On the Motion and Appearance of Specularities in Image Sequences

Rahul Swaminathan[1], Sing Bing Kang[2], Richard Szeliski[2],
Antonio Criminisi[2], and Shree K. Nayar[1]

[1] Columbia University, New York NY 10027, USA,
srahul@cs.columbia.edu
[2] Microsoft Research,
One Microsoft Way, Redmond 98052, WA, USA

Abstract. Real scenes are full of specularities (highlights and reflections), and yet most vision algorithms ignore them. In order to capture the appearance of realistic scenes, we need to model specularities as separate layers. In this paper, we study the behavior of specularities in static scenes as the camera moves, and describe their dependence on varying surface geometry, orientation, and scene point and camera locations. For a rectilinear camera motion with constant velocity, we study how the specular motion deviates from a straight trajectory (*disparity deviation*) and how much it violates the epipolar constraint (*epipolar deviation*). Surprisingly, for surfaces that are convex or not highly undulating, these deviations are usually quite small. We also study the appearance of specularities, i.e., how they interact with the body reflection, and with the usual occlusion ordering constraints applicable to diffuse opaque layers. We present a taxonomy of specularities based on their photometric properties as a guide for designing separation techniques. Finally, we propose a technique to extract specularities as a separate layer, and demonstrate it using an image sequence of a complex scene.

1   Introduction

Specularities are all around us. There are reflections in windows, monitors, and picture frames, glossy sheen on books and furniture, and bright highlights on coffee mugs and fruits. The presence of specular reflections[1] in an environment is what gives it a true sense of realism. Understanding the behavior of specularities has been one of the long-standing research areas in optics (see Euclid's Optica). In the 18th century, the results obtained in optics became the basis of new and powerful descriptive geometry tools (page 60 of [Kem01]). Descriptive geometry and its novel ways of representing 3D objects and specular reflections, combined

---

[1] In this paper, the term *specularity* or specular reflection is used to describe any non-Lambertian component. *Highlights* are isolated bright spots due to reflections of point or small area light sources, and *reflections* are virtual images with discernible structures.

with the well-established rules of linear perspective, allowed artists and architects to greatly enhance the sense of realism of the environments they depicted.

From an *image-based rendering* perspective, realistic scene synthesis requires extensive appearance capture. However, the number of images required can be reduced using geometry. Also, one must deal with specularities and reflections in real scenes. One approach is that of modeling specularities and reflections as separate layers [LS97]. Thus, the diffuse component of the image sequence contains the structure information while the specular component contains higher order appearance information. We propose a simple technique to recover scene structure as layers while at the same time detecting specularities and estimating the underlying diffuse component.

Many approaches have been suggested to separate the specular and diffuse layers. These can be broadly categorized into physics-based approaches and image-based approaches. Physics-based methods either use the Dichromatic Reflectance Model to recover specularities [KSK88,SHW92] or use the polarizing effect on specularities [Wol90,NFB97,SSK98,FA99]. Image-based approaches include dominant motion based techniques such as [IRP94,SAA00]. However, current techniques (except [KSK88,SHW92]) can only handle planar objects reflected off planar surfaces or require the use of added imaging components such as polarizer filters. Lee and Bajcsy [LB92] proposed a spectral differencing technique to detect specularities from multiple views. However, the diffuse component could not be estimated nor could it deal with dis-occlusion events.

Many vision algorithms that compute scene structure typically ignore specularities. For example, it has been common practice in stereo to ignore specular scenes or to treat specular pixels as outliers. This is because specularities alter the appearance of a scene point from multiple views, which can lead to errors during the matching process. However, attempts to make use of specularities to recover the shape of objects have also been demonstrated in the past [BN98, BB88,Ike81,HB88,Sch94,ON97]. In [ON97], the estimated 3D caustic surface is used to recover shape and classify image features as being either specular or diffuse.

The problem of recovering structure while at the same time separating the diffuse and specular layers is very hard under general camera motions. We simplify our analysis by assuming linear camera motion with constant velocity. This enables us to illustrate our ideas using the concept of Epipolar Plane Image (EPI) [BB89] in a spatial-temporal volume (a stacked image sequence). An EPI is a cross-section of the spatial-temporal volume corresponding to a particular scan-line.

