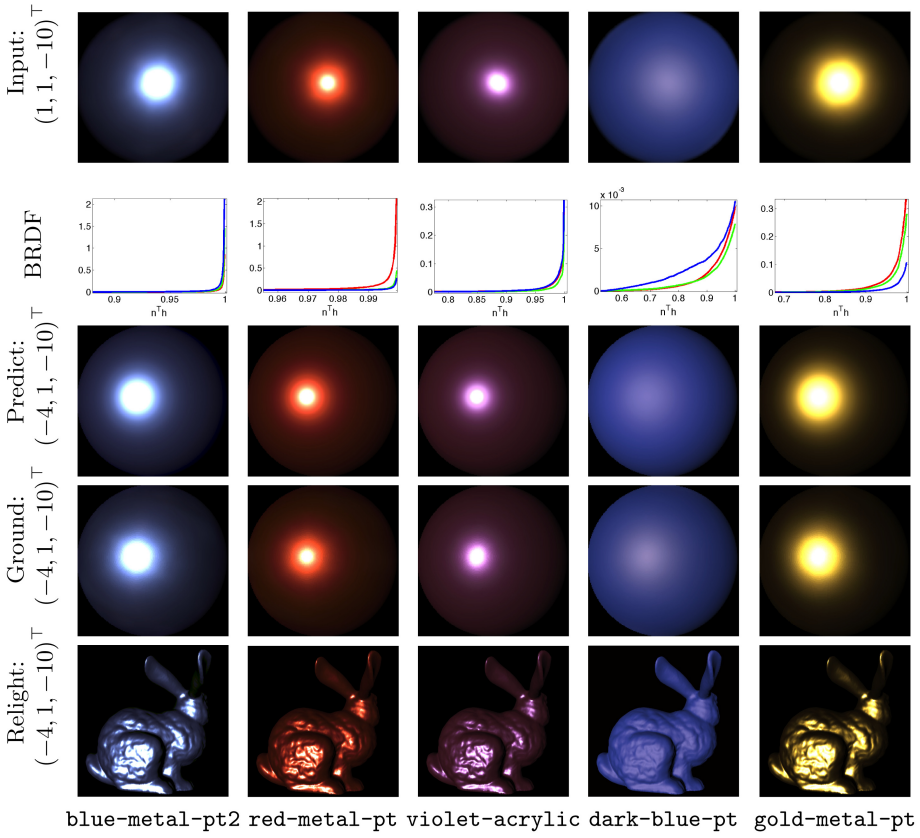


Fig. 2. (Row 1) One of six input images from five small object motions for MERL database BRDFs. (Row 2) BRDF recovered in each color channel. (Row 3) Predicted appearance for a novel light direction. (Row 4) Ground truth appearance for the novel lighting. (Row 5) A novel geometry relighted using the estimated BRDF. Percentage image errors of row 3 relative to row 4 are 5.8%, 2.1%, 1.8%, 4.7% and 7.8%, respectively.

We also note that several interesting relationships exist between the above results and [7], where uniqueness conditions for BRDF estimation are established in a single image setting, with known shape and uniform albedo. The results of our theory further support the conclusions of [7] and extend them to a multiple image setting. A discussion of those relationships is presented in Section 7.

Experimental Validation. We validate our theory using real measured BRDFs from the MERL database [14]. In the top row of Figure 3, we show one of six input images, corresponding to five differential object motions. Spatial and temporal image derivatives are computed, following which depth and $\boldsymbol{\pi}$ are determined by Prop. 3. From $\boldsymbol{\pi}$, the BRDF is estimated using the theory of this section.

