Specular Flow and the Recovery of Surface Structure

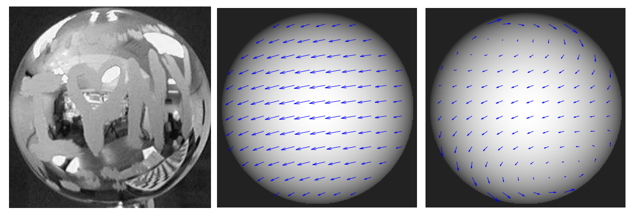
Stefan Roth Michael J. Black
Department of Computer Science, Brown University, Providence, RI, USA
{roth,black}@cs.brown.edu

Abstract

In scenes containing specular objects, the image motion observed by a moving camera may be an intermixed combination of optical flow resulting from diffuse reflectance (diffuse flow) and specular reflection (specular flow). Here, with few assumptions, we formalize the notion of specular flow, show how it relates to the 3D structure of the world, and develop an algorithm for estimating scene structure from 2D image motion. Unlike previous work on isolated specular highlights we use two image frames and estimate the semi-dense flow arising from the specular reflections of textured scenes. We parametrically model the image motion of a quadratic surface patch viewed from a moving camera. The flow is modeled as a probabilistic mixture of diffuse and specular components and the 3D shape is recovered using an Expectation-Maximization algorithm. Rather than treating specular reflections as noise to be removed or ignored, we show that the specular flow provides additional constraints on scene geometry that improve estimation of 3D structure when compared with reconstruction from diffuse flow alone. We demonstrate this for a set of synthetic and real sequences of mixed specular-diffuse objects.

1. Introduction

We address the problem of recovering the image motion and 3D shape of a specular surface viewed by a moving camera. Previous work on this topic has focused on the shape or motion of isolated specular features or specular highlights. While local shape information can be recovered from both static [12] and moving highlights [19], these represent a subset of the general class of specular reflections. More generally, a specular surface may reflect a distorted view of the illumination field surrounding it. In this case the specular reflections may be "dense" or "semi-dense" in that they appear over a relatively large area. When the camera or scene is in motion, these specular reflections give rise to 2D image motion which we call *specular flow* [19, 30]. In this paper we formalize the notion of specular flow, relate it mathematically to the 3D structure of the scene and combine diffuse and specular flow to estimate surface curvature.



(a) *(b)* *(c)*

Figure 1. Moving surface with diffuse markings and specular reflections from an unknown textured scene. Our method relates the 3D surface shape and camera motion to the 2D image motion of diffuse and specular pixels. The algorithm automatically recovers the surface shape and flow and probabilistically classifies pixels as diffuse or specular. *(a)* One image from a sequence. *(b)* Recovered diffuse flow (magnified) corresponding to the motion of the surface texture. *(c)* Recovered specular flow (magnified) corresponding to the motion of the specular texture.

Consider the image in Figure 1 of a reflective, textured surface viewed from a moving camera. The motion of the surface texture gives rise to the diffuse flow, the mathematics of which is well understood. The reflected scene on the other hand gives rise to specular image motion where the relationship to 3D structure is significantly more complex. Our key insight is that the spatial pattern of specular flow from a textured scene provides rich information that can constrain estimates of surface shape. Consequently, rather than treating specular reflections as "noise" to be removed, we treat them as a rich signal which provides information about 3D shape. It is worth emphasizing that, in contrast to prior work, we are *not* focusing on the motion of specular *highlights* but rather on the spatially extended flow fields arising from the reflection of textured scenes.

The paper makes a number of contributions that go beyond the state of the art. (1) We relate specular flow to 3D shape and camera motion. Solving for the specular image motion given shape and camera motion is a difficult problem and we exploit an approximation that makes the estimation of specular flow tractable. (2) We exploit this formulation to recover optical flow using a direct method. This is an extension of layered parametric motion estimation methods [24] and in that spirit we exploit a probabilistic mixture model formulation and solve for the motion using an Expectation-Maximization (EM) algorithm. (3) In solving for the image motion the mixture model provides a classification of the scene pixels as specular or diffuse. (4) The EM algorithm is automatically initialized using an estimate of 3D shape obtained from a dense optical flow field that contains both diffuse and specular flow. To do so we exploit the epipolar constraint to identify flow vectors consistent with the diffuse flow and use these to obtain an initial estimate of 3D structure. (5) We show that 2-frame specular flow constrains the 3D shape and we provide the first formal characterization of these constraints. These constraints are, in general, different from the diffuse motion constraints hence combining diffuse and specular flow may improve the accuracy of shape estimation.

Our technique makes weak assumptions compared with previous work. In particular we do not need to know the illumination field or the illumination direction. Like previous methods we assume known camera motion. Furthermore, frame-to-frame camera motions are assumed to be small. We assume surfaces have distinct regions of specular and diffuse reflectance and that their shape can be approximated by a smooth implicit function or by piecewise quadratic functions [3]. The resulting theory of specular flow extends the prior art to spatially extended, unknown light sources and leads to the first complete algorithm to exploit specular flow for shape recovery.