Using the EPI framework, we study the motion of specularities not only in the scene but also within each EPI. Unlike in [ON97], we perform the analysis within the EPI framework. We define metrics to distinguish between traces of specular and diffuse features in the EPI and study the factors on which they depend (§ 2 and § 3). These metrics are also experimentally validated for real specular objects.
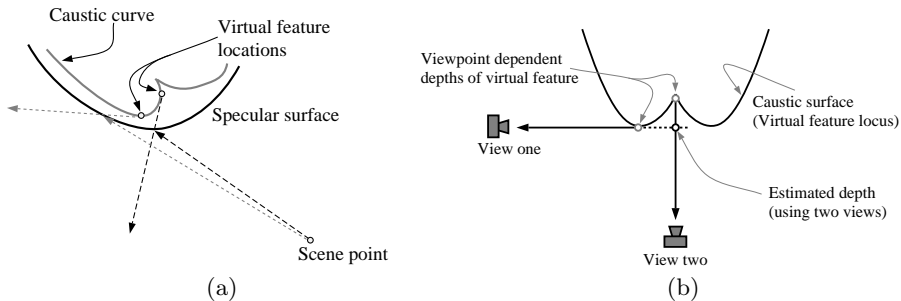
Fig. 1. Reflections on curved surfaces: (a) The geometry of reflection on curved specular surfaces. The position of the virtual feature at two viewpoints lies on the caustic curve at two distinct points. Any point on the caustic is visible only along the tangent to the caustic at the point. (b) Simple two view stereo algorithms estimate an erroneous depth for the virtual feature.

We also show the limits to which geometry alone can be used to separate the two layers and propose the use of photometry as well. We build a taxonomy of specular reflections in § 4, which aids in the design of hybrid algorithms to separate the diffuse and specular layers of a scene. Finally, we demonstrate the effectiveness of our approach on a real scene with specularities (§ 5).

2   Specular Motion in 2D

In general, for flat surfaces, we know that the reflected scene point (virtual feature) lies at a single point behind the surface. However, for curved surfaces, the position of the virtual feature is viewpoint dependent (Fig. 1)[ON97]. The locus of the virtual feature is a *catacaustic* [Ham28], referred to in this paper as just a caustic.

2.1   A Circular Reflector

For purposes of demonstration we assume the specular curve (in 2D) to be circular. The caustic is defined by the geometry of the specular curve and the scene point being reflected. Thus, we can compute the caustic curve in closed form [BS73,BG84].

Given a camera position, we derive the point on the caustic where the virtual feature is seen. It's image location is simply a projection of the caustic point onto the image plane. We derive the image location of a virtual feature as a function of camera pose, specular surface geometry and the scene point.

To compute the EPI trace of the specularities, we assume that the camera motion is linear in the plane parallel to the imaging plane. The linear camera motion implies that the EPI trace of any scene point must lie along straight lines within the EPI-slice. However, reflected points move along their caustic. Thus, their EPI traces would be expected to be curved.
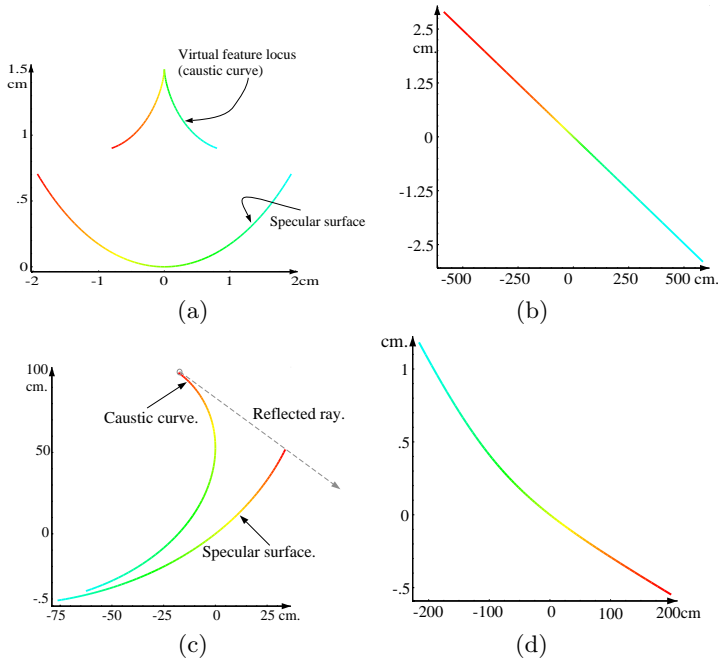
Fig. 2. Plots of actual surfaces, associated caustic curves, and EPI traces. Please note that correspondence between points on the actual surface, caustic curve, and EPI trace, is color-coded. (a) A high curvature surface, such as a soda can, for which the caustic curve is also small and has high curvature. (b) The corresponding EPI trace is almost linear since the virtual feature undergoes minimal motion. (c) An extreme case: In the vicinity of the drawn reflected ray, the camera observes reflection on an almost flat surface (e.g., a monitor screen) at an oblique angle. The corresponding part of the caustic has the least curvature. Thus for small viewpoint changes, the virtual feature moves significantly. (d) The corresponding EPI trace is noticeably bent and therefore appears non-rigid.