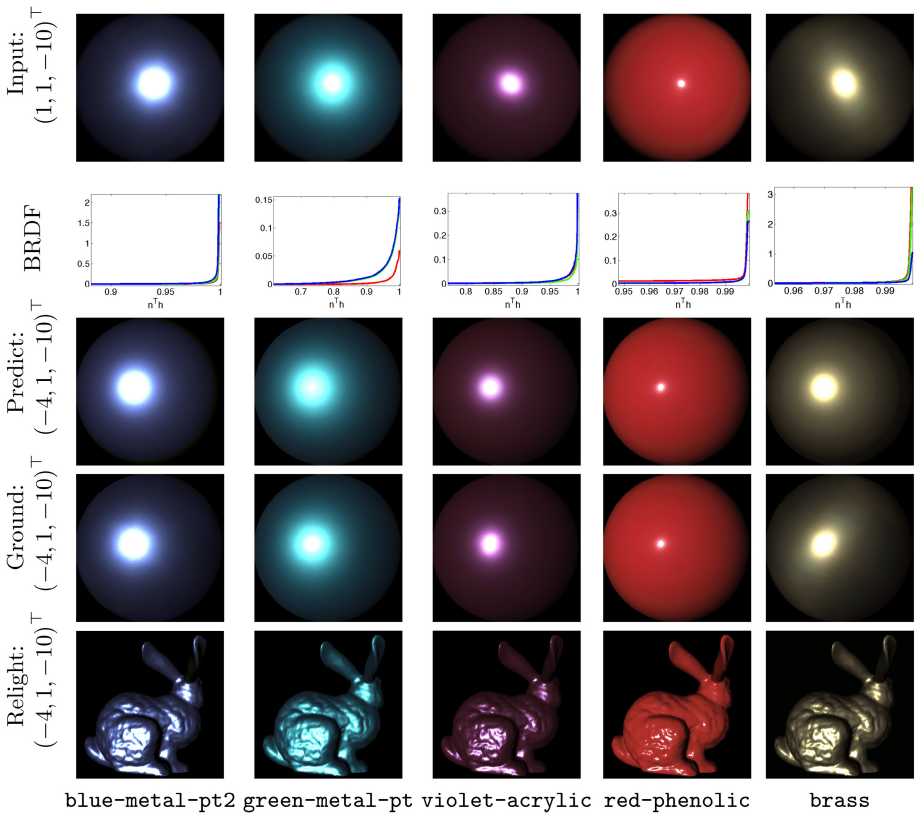


Fig. 3. (Row 1) One of six input images from five small camera motions for MERL database BRDFs. (Row 2) BRDF recovered in each color channel. (Row 3) Predicted appearance for a novel light direction. (Row 4) Ground truth appearance for the novel lighting. (Row 5) A novel geometry relighted using the estimated BRDF. Note the reasonable approximation obtained even for the anisotropic `brass` material. Percentage image errors of row 3 relative to row 4 are 3.4%, 2.9%, 1.6%, 4.8% and 15.8%, respectively.

The second row shows the estimated BRDF curves in each color channel. Notice the qualitative accuracy of the estimation, as more specular materials have curves with sharper rise. With the recovered shape and BRDF, appearance is predicted from a novel lighting direction, shown in the third row. It is found to closely match the ground truth, as shown in the fourth row. The final row shows a novel geometry relighted using the estimated BRDF, from a novel lighting direction. Further experiments are included in [9].

6.2 Camera Motion

We now briefly study the case of camera motion, while refering the reader to [9] for details. We have seen in (15) that $m \geq 3$ motions determine the entity $\boldsymbol{\pi}$ that encodes BRDF-derivatives. We specify what BRDF information may be recovered from $\boldsymbol{\pi}$, given its form in (7):

$$\boldsymbol{\pi} = \mathbf{n} \times \nabla_{\mathbf{n}} \rho + \mathbf{s} \times \nabla_{\mathbf{s}} \rho. \tag{23}$$

Recall from (13) that for any isotropic BRDF where $\rho(\mathbf{n}, \mathbf{s}, \mathbf{v}) = \bar{\rho}(\mathbf{n}^\top \mathbf{s}, \mathbf{s}^\top \mathbf{v}, \mathbf{n}^\top \mathbf{v})$, the BRDF-derivative $\bar{\rho}_{\mathbf{n}^\top \mathbf{s}}$ vanishes. Thus, a full isotropic BRDF may not be recovered using camera motion. However, one may still recover restricted forms of isotropic BRDFs, such as the K-lobe model, as shown next.