2. Previous work

The use of specular reflections in estimating scene structure has a long history in computer vision. Many techniques recover information about surface shape by assuming knowledge of the light source shape, its position, and/or its orientation [12, 13, 16, 22, 23]; the illumination is treated as a form of structured light. In general however, we may not know the structure of the illumination field, or even the direction of illumination. In this case a single image provides little information about the surface shape.

Alternatively, many techniques for estimating structure treat specular reflections as violations of Lambertian reflectance, detect them, and remove them from consideration [14, 15, 28]. Treating specular reflections as noise however fails to exploit the relationship between specular motion and surface shape. Methods that do exploit specular motion have typically exploited specular highlights (points or localized regions of specular reflection) and their motion over many frames [19, 27].

2-frames and a known light source. Early specular reconstruction techniques assumed the illuminant was known and used 2-frames with known camera motion (i.e., specular stereo). They showed that the motion of a specular highlight can be combined with the known depth of a nearby point to obtain constraints on local surface curvature [5, 6, 31]. Furthermore, the convexity/concavity of the surface can be determined using the relative motion between a specular highlight and a nearby surface point [31].

Extended temporal constraints. It is possible to recover parts of the surface geometry by tracking a single specular feature over a long series of frames with known camera motion [19]. But in order to uniquely recover the surface profile along the trajectory, the feature needs to be tracked to and from occlusion boundaries. This only provides the surface geometry along the path of the feature.

Optical flow and specular reflections. Many methods have looked at the estimation of optical flow with reflective or transparent surfaces where the traditional brightness constancy assumption is violated. [2, 4, 11, 21, 25, 29]. These methods however have focused on the 2D motion problem and have not looked at using specular motion to estimate 3D shape. Waldon and Dyer [30] introduced the notion of the "reflection flow field" (what we call specular flow) in the case of a stationary camera and moving specular object. They extend the brightness constancy assumption to include two terms: one for diffuse reflectance and one for specular reflectance. They show that the specular flow can dominate the diffuse flow depending on the surface curvature and point out that in such cases the specular flow can be recovered using standard gradient-based optical flow techniques, but did not couple it to the shape of the surface.

In contrast to previous work, we formulate a parametric model of specular flow that directly relates image motion to 3D shape. Rather than focusing on specular highlights over many frames we use the spatially extended specular flow of a textured scene between two frames to constrain surface shape. Finally our only assumption about the illumination field is that it is far from the reflecting surface.

3. Parametric models of diffuse and specular motion

We derive a parametric model of diffuse and specular image motion between two image frames as a function of camera motion and the shape of the viewed object. We assume that the surface is given by a known, twice differentiable implicit function $g(\mathbf{x} \in \mathbb{R}^3; \theta) = 0$, where θ are the parameters of the implicit representation. Furthermore, we assume that both the internal and the external camera parameters are known. We denote the fixed internal parameters of the camera as $\mathbf{K} \in \mathbb{R}^{3\times3}$; the external parameters for frame i are denoted as $(\mathbf{R}_i, \mathbf{t}_i)$, where $\mathbf{R}_i \in \mathbb{R}^{3\times3}$ is the camera rotation and $\mathbf{t}_i \in \mathbb{R}^3$ is the camera translation. The resulting projection matrix for frame i is then written as $\mathbf{P}_i = \mathbf{K} \cdot (\mathbf{R}_i, \mathbf{t}_i)$.

We make the important assumption that the features in the illumination field (*e.g.* traditional light sources as well as

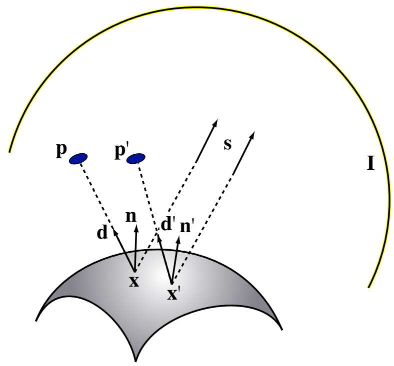


Figure 2. A specular object reflects a feature from an illumination field $\mathbf{I}$ at infinity in direction $\mathbf{s}$. The object is viewed from two different viewpoints $\mathbf{p}$ and $\mathbf{p}'$. The feature is reflected from the surface at points $\mathbf{x}$ and $\mathbf{x}'$ in two different frames with the corresponding surface normals $\mathbf{n}$ and $\mathbf{n}'$. The vectors $\mathbf{d}$ and $\mathbf{d}'$ specify the direction from the reflection point to the respective camera.

illuminated objects in the surroundings) are sufficiently far away so that the lighting direction that induces a particular specular feature in multiple frames remains approximately constant across frames. This is illustrated in Figure 2. The assumption of distant illumination allows us to model the specular image motion without making additional assumptions about the scene surrounding the viewed object.