We define the deviation of an EPI-trace from a straight line as *disparity deviation* (DD). Disparity deviation depends entirely on the movement of the virtual feature and distance of the viewer from the scene. Motion along the caustic in turn depends on the curvature of the surface, surface orientation and the distance of the reflected point from the surface. The greater this distance, the greater the motion along the caustic surface.

Fig. 2 shows sample EPI curves for two specular curves. Surprisingly, the curve with higher curvature shows little disparity deviation. Although high curvatures lead to faster angular motion along the caustic, this motion is contained within a very small area. Lower curvatures, on the other hand, can produce noticeable disparity deviation in the EPI. For a given curvature, disparity deviation is accentuated at grazing angles of reflections (as we show below).
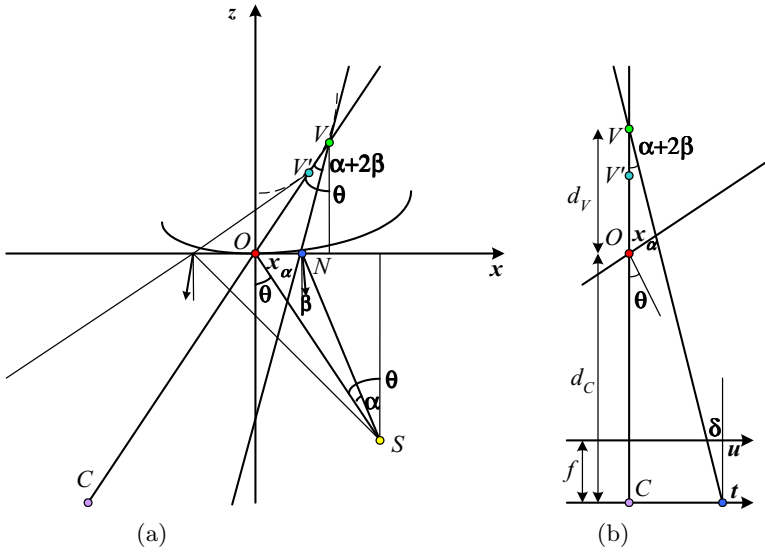
Fig. 3. 2D analysis of specular reflection: (a) Reflection of surface point S by the curved surface at point O as seen by camera C; (b) Projection of the reflected image into the camera C with image plane u moving along the t axis.

2.2   Infinitesimal Motion

We now characterize the local behavior of a specularity, starting with the 2D case (Fig. 3). The scene point being reflected is at S, the camera is at C, and the reflected surface point O is at the origin, with the surface normal along the Z-axis. The incident angle to the surface is θ, while the surface itself has a curvature $\kappa = 1/\rho$.

Consider an infinitesimal change of angle $\alpha = \angle OSN$ in the direction of the light ray leaving S. This corresponds to a motion along the surface from O to N of length x_α,

$$x_\alpha = d_S[\sin\theta - \sin(\theta - \alpha)], \tag{1}$$

where, d_S is the distance from S to O. At the new reflection point N, the surface normal has changed by an angle $\beta = \kappa x_\alpha + \frac{1}{2}\dot\kappa x_\alpha^2 + O(x_\alpha^3)$. Thus, while the incidence angle is $\theta - \alpha$, the emittance angle is $\theta - \alpha - 2\beta$.