It also follows from (13) that $\boldsymbol{\pi}^\top \mathbf{v} = \pi_3 = 0$. Thus, only two independent constraints on the BRDF are available through differential motion of the camera. Consider a K-lobe image formation $I = \sigma \mathbf{n}^\top \mathbf{s} + \rho(\eta_1, \cdots, \eta_K)$, where $\eta_i = \mathbf{n}^\top \mathbf{r}_i$. From the linearity of differentiation, π_j are of the form $\sum_{i=1}^{K} \rho_{\eta_i} f_i^j(\mathbf{n}, \mathbf{s}, \mathbf{r}_i)$, for some analytic functions f_i^j and $j = 1, 2$. Clearly, for $K > 2$, one may not determine all the ρ_{η_i}, since only two constraints on $\boldsymbol{\pi}$ are available. Further, note that there is no dependence of $\boldsymbol{\pi}$ on σ, unlike the case of object motion. Thus, for $K = 2$, when $\mathbf{r}_1$ and $\mathbf{r}_2$ are independent and "general" (that is, with no special dependencies for f_i), both ρ_{η_1} and ρ_{η_2} may be determined. Thus, the BRDF ρ can be recovered by integration. For known lighting, the albedo may subsequently be estimated by subtracting the non-diffuse component. Thus, we have:

Proposition 7. *Three or more differential motions of the camera suffice to uniquely determine the shape, albedo and reflectance of an object with a general K-lobe dichromatic BRDF, for $K \leq 2$.*

An important exception is the case of retroreflection, when one may have $\eta_i = \mathbf{n}^\top \mathbf{s}$. From the symmetry of the expression for $\boldsymbol{\pi}$ in (23), it follows that $\rho_{\eta_i} = 0$. Consequently, the BRDF may not be uniquely determined in this case.

Experimental Validation. To show that BRDF estimation is possible using camera motion, we again use measured real BRDFs from the MERL dataset. As before, the top row of Figure 3, shows one of six input images, corresponding to five differential motions of the camera. Depth and BRDF are estimated using the theories of Sections 5.2 and 6.2. Compare the similarities in BRDF curves to those recovered using object motion, for the repeated materials `blue-metallic-paint2` and `violet-acrylic`. Appearance from a novel lighting direction is accurately predicted in the third row and a novel geometry is relighted in the fifth row.

The final column in Figure 3 shows results for the `brass` material. From the elongated shape of the specular lobe in the input images, it is clear that the material is anisotropic. However, the estimation using camera motion still yields an isotropic BRDF whose appearance is a good approximation to the original. Further experiments are included in [9].

7 Discussion: Shape, Reflectance and Lighting Recovery

We now consider the problem of jointly recovering shape, reflectance and lighting using motion cues (for convenience, "light direction" in this section also refers to the reflection directions). We show that the linear separability of shape, reflectance and lighting imposed by Propositions 1 and 2 allows a characterization of the hardness of such joint recovery problems. Further, we show how our theory is consistent with prior works like [24,7] and also extends them.

Object Motion. For a BRDF dependent on K reflection directions, image formation is given by (20) and shape recovered as $z = \epsilon_1$ using Proposition 3. Three additional equations of the form (21) are available relating the remaining unknowns $\{\sigma, \rho_{\eta_1}, \cdots, \rho_{\eta_K}, \mathbf{s}, \mathbf{r}_1, \cdots, \mathbf{r}_K\}$, reproduced here for convenience:

$$[\mathbf{n}(\epsilon_1)]_\times \left(\sigma \mathbf{s} + \sum_{i=1}^{K} \rho_{\eta_i} \mathbf{r}_i\right) = \frac{\boldsymbol{\epsilon}'}{1 + \beta \epsilon_1}. \tag{24}$$

Since $[\mathbf{n}(\epsilon_1)]_\times$ is skew-symmetric, only two of the three relations in (24) are independent. Thus, for N pixels (or more precisely, N independent normals), we have $2N$ equations in $(K+1)(N+2)$ unknowns (N unknowns for each of albedo and BRDF-derivatives, two unknowns for each direction). Clearly, the system of equations (24) is underdetermined for any $K \geq 1$. Thus, we may state:

Proposition 8. *With unknown albedo and non-Lambertian dichromatic BRDF, the problem of joint recovery of shape, reflectance and lighting using object motion is underconstrained.*

Despite this apparently negative result, our framework is fruitful for understanding and extending several prior works on shape and reflectance recovery:

(a) First, it matches intuition that joint recovery problems are hard in general cases. For example, estimating even a one-lobe dichromatic BRDF with unknown albedo and light source is ambiguous in a single-image setup [7]. Our theory shows that it stays ambiguous even with object motion.
(b) Second, we observe that for Lambertian surfaces ($K = 0$), we have $2N$ equations in $N + 2$ unknowns, so such joint recovery is well-posed, which validates the solutions obtained by prior works like [24].
(c) Third, for uniform albedo and unknown lighting, reflectance may be recovered for single-lobe dichromatic BRDFs, since we have $2N$ equations in $N+5$ unknowns. This shows that motion cues can help reflectance recovery beyond the single-image setup of [7], where uniqueness may be shown only for the case of known albedo and known lighting.
(d) Next, for the case of uniform albedo and known lighting, reflectance recovery for a dichromatic BRDF with $K = 2$ lobes is mildly ill-posed, since we have $2N$ equations in $2N + 5$ unknowns. Thus, mild external constraints or regularization suffice to recover BRDF in such cases. Additional conditions are imposed in [7] by assuming non-negative and monotonic functions, while estimation is regularized by performing a smooth regression.

(e) Finally, it is conjectured in prior works like [16,7] that BRDF estimation is ill-posed for $K > 2$, even with known shape and lighting. Indeed, when considering motion cues, while object shape may be recovered for such BRDFs, the reflectance recovery involves $2N$ equations in $3N + 5$ unknowns, which is severely ill-posed. Thus, our theory establishes that even motion cues cannot unambiguously recover such BRDFs.

Camera Motion. Considering image formation in (20) dependent on a K-lobe BRDF, shape may always be recovered using Proposition 4. By definition in (23), $\boldsymbol{\pi}$ is independent of albedo. As in Section 6.2, from the definitions of $\boldsymbol{\pi}$ in (15) and (23), the relations for camera motion corresponding to (24) are of the form

$$\sum_{i=1}^{K} \rho_{\eta_i} f_i^j(\mathbf{n}(\tilde{\epsilon_1}), \mathbf{s}, \mathbf{r}_i) \;=\; \frac{\tilde{\epsilon}_{j+1}}{1 + \beta\tilde{\epsilon}_1}, \text{ for known functions } f_i^j \text{ and } j = 1, 2. \quad (25)$$

Since $\pi_3 = 0$ by definition in (15), only two independent relations are available. Thus, for N pixels, we have $2N$ equations in $K(N + 2) + 2$ unknowns.

Proposition 9. *With unknown albedo and a K-lobe dichromatic BRDF, the problem of joint recovery of shape, reflectance and lighting using camera motion is well-posed for $K \leq 1$ and ill-posed for $K > 1$.*

This is a surprising result, since joint recovery of shape, reflectance and lighting has traditionally been considered hard. The above shows that even beyond the traditionally studied Lambertian cases, for many common materials like metals and plastics whose BRDF shows a strong half-angle dependence ($K = 1$), there are enough constraints available to solve such joint recovery problems.

For a BRDF with two lobes, we have $2N + 6$ unknowns, so the system (25) is only mildly ill-posed and may be solved for shape, relfectance and lighting under regularization. Finally, we note that the problem is severely ill-posed for $K > 2$.

8 Conclusions and Future Work

We have presented a framework that helps understand the extent to which object or camera motion cues enable recovery of shape, reflectance and lighting. The theoretical results of Sec. 5, 6 and 7 are summarized in Table 1. These results reflect the intrinsic difficulty of shape and reflectance recovery from motion cues, independent of choice of estimation method. Our framework yields some surprising results on shape and reflectance recovery. In particular, we show both theoretically and in experiments that motion cues can decouple shape and BRDF, allowing both to be simultaneously (rather than alternatingly) estimated for many common materials. Even more unexpectedly, it can be shown that under camera motion, joint recovery of shape, albedo, reflectance functions, lighting and reflection directions is well-posed for some materials (and only mildly ill-posed under object motion). Our future work will explore estimation algorithms that exploit this well-posedness for joint recovery of shape, reflectance and lighting.

Acknowledgments. We thank Ravi Ramamoorthi for helpful discussions and Shen Tian for help with preparing the figures.