3.1. Diffuse motion

Given the parametric surface as well as the camera parameters, we compute the parametric motion field that corresponds to motion of diffuse surface features, which we call diffuse flow. For every pixel $\mathbf{u} = (u_1, u_2)$ in frame i, we identify a ray along which the corresponding diffuse surface feature must lie. This ray is given as

$$\mathbf{r}(\lambda) = \mathbf{p} + \lambda \cdot \mathbf{d}, \tag{1}$$

where $\mathbf{p} = -\mathbf{R}_i^{\mathrm{T}} \mathbf{t}_i$ is the camera center and $\mathbf{d} = -\mathbf{R}_i^{\mathrm{T}} \mathbf{K}^{-1} (u_1, u_2, 1)^{\mathrm{T}}$ is the ray direction from the surface toward the viewpoint. To find the surface point $\mathbf{x}$ in 3D that projects to $\mathbf{u}$ we determine the intersection of the ray with the implicit surface so that $\mathbf{x} = \mathbf{r}(\lambda)$ and $g(\mathbf{x};\, \theta) = 0$. In case of multiple intersections, the surface feature corresponds to the intersection that is closest to the camera. Given the surface point $\mathbf{x}$, it can be projected into another view (assuming that it is visible), *e.g.* for frame $i+1$ the image location is given by $\mathbf{u}' = (y_1/y_3, y_2/y_3)$ where $\mathbf{y} = \mathbf{P}_{i+1}\mathbf{x}$. The diffuse flow from frame i to frame $i+1$ is finally given as the difference in image coordinates: $\mathbf{w}_{\mathrm{D}}(\mathbf{u}) = \mathbf{u}' - \mathbf{u}$.

3.2. Specular motion

Computing the specular motion is unfortunately not as straightforward as the diffuse case, even with the assumption that the illumination field is at infinity. To compute the specular motion for a pixel $\mathbf{u}$, we first identify the 3D surface point $\mathbf{x}$ that reflects a specular feature. As for diffuse motion, we find $\mathbf{x}$ by intersecting the viewing ray with the implicit shape. In case of specular reflection the observed feature comes from the unknown, distant illumination field and not directly from the surface of the object. This means that to compute the image motion of the feature, we first have to find the surface point $\mathbf{x}'$ that reflects the same specular feature in frame $i+1$ in which the camera is centered at $\mathbf{p}'$. This is illustrated in Figure 2. To find $\mathbf{x}'$ it is helpful to define the following: The surface normal at point $\mathbf{x}$ is given as the gradient of the implicit surface description, *i.e.* $\mathbf{n}(\mathbf{x}) = \nabla g(\mathbf{x})$. We write the unit normal as $\hat{\mathbf{n}} = \mathbf{n}/||\mathbf{n}||$ and the unit ray direction (from above) as $\hat{\mathbf{d}} = \mathbf{d}/||\mathbf{d}||$. From the law of reflection, the direction $\mathbf{s}$ of the specular feature on the illumination field is

$$\mathbf{s} = 2\hat{\mathbf{d}}^{\mathrm{T}}\hat{\mathbf{n}} \cdot \hat{\mathbf{n}} - \hat{\mathbf{d}}. \tag{2}$$

Based on our assumption of distant illumination, the same illumination direction gives rise to the specular feature in frame $i+1$. For the purpose of finding $\mathbf{x}'$ it will be easier to restate the law of reflection, in particular that the normal direction bisects the viewing and the lighting directions:

$$\hat{\mathbf{d}}' + \mathbf{s} = \tilde{\mathbf{n}}' = \nu \cdot \mathbf{n}', \tag{3}$$

where ν is a scaling factor. Since $\mathbf{x}'$ is unknown at this point, we rewrite this as

$$\frac{\mathbf{p}' - \mathbf{x}'}{||\mathbf{p}' - \mathbf{x}'||} + \mathbf{s} = \nu \cdot \nabla g(\mathbf{x}'). \tag{4}$$

Determining the reflection point amounts to finding an $\mathbf{x}'$ so that (4) holds for some ν under the constraint that $\mathbf{x}'$ is on the surface, *i.e.* $g(\mathbf{x}';\theta) = 0$. Even for simple implicit surfaces such as a sphere this is difficult to do analytically. For finite distance light sources this problem is known as Alhazen's Billiard problem, a classical problem in mathematics first posed by Ptolemy [18, 26].

Even for a circle in 2D the problem is hard: Given a camera at a point **a** and a light source at **b** find the points on the circle which reflect a ray from **a** into **b**. There are in general 4 such points; at most two actually amount to reflection rather than refraction [26]. For objects more complex than a sphere, no closed form solution exists [10]. However, Chen and Arvo [10] propose a local perturbation method that describes the differential behavior of specularities given either camera or light source motion. This has been exploited in computer graphics to render specular reflections efficiently; we use their method to approximate the specular flow.

The method is based on the path that the reflection point takes on the surface as the camera moves from $\mathbf{p}$ to $\mathbf{p}'$. This path can be described using a smooth path function