This emittance angle determines the angle $\angle OVN = \alpha + 2\beta$, where V is the virtual image point, formed by the intersection of the reflected ray at the origin and the reflected ray at the new point N. We obtain

$$x_\alpha = d_V[\sin\theta - \sin(\theta - \alpha - 2\beta)], \tag{2}$$

where, d_V is the distance from V to O.

Equating (1) and (2) and taking the limit of d_V as $\alpha \to 0$ gives us

$$\lim_{\alpha\to 0} d_V = \frac{d_S}{1 + 2d_S\kappa\cos\theta}. \tag{3}$$

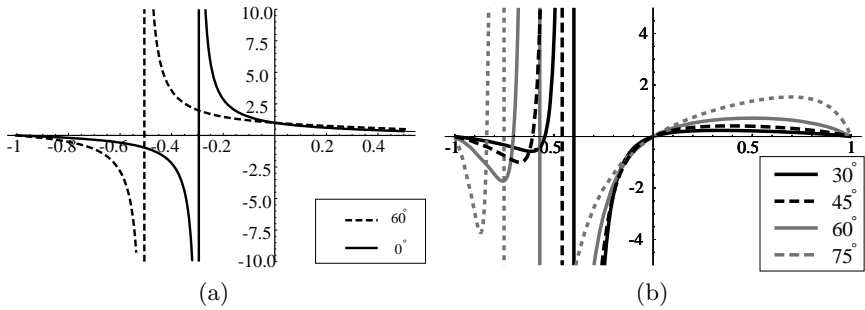**Fig. 4.** (a) Plot of virtual depth d_V as a function of curvature κ for $d_S = 1$ and $\theta = 0°$ and $60°$. (b) Disparity deviation for $f = 100$ as a function of κ for $d_S = 1$, $d_C = 4$, and $\theta = 30°, 45°, 60°, 75°$. The horizontal axis in both cases is actually $2/\pi \tan^{-1} \kappa$, so that the full range $\kappa = (-\infty, 0, \infty)$ can be visualized.

In the limiting case as $d_S \to \infty$ or $\kappa \to \infty$ ($\rho \to 0$), i.e., as the scene point distance becomes large relative to the radius of curvature, we get $d_V = \frac{\rho}{2} \sec \theta$. This result is quite intuitive: the virtual image sits at the focal point behind (or in front of) the reflector for head-on viewing condition, and further away for tilted surfaces.

The behavior in the general case when the source is closer to the surface is plotted in Fig. 4(a). The virtual depth slowly decreases for a convex reflector as the curvature increases. For a concave reflector, the virtual depth decreases, moving rapidly towards negative infinity as the radius of curvature approaches the object distance (as the object approaches the focal point), and then jumps back to positive virtual depths. The actual distance seen by the camera is $d_V + d_C$, so that impossible apparent depths only occur when $d_V < -d_C$.

These results are consistent with the shapes of the caustics presented previously for the circular reflector. Now, how does the disparity (curvature in the EPI) change as we vary the camera position? In other words, what is the disparity deviation of a specular feature? From Fig. 3(b), we see that the disparity D is given by $D = \delta/t = f/(d_V + d_C)$. To see how D varies with t, we apply the chain rule to obtain

$$\dot{D} = \frac{\partial D}{\partial t} = \frac{f d_V^3}{(d_V + d_C)^3} \left(\dot{\kappa}(1 + \cos 2\theta) + 4\frac{\kappa}{d_S} \sin \theta + 2\kappa^2 \sin 2\theta \right). \qquad (4)$$

(The full derivation is given in our technical report [CKSS02].) Notice that there is no disparity deviation for planar reflection, i.e., $\dot{D} = 0$ when $\kappa = \dot{\kappa} = 0$, as expected.