References

1. Alldrin, N., Zickler, T., Kriegman, D.: Photometric stereo with non-parametric and spatially-varying reflectance. In: CVPR (2008)
2. Barron, J.T., Malik, J.: Shape, albedo, and illumination from a single image of an unknown object. In: CVPR, pp. 334–341 (2012)
3. Blinn, J.F., Newell, M.E.: Texture and reflection in computer generated images. Comm. ACM 19, 542–547 (1976)
4. Canas, G.D., Vasilyev, Y., Adato, Y., Zickler, T., Gortler, S.J., Ben-Shahar, O.: A linear formulation of shape from specular flow. In: ICCV, pp. 191–198 (2009)
5. Chandraker, M.: What camera motion reveals about shape with unknown BRDF. In: CVPR, pp. 2179–2186 (2014)
6. Chandraker, M., Bai, J., Ramamoorthi, R.: On differential photometric reconstruction for unknown, isotropic BRDFs. PAMI 35(12), 2941–2955 (2013)
7. Chandraker, M., Ramamoorthi, R.: What an image reveals about material reflectance. In: ICCV, pp. 1076–1083 (2011)
8. Chandraker, M., Reddy, D., Wang, Y., Ramamoorthi, R.: What object motion reveals about shape with unknown BRDF and lighting. In: CVPR, pp. 2523–2530 (2013)
9. Chandraker, M.: On joint shape and material recovery from motion cues. Tech. rep., NEC Labs America (2014)
10. Goldman, D.B., Curless, B., Hertzmann, A., Seitz, S.M.: Shape and spatially-varying BRDFs from photometric stereo. PAMI 32(6), 1060–1071 (2010)
11. Horn, B., Schunck, B.: Determining optical flow. Art. Intell. 17, 185–203 (1981)
12. Lawrence, J., Ben-Artzi, A., Decoro, C., Matusik, W., Pfister, H., Ramamoorthi, R., Rusinkiewicz, S.: Inverse shade trees for non-parametric material representation and editing. In: ACM ToG (SIGGRAPH), pp. 735–745 (2006)
13. Lucas, B., Kanade, T.: An iterative image registration technique with an application to stereo vision. In: Image Understanding Workshop, pp. 121–130 (1981)
14. Matusik, W., Pfister, H., Brand, M., McMillan, L.: A data-driven reflectance model. ToG 22(3), 759–769 (2003)
15. Nagel, H.H.: On a constraint equation for the estimation of displacement rates in image sequences. PAMI 11(1), 13–30 (1989)
16. Ngan, A., Durand, F., Matusik, W.: Experimental analysis of BRDF models. In: EGSR, pp. 117–126 (2005)
17. Oxholm, G., Nishino, K.: Shape and reflectance from natural illumination. In: Fitzgibbon, A., Lazebnik, S., Perona, P., Sato, Y., Schmid, C. (eds.) ECCV 2012, Part I. LNCS, vol. 7572, pp. 528–541. Springer, Heidelberg (2012)
18. Romeiro, F., Zickler, T.: Blind reflectometry. In: Daniilidis, K., Maragos, P., Paragios, N. (eds.) ECCV 2010, Part I. LNCS, vol. 6311, pp. 45–58. Springer, Heidelberg (2010)
19. Romeiro, F., Vasilyev, Y., Zickler, T.: Passive reflectometry. In: Forsyth, D., Torr, P., Zisserman, A. (eds.) ECCV 2008, Part IV. LNCS, vol. 5305, pp. 859–872. Springer, Heidelberg (2008)
20. Sato, I., Okabe, T., Yu, Q., Sato, Y.: Shape reconstruction based on similarity in radiance changes under varying illumination. In: ICCV, pp. 1–8 (2007)

21. Seitz, S., Curless, B., Diebel, J., Scharstein, D., Szeliski, R.: A comparison and evaluation of multiview stereo algorithms. In: CVPR, pp. 519–526 (2006)
22. Torrance, K.E., Sparrow, E.M.: Theory for off-specular reflection from roughened surfaces. JOSA 57, 1105–1112 (1967)
23. Verri, A., Poggio, T.: Motion field and optical flow: Qualitative properties. PAMI 11(5), 490–498 (1989)
24. Zhang, L., Curless, B., Hertzmann, A., Seitz, S.: Shape from motion under varying illumination. In: ICCV, pp. 618–625 (2003)
25. Zickler, T., Belhumeur, P., Kriegman, D.: Helmholtz stereopsis: Exploiting reciprocity for surface reconstruction. IJCV 49(2/3), 1215–1227 (2002)