$\mathbf{f}_\theta : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ that assigns a surface point for each camera location (given the surface parameters θ). It holds that $\mathbf{f}_\theta(\mathbf{p}) = \mathbf{x}$ and $\mathbf{f}_\theta(\mathbf{p}') = \mathbf{x}'$. While evaluating $\mathbf{f}_\theta(\mathbf{p}')$ amounts to the same, difficult problem as above, here we perform a Taylor approximation of the path function around $\mathbf{p}$ based on the assumption that $\mathbf{p}$ and $\mathbf{p}'$ are close:

$$\begin{aligned} \mathbf{x}' &= \mathbf{f}_\theta(\mathbf{p}') && (5) \\ &\approx \mathbf{f}_\theta(\mathbf{p}) + \frac{\partial \mathbf{f}_\theta(\mathbf{p})}{\partial \mathbf{p}} \cdot (\mathbf{p}' - \mathbf{p}) && (6) \\ &\approx \mathbf{x} + \frac{\partial \mathbf{f}_\theta(\mathbf{p})}{\partial \mathbf{p}} \cdot (\mathbf{p}' - \mathbf{p}), && (7) \end{aligned}$$

where $\frac{\partial \mathbf{f}_\theta(\mathbf{p})}{\partial \mathbf{p}}$ is called the path Jacobian, which in contrast to $\mathbf{f}_\theta(\mathbf{p}')$ itself can be expressed in closed form. We omit a detailed derivation of the path Jacobian as it is almost identical to the derivation in [10]. The main difference is that we assume the illumination field to be infinitely far away, which actually simplifies the derivation and makes is possible to compute the path Jacobian without knowing the position of the illumination. This sets the proposed method apart from previous work that often made more restrictive assumptions about the illumination. It is important to note that the only other requirement for computing the Jacobian is that we can evaluate the gradient and the Hessian of the implicit surface equation. This will hold for quadratic surfaces as considered later, but the method extends to more flexible surface representations, such as the ones used in [27].

Summarizing the parametric model of specular flow: Given a pixel $\mathbf{u}$ in frame i, find the associated surface point $\mathbf{x}$. Then approximately find the nearby surface point $\mathbf{x}'$ that causes the same specular feature to appear in frame $i+1$ using the path perturbation method described above. The surface point $\mathbf{x}'$ can be projected into the view at frame $i+1$, to give $\mathbf{u}' = (y_1/y_3, y_2/y_3)$ where $\mathbf{y} = \mathbf{P}_{i+1}\mathbf{x}'$. The specular flow for distant illumination from frame i to frame $i+1$ with small camera motion is finally given as the difference in image coordinates: $\mathbf{w}_\mathrm{S}(\mathbf{u}) = \mathbf{u}' - \mathbf{u}$.

4. Properties of specular flow

We have described how specular flow relates to the 3D geometry of the scene and to the camera motion under the assumption of distant illumination and small camera motion. Now we show that the observed specular flow can be used to constrain the geometry of an unknown object and that it provides additional constraints on surface shape not present in the diffuse flow. For clarity and brevity of exposition we present the insights in an informal way; the formal expression of these is straightforward.

Definition 1. A specular feature viewed from two different viewpoints is said to fulfill the **weak specular constraint**, if the object causing the specular feature is infinitely far away from the specular surface. In this case the reflected feature appears under the same illumination direction at any point on the surface.

Observation 1. *Assume that a single specular feature is observed from two different viewing directions. In general the feature will be seen reflected from two surface points. Then the weak specular constraint constrains two of the four parameters of the associated surface normals.*

Proof. From the law of reflection we know that for each of the two viewpoints the viewing direction, the surface normal, and the lighting direction all lie in a plane. Because the viewpoint and the viewing direction lie in this plane, we can parametrize the plane for each viewpoint using a single parameter, *i.e.* the rotation angle around the viewing ray. By virtue of the weak specular constraint, the lighting direction has to be identical in both cases, so the only admissible lighting direction is given as the intersection of the two planes. The planes intersect because we assumed the viewing directions to differ. Once the lighting direction is known, the two surface normals are given as the vectors that bisect the lighting direction and each viewing direction. □

If we interpret the motion of a specular feature as arising from the motion of a diffuse surface marking on a rigid object, then this virtual feature will generally appear at an incorrect depth. The distance of the feature to the actual object surface depends on the surface curvature as well as the viewing angle [5, 19, 28]. Specular reflections from convex surfaces appear behind the surface; for concave surfaces they lie in front. Consider a simple case where light is reflected off a spherical mirror so that it appears to come from the center of the sphere. As the camera motion goes to zero the location of the virtual feature converges to the point half-way between the surface and the center of the sphere [19, 28].

To understand how diffuse and specular motion constrain the surface geometry, we perform a simple experiment relating the depth and curvature of a spherical surface patch to the camera motion. For known camera motion and known parameters of the surface we can compute the parametric flow as above. We now vary the distance of the patch to the camera as well as the curvature of the patch and compute the average angular error [1] between the associated image motion and the true image motion. Figure 3 shows the results for diffuse and specular flow. The left plot shows that, for small camera motions, diffuse flow constrains the depth of the patch much better than the curvature. The specular flow on the other hand provides a different, and complimentary, constraint that comes from the fact that the virtual depth of a specular feature is a function of both the depth and curvature of the surface. This motivates using the combination of diffuse *and* specular constraints for recovering surface geometry of mixed diffuse/specular surfaces.

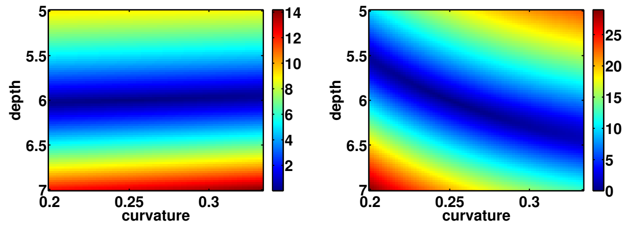


Figure 3. Geometrical constraints of diffuse and specular flow: Average angular error of parametric flow for a spherical surface patch compared to the true parametric flow (depth $= 6$, curvature $= 0.25$). *(left)* Diffuse flow. *(right)* Specular flow.