We can now examine each component in (4). The first ratio $(d_V/(d_C + d_V))$ becomes large when $d_C \approx -d_V$, i.e., when the virtual image appears very close to the camera, which is also when the disparity itself becomes very large. The term that depends on the curvature variation $\dot{\kappa}$ decreases for slanted surfaces. It is most significant for undulating surfaces. At their inflection points, the apparent location of the virtual image can move very rapidly.
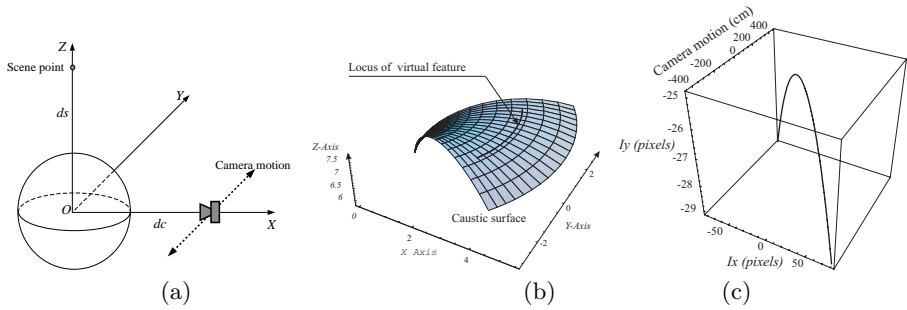
(a)   (b)   (c)

Fig. 5. (a) Analytic setup showing the location of the scene point in relation to the specular surface and camera path. (b) Section of the 3D caustic surface associated with (a). The thin curve on this surface is the locus of virtual features and is neither stationary nor planar. (c) The corresponding EPI-curve clearly exhibits significant epipolar deviations.

The term κ/d_S might at first appear to blow up for $d_S \to 0$, but since d_V is proportional to d_S, this behavior is annihilated. However, for moderate values of d_S, we can get a strong disparity deviation for slanted surfaces. The last term is strongest at a 45° surface slant. It would appear that this term would blow up for large κ, but since d_V is inversely proportionally to κ in these cases, it does not.

To summarize, the two factors that influence the disparity deviation the most are (1) when $d_C + d_V \approx 0$, which is when disparities are very large to start with (because the camera is near the reflector's focal point), and (2) fast undulations in the surface. Ignoring undulations, Fig. 4(b) shows how $\dot{D}$ varies as a function of κ for a variety of slants, with $d_S = 1$ and $d_C = 4$. Therefore, under many real-world conditions, we expect the disparity deviation to be small enough that treating virtual features as if they were real features should work in practice.

3   Specular Motion in 3D

We now discuss the effect of specularities in 3D, again using the caustic surface to perform our analysis. We present our results for a spherical reflector although the results can be extended to arbitrary surface geometries.

Consider a spherical specular surface whose center lies at the origin. The scene point being reflected is located along the positive Z-axis at a distance d_S from the origin. We again derive the caustic surface using the Jacobian technique [BS73,SGN01]. To study the motion of specularities, we assume the camera to move in the X, Y-plane, parallel to the Y-axis at a distance d_C from the origin (Fig. 5(a)). Note, this camera path is not critical to the results we derive.

We need to derive the image location of a virtual feature as a function of camera pose. From the setup in Fig. 5(a), the Z-axis forms the axis of symmetry for the caustic surface (see [SGN01]). Thus, all rays that reflect on the spherical surface, must pass through the axis at some point. Hence, for any camera position, the Z-axis and the camera location determine a plane on which the virtual

feature must also lie. Given the camera pose and caustic surface, determining the position of the virtual feature is now reduced to a 2D problem for which an analytic solution exists (details are in [CKSS02]). The image location of the virtual feature is simply a projection of the derived virtual feature onto the image plane.

3.1   Epipolar Deviations for a Spherical Reflector

Under linear camera motion, the images of a rigid scene point must all lie on the same epipolar line (or plane). However, the motion of a virtual feature on the caustic surface (Fig. 5(b)) violates this constraint. As seen in Fig. 5(c), the image of the virtual feature does not lie on a single scan-line. We refer to this phenomenon as *epipolar deviation (ED)*.