5. Mixture model: diffuse and specular flow

There are a number of ways to exploit specular flow to recover surface shape. One might first compute dense optical flow and then classify the flow vectors as diffuse or specular using the epipolar constraint [28]. While diffuse motion (of rigid objects) obeys the epipolar geometry, the image motion of specular features does not always obey the epipolar constraint. Using the classified flow vectors we could find the surface parameters by matching the parametric motions derived in Section 3 with the measured diffuse and specular flow. This method has the disadvantage that the classification of flow vectors into specular and diffuse is not always reliable. As [28] notes, the amount of epipolar deviation is dependent on the curvature of the surface. Additionally, the dense estimation of optical flow in the case of mixed diffuse and specular reflection is challenging.

We follow a different approach here and develop a direct method for recovering the surface geometry. In particular, we propose a mixture model of parametric motions, where each pixel can either be diffuse, specular, or belong to an outlier class. The mixture model extends previous work on multiple parametric motions [24]. In the following we assume that the external camera parameters have been determined ahead of time through a separate process.

5.1. Model formulation

We formulate the problem of recovering surface shape from an image pair as one of probabilistic inference. In particular, we find the parameters of the surface shape θ that maximize the likelihood of the image $\mathbf{O}_i$ given the subsequent frame $\mathbf{O}_{i+1}$:

$$\theta^* = \arg\max_\theta p(\mathbf{O}_i \mid \mathbf{O}_{i+1}, \theta). \tag{8}$$

As in [24] we assume conditional independence of the pixels in frame i so that we can rewrite (8) using

$$p(\mathbf{O}_i \mid \mathbf{O}_{i+1}, \theta) = \prod_{\mathbf{u}} p(O_i(\mathbf{u}) \mid \mathbf{O}_{i+1}, \theta), \tag{9}$$

where $O_i(\mathbf{u})$ is the grey-value of image $\mathbf{O}_i$ at pixel $\mathbf{u}$ and we take the product over all the pixels $\mathbf{u}$. Furthermore, we assume that every pixel can be either diffuse, specular, or it can be an outlier, which we formulate using the per-pixel mixture model

$$\begin{aligned} p(O_i(\mathbf{u}) \mid \mathbf{O}_{i+1}, \theta) \; &= m_{\mathrm{D}} \cdot p_{\mathrm{D}}(O_i(\mathbf{u}) \mid \mathbf{O}_{i+1}, \theta) + \\ &\quad m_{\mathrm{S}} \cdot p_{\mathrm{S}}(O_i(\mathbf{u}) \mid \mathbf{O}_{i+1}, \theta) + \\ &\quad m_{\mathrm{O}} \cdot p_{\mathrm{O}}. \end{aligned} \tag{10}$$

m_{D}, m_{S}, and m_{O} are the a-priori mixture weights of the diffuse, specular, and outlier components; p_{D} and p_{S} are the component likelihoods for diffuse and specular pixels, and p_{O} is a fixed outlier probability. In order to specify the component likelihoods, we denote the parametric diffuse flow for a surface with parameters θ at pixel $\mathbf{u}$ as $\mathbf{w}_{\mathrm{D}}(\mathbf{u})$ and the specular flow as $\mathbf{w}_{\mathrm{S}}(\mathbf{u})$. Assuming Gaussian noise, the component likelihood for specular flow is written as

$$\begin{aligned} &p_{\mathrm{S}}(O_i(\mathbf{u}) \mid \mathbf{O}_{i+1}, \theta) = \\ &\qquad \mathcal{N}(O_i(\mathbf{u});\, O_{i+1}(\mathbf{u} + \mathbf{w}_{\mathrm{S}}(\mathbf{u})), \sigma); \end{aligned} \tag{11}$$

the diffuse component likelihood is defined analogously.

Note that, unlike previous mixture model approaches we do not assume a linearized brightness constancy assumption. Instead we adopt a warping approach (see *e.g.* [9]), which assumes $O_i(\mathbf{u}) = O_{i+1}(\mathbf{u} + \mathbf{w}_{\mathrm{S}}(\mathbf{u}))$. This is more appropriate for specular motions where the image motion may be large and a linearization of the brightness constancy assumption would be inappropriate.

5.2. Inference algorithm

Inference of the model parameters is performed using the Expectation-Maximization (EM) algorithm [17], which alternates between probabilistically assigning pixels to causes (specular flow, diffuse flow, or outlier) and maximizing the expected log-likelihood of the model parameters given the assignments.

We omit the detailed derivation here, as our model is similar to other mixture models of parametric motion [24]. It is interesting though to note that the E-step yields ownership weights that segment pixels as being consistent with either diffuse or specular motion. If we denote the components as $c \in C = \{\mathrm{S}, \mathrm{D}, \mathrm{O}\}$, the ownership weights for a pixel $\mathbf{u}$ are given as

$$\pi(c, \mathbf{u};\, \mathbf{O}_i, \mathbf{O}_{i+1}, \theta) = \frac{m_c \cdot p_c(O_i(\mathbf{u}) \mid \mathbf{O}_{i+1}, \theta)}{\sum_{\mathbf{c}' \in C} m_{c'} \cdot p_{c'}(O_i(\mathbf{u}) \mid \mathbf{O}_{i+1}, \theta)}.$$

To simplify inference, we will keep the prior probabilities over the mixture components m_c as well as the component variance σ fixed ($m_c = 1/|C|$ and $\sigma = 10$). We should also point out that maximizing the expected likelihood in the M-step is difficult given the developed model because the specular motion depends on the surface parameters in highly nonlinear ways. With no closed form solution for the model parameters θ we perform a local optimization of the expected likelihood.

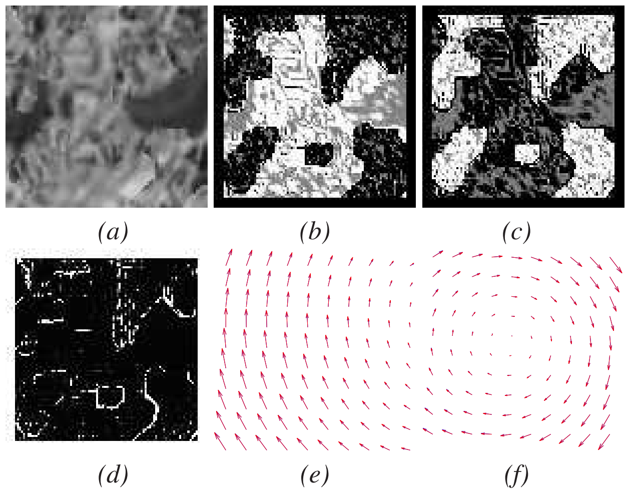


Figure 4. Reconstruction from a synthetic specular/diffuse sequence. *(a)* Example frame from a pair. *(b)* Ownership weights for reconstructed diffuse motion. *(c)* Ownership weights for reconstructed specular motion. *(d)* Outlier ownerships weights corresponding to boundary between specular and diffuse regions. *(e)* Diffuse flow (magnified). *(f)* Specular flow (magnified).

6. Experimental evaluation

We evaluate the performance of surface reconstruction using specular flow with synthetic and real image sequences for which we know the ground truth. In these experiments, we restrict the surface geometry to spherical patches, but use both convex and concave shapes. The implicit surface description in this case is given as

$$g(\mathbf{x};\, \theta) = ||\mathbf{x} - \mathbf{c}||^2 - r^2 = 0, \tag{12}$$

where the reconstruction needs to determine the center and radius of the sphere $\theta = \{\mathbf{c}, r\}$. It is worth noting that the method can be applied in local image patches and does not exploit the outline of the object being viewed.

6.1. Experiments with synthetic data

We rendered 20 image sequences (100×100 pixels) of spherical surface patches with various positions and radii; 10 of them convex and 10 concave. The camera motion is a random rotation around a point near the sphere with a uniformly random rotation axis and a small, random angle. The surface patches exhibit contiguous regions of either diffuse reflectance or specular reflection, which are determined by a random binary surface texture. Both diffuse and specular parts are textured with natural textures. Figure 4 shows such a synthetic frame. To remove a potential bias from the outline of the sphere, the patches are chosen so that the images show only a part of the full sphere. Furthermore, the motion in a boundary of 5 pixels around the edges of the image is ignored to avoid effects from occlusion and disocclusion. Boundary pixels are forced to be explained by the outlier layer and thus do not affect the estimation. Nevertheless, occlusion and disocclusion may still happen between the various specular and diffuse regions since they each have different apparent motions; this is automatically dealt with by the outlier component (Fig. 4 (*d*)).

The EM algorithm is initialized by perturbing the ground truth with random noise. The surface shape is then reconstructed from diffuse and specular motion using 25 iterations of EM as described above. Figure 4 shows the ownership weights (*b* – *d*) and parametric flow fields (*e* – *f*) associated with one of the reconstructed shapes. For the ownership weights, white/black indicates high/low ownership respectively. Note that gray ownership weights correspond to homogeneous regions where both specular and diffuse motion explain the image data.

To determine whether the specular flow helps constrain surface reconstruction we compared against a similar mixture model with only 2 classes: diffuse motion and an outlier class which can account for parts of the image that are not consistent with the diffuse motion. To measure the reconstruction accuracy we computed the error in locating the sphere center and in determining the radius. The median error over all 20 shapes using the combined specular/diffuse model was 0.074 units for the center, and 0.019 units for the radius; for the purely diffuse model, the error was 1.03 and 0.90 units respectively. The difference of the medians is statistically significant according to a rank sum test; the p-values are $1.9 \cdot 10^{-7}$ for the center and $9.1 \cdot 10^{-8}$ for the radius. The diffuse model was sometimes severely affected by the specular motion and converged to solutions far from the true solution. We conclude that for scenes containing significant specular and diffuse motions, the combined model significantly improves the accuracy of surface recovery.