In general, epipolar deviations depend on three primary factors: surface curvature, orientation of surface, and distance of the camera from the reflecting surface. We only consider scene points distant from the surface as they usually produce the largest caustic surfaces [CKSS02]. We now analyze each factor for its contribution to ED. This study helps determine situations when ED effects can be neglected and when they provide significant cues to the presence of highlights and specularities.

Surface Curvature: We know that for planar mirrors, the virtual feature is stationary at a single point behind the surface. Similarly, high curvature surfaces such as sharp corners have very localized tiny caustic surfaces. Between these two extreme curvatures, surfaces exhibit higher epipolar deviations as seen in Fig. 6(a).

Surface Orientation: The angle of incidence of an observed reflection is also critical to epipolar deviation. The more oblique the incidence, the greater the motion of the virtual feature along the caustic surface, causing larger ED. From Fig. 6(b) we can see how ED drops to zero at an angle which corresponds to the plane in which the caustic curve is planar. Beyond this point, the virtual feature locus is again non-planar and causes epipolar deviations. As one moves to near-normal reflections, we see that the feature locus is restricted to the cusp region of the caustic. This implies very small feature motion, in turn reducing ED.

Camera Distance: As camera distance from the scene increases, disparity between scene points decreases. Thus, decreasing disparities imply lower virtual feature motions, in turn decreasing epipolar deviation (Fig. 6(c)).

To empirically validate these analytical results, we took a series of pictures of a mirrored ball at different distances and orientations, and manually plotted the specularity trajectories (measured to the nearest pixel). As seen in Fig. 6(d–f), the results of our experiments are in agreement with our theoretical prediction.

In general, specular reflections or virtual features do not adhere to epipolar geometry. In our geometric analysis, we assume large camera field of view and range of motion, and on occasion, large scene distances. However, in typical real situations, both the camera's range of motion and field of view are limited; as
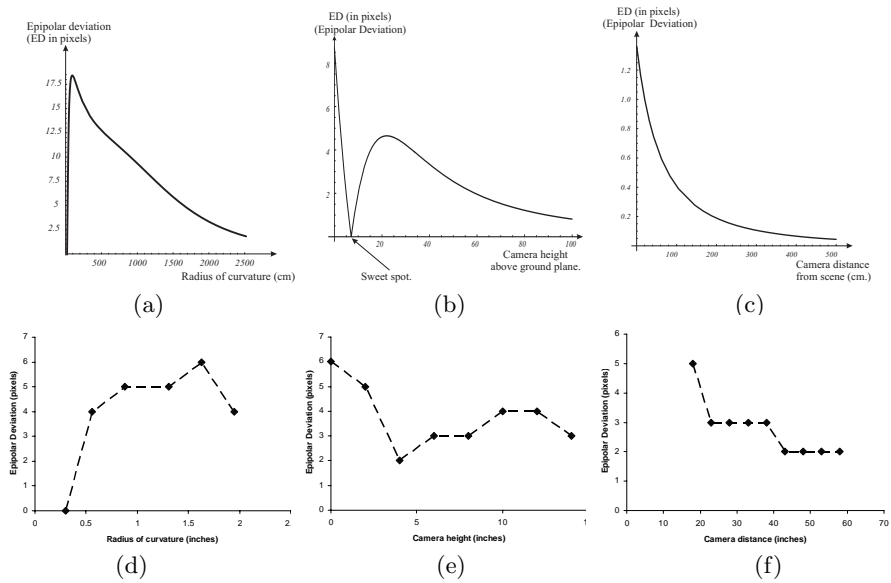
Fig. 6. Epipolar Deviations as a function of the three most significant factors: (a) surface curvature (b) surface orientation (c) camera distance from scene. (d-f) are the corresponding results of experiments using real objects. (d) We used reflective balls with radii ranging from 1.95 to 0.3 inches; each was placed about 3 feet away from the camera. (e) The ball of radius 1.95 inches was placed 3 feet away from the camera. The height of the ball was changed up to 14 inches. (f) The same ball was used, with the distance of the camera to the ball varied from 1.5 to 5 feet. Notice the similar trends between the theoretical and experimental plots.

a result, the specular features appear to adhere closely to epipolar constraints. This makes it hard to disambiguate between specular and diffuse streaks in the EPI. Thus, for any diffuse-specular separation technique to be effective, pure geometric analysis may be insufficient. Of course, other effects such as changing shape of the reflection is a geometric cue towards the reflector's geometry. However, such area based approaches are beyond the scope of this paper. We propose the use of photometric constraints to dis-ambiguate the problem.