6.2. Reconstruction of a shiny ball

To demonstrate the applicability to real imagery, we captured a shiny ball from 6 different views and reconstructed its position and radius from two of the views. The ball as shown in Figure 1 is highly reflective but also exhibits diffuse reflectance corresponding to surface paint. We captured the ball in a textured environment using a standard digital camera, which was calibrated using the software from [8]. A calibration grid was captured alongside the shiny ball in order to estimate the external parameters of the camera, which was done using the same software. We manually determined the center of the ball in the 6 views and used this to estimate the ground truth position of the ball's center in 3D. The radius of the ball was measured by hand (45mm). To avoid a potential bias from the outline of the sphere, we reconstructed the object only from a masked portion of the image as shown in Figure 5. Pixels outside of the masked region are again forced to be explained by the outlier layer.

Figure 5. Specular reconstruction of a real, reflective object. *(top)* Two views of a specular ball viewed through a circular aperture. Note how the surface paint and the reflections move differently. *(bottom)* Ownership weights of the diffuse flow (left) and the specular flow (right) computed from the reconstructed shape.

To automatically initialize EM with approximate surface geometry, we first compute dense optical flow between the image pair using a standard method [4]. Given the known camera geometry, we classify the flow vectors into specular or diffuse based on the epipolar deviation [28] and reject the putative specular flow vectors with large epipolar deviation from consideration during initialization. From the camera motion and diffuse flow we then estimate the depth of the diffuse surface points. Finally we fit the spherical surface that minimizes the distance to the estimated depth values.

Reconstruction from the initialization was performed as for the synthetic experiments. The reconstruction error of the initialization alone was 15.1mm for the center of the sphere, and 5.1mm for the radius. Running the EM algorithm based on the parametric mixture of specular and diffuse flow lead to substantially improved results: The error for the center and radius was 2.8mm and 0.4mm, respectively. Figure 5 shows the ownership weights as determined by the EM algorithm and Figure 1 shows the parametric diffuse and specular flow fields for the reconstructed shape. We can clearly see that the diffuse layer picks up diffuse features such as the outline of the surface paint, and that the specular flow brings out reflected structures such as the checkerboard, ceiling lamps etc. Parts that do not exhibit any significant spatial structure are assigned equal probability for either being specular or diffuse.

Finally, we compared this against a mixture model with only diffuse flow and an outlier layer. The reconstruction error increased to 20.0mm for the center and 8.9mm for the radius. Again this suggests the benefit of the weak specular flow constraint when reconstructing objects with mixed specular-diffuse appearance.

7. Discussion and future work

We have formalized the notion of specular flow, *i.e.* the image motion that results from a camera moving around a reflective object. Unlike previous work that dealt with single specular highlights, specular flow represents the dense (or semi-dense) flow field that arises from reflective surfaces in textured environments. We showed how, under the assumption of distant illumination fields, specular flow can be related to the 3D geometry of reflective objects defined by parametric implicit functions, and we furthermore presented a method for computing the specular flow approximately under the assumption of small camera motion. Moreover, we analyzed how the assumption of distant illumination, which we termed the weak specular constraint, can be used to constrain the geometry of an unknown reflective object.

Many reflective objects in the world show both diffuse reflectance as well as specular reflection, and thus exhibit a mixture of diffuse flow (the image motion of diffuse surface markings) and specular flow. We explored whether specular flow provides constraints on the geometry of such objects that go beyond those provided by diffuse flow, and found that diffuse and specular flow provide complementary constraints. This motivates combining diffuse and specular cues for surface reconstruction. We proposed a parametric mixture model for surface recovery that is based on parametric models of diffuse and specular flow. Approximate inference is performed using the EM algorithm, which also provides ownership weights that segment the image into diffuse and specular pixels.

We applied the model to synthetic image sequences as well as a real sequence and found that using both specular and diffuse flow leads to a better reconstruction accuracy than reconstruction from the diffuse flow alone.

Given the complexity of the specular flow field in relation to 3D surface geometry it is interesting to ask whether humans use such a cue in judgments of surface shape. Blake and Bülthoff [7] studied the use of specularities in static stereo perception and hypothesized that there is a straightforward connection between specular stereo and specular motion perception. To explore whether humans use specular motion in judging surface shape we performed a series of experiments [20] on the perception of specular flow with random dot displays that remove the dependence on any cues besides diffuse and specular motion. Random dot displays corresponding to diffuse and specular flow were not perceived as coming from a shiny object. Despite this we found that the combination of diffuse and specular flow improved discrimination between convex and concave shapes. Further studies are needed to better understand how well humans are capable of perceiving geometry from specular flow.

For machine vision there are a number of promising directions for future work. The proposed model assumes that the diffuse and specular motions occur in distinct image regions. Future work should consider more general situations related to transparency, where regions can contain both diffuse and specular components at the same time. Furthermore, we have only applied our model to simple parametric surfaces. It will be interesting to study more flexible surface representations, for example a variational framework such as [27]. The method here may also be useful for classifying material properties of surfaces from video. Finally, the reconstruction method assumes that the external camera parameters are recovered through an independent process. Future work should address the joint recovery of camera motion and surface geometry.

Acknowledgments. We thank Tomer Moscovich for assisting with the image capture for the experiments, and Fulvio Domini for helpful discussions. This work was supported by NSF IIS-0535075.

References

[1] J. L. Barron, D. J. Fleet, and S. S. Beauchemin. Performance of optical flow techniques. *IJCV*, 12(1):43–77, 1994.

[2] J. R. Bergen, P. J. Burt, R. Hingorani, and S. Peleg. A three-frame algorithm for estimating two-component image motion. *PAMI*, 14(9):886–896, 1992.

[3] P. J. Besl and R. C. Jain. Segmentation through variable-order surface fitting. *PAMI*, 10(2):167–192, 1988.

[4] M. J. Black and P. Anandan. The robust estimation of multiple motions: Parametric and piecewise-smooth flow fields. *CVIU*, 63(1):75–104, 1996.

[5] A. Blake. Specular stereo. *IJCAI*, pp. 973–976, 1985.

[6] A. Blake and G. Brelstaff. Geometry from specularities. *ICCV 1988*, pp. 394–403.

[7] A. Blake and H. Bülthoff. Does the brain know the physics of specular reflection? *Nature*, 343:165–168, 1990.

[8] J.-Y. Bouguet. Camera calibration toolbox for Matlab. `http://www.vision.caltech.edu/bouguetj/calib_doc/`.

[9] T. Brox, A. Bruhn, N. Papenberg, and J. Weickert. High accuracy optical flow estimation based on a theory for warping. *ECCV 2004*, v. 4, pp. 25–36.

[10] M. Chen and J. Arvo. Theory and application of specular path perturbation. *ACM Transactions on Graphics*, 19(4):246–278, 2000.

[11] N. Cornelius and T. Kanade. Adapting optical flow to measure object motion in reflectance and X-ray image sequences. *ACM Siggraph/Sigart Interd. Workshop on Motion: Representation and Perception*, pp. 50–58, 1983.

[12] G. Healy and T. O. Binford. Local shape from specularity. *Computer Vision, Graphics, and Image Processing*, 42:62–86, 1988.

[13] K. Ikeuchi. Determining surface orientations of specular surfaces by using the photometric stereo method. *PAMI*, 3(6):661–669, 1981.

[14] H. Jin, A. Yezzi, and S. Soatto. Variational multiframe stereo in the presence of specular reflections. *3DPVT 2002*, pp. 626–630.

[15] S. Lin, Y. Li, S. B. Kang, X. Tong, and H.-Y. Shum. Diffuse-specular separation and depth recovery from image sequences. *ECCV 2002*, v. 3, pp. 210–224.

[16] S. K. Nayar, K. Ikeuchi, and T. Kanade. Determining shape and reflectance of Lambertian, specular, and hybrid surfaces using extended sources. *International Workshop on Industrial Applications of Machine Intelligence and Vision*, pp. 169–175, 1989.

[17] R. M. Neal and G. E. Hinton. A view of the EM algorithm that justifies incremental, sparse, and other variants. M. I. Jordan, editor, *Learning in Graphical Models*, pp. 355–368. Kluwer, 1998.

[18] P. M. Neumann. Reflections on reflection in a spherical mirror. *Am. Math. Monthly*, 105(6):523–528, 1998.

[19] M. Oren and S. K. Nayar. A theory of specular surface geometry. *IJCV*, 24(2):105–124, 1996.

[20] S. Roth, F. Domini, and M. J. Black. Specular flow and the perception of surface reflectance. *Journal of Vision*, 3(9):413a, 2003.

[21] B. Sarel and M. Irani. Separating transparent layers through layer information exchange. *ECCV 2004*, v. 4, pp. 328–341.

[22] S. Savarese and P. Perona. Local analysis for 3D reconstruction of specular surfaces. *CVPR 2001*, v. 2, pp. 738–745.

[23] S. Savarese and P. Perona. Local analysis for 3D reconstruction of specular surfaces - Part II. *ECCV 2002*, v. 2, pp. 759–774.

[24] H. S. Sawhney and S. Ayer. Compact representation of videos through dominant and multiple motion estimation. *PAMI*, 18(8):814–830, 1996.

[25] M. Shizawa and K. Mase. Principle of superposition: A common computational framework for analysis of multiple motion. *IEEE Workshop on Visual Motion*, pp. 164–172, 1991.

[26] J. D. Smith. The remarkable Ibn al-Haytham. *Mathematical Gazette*, 76(475):189–198, 1992.

[27] J. E. Solem, H. Aanæs, and A. Heyden. A variational analysis of shape from specularities using sparse data. *3DPVT 2004*, pp. 26–33.

[28] R. Swaminathan, S. B. Kang, R. Szeliski, A. Criminisi, and S. K. Nayar. On the motion and appearance of specularities in image sequences. *ECCV 2002*, v. 1, pp. 508–523.

[29] R. Szeliski, S. Avidan, and P. Anandan. Layer extraction from multiple images containing reflections and transparency. *CVPR 2000*, v. 1, pp. 246–253.

[30] S. Waldon and C. R. Dyer. Dynamic shading, motion parallax and qualitative shape. *IEEE Workshop on Qualitative Vision*, pp. 61–70, 1993.

[31] A. Zisserman, P. Giblin, and A. Blake. The information available to a moving observer from specularities. *Image and Vision Computing*, 7(1):38–42, 1989.