4   Photometry of Specularities

We now present a photometric analysis of specularities under linear camera motion. Within the framework of EPIs, we develop a taxonomy of specularities and motivate the need for hybrid algorithms that use geometric and photometric constraints to separate the diffuse and specular layers. In the following discussion, we define an *EPI-strip* to be a strip within an EPI where all the associated physical points are contiguous and share a *common fate* (e.g., similar color or true depth).

Fig. 7. Taxonomy of specularities with example snapshots of sequences. Below each image is the EPI associated with the marked scan-line. Note that all of the EPIs were sheared for visual clarity.

4.1 Taxonomy of Specularities

We categorize the type of observed specularities based on whether the reflecting and reflected surfaces (which we term *reflector* and *source* respectively) are textured (Fig. 7). Furthermore, we differentiate between area and point sources, since this has an impact on how separation can be accomplished.

Textured reflector–Textured source: The EPI-strip associated with this type of specularity is characterized by a blending between the reflector and source textures leading to a criss-cross pattern. Effective techniques for separation tend to analyze the entire area (e.g., [SAA00,Wei01]).

Textured reflector–Textureless area source: In this case, most of the EPI strip is brightened by a uniform color associated with the source. This causes ambiguity in separation, as the modified part of the strip may be misclassified as a separate strip.

Textured reflector–Textureless point source: In principle, this is similar with the previous case, except that the source is highly localized (Fig. 7). As a result, separation can be accomplished by analyzing constant color sub-strips of the EPI-strip, e.g., using the Dichromatic Reflectance Model [KSK88].

Textureless reflector–Textured source: This presents a similar problem as the textured reflector-textureless source case. If, however, we are able to extract
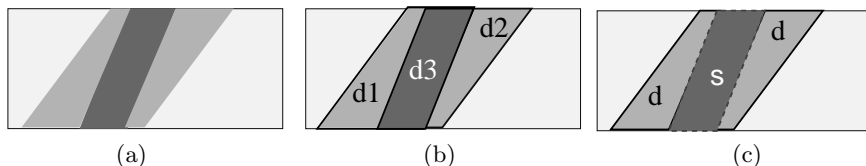
Fig. 8. (a) A typical EPI in which a smaller EPI-strip (darker region) is enclosed with another EPI-strip (lighter region). This EPI has many interpretations in the absence of prior scene knowledge. (b) One interpretation could be that each thin strip of alternating colors, represents a "region" in the scene. Thus, each (d_1, d_2, d_3) is understood to be an unique Lambertian region. (c) Another interpretation could segment the larger EPI-strip (d) includes the darker EPI-strip (s) within it. In this case, we analyze the inner strip to check for specularities and would be separated as a separate layer if necessary.

the EPI-strip associated with the textureless reflector, we can then apply Bajcsy *et al.*'s multi-chromatic model [BLL96] to extract the underlying layers.

Textureless reflector–Textureless area source: If most of the EPI-strip associated with the textureless reflector is superimposed by the textureless source, then again separation is difficult without prior knowledge of the scene. A more detailed explanation of this case is given in the following section (§4.2).

Textureless reflector–Textureless point source: As with the textured reflector – textureless point source case, there exist separation techniques, including the Dichromatic Reflectance Model [KSK88].

4.2   EPI-Strips and Their Inherent Ambiguity

There exists an inherent ambiguity in EPI analysis for specularities and diffuse regions when considering individual EPI. Fig. 8(a) illustrates such an EPI. One EPI-strip (darker) is completely enclosed by another EPI-strip (lighter). Individual layers can now be extracted in one of many ways leading to valid and unique interpretations.

Fig. 8(b) is one interpretation where each EPI-strip was extracted separately representing three unique diffuse layers $(d_1 \ldots d_3)$. The varying tilts of their bordering edges in the EPI lead to slanted segments in the scene of varying depths. In contrast, another equally valid extraction includes the inner EPI-strip (Fig. 8(c)). If this inner strip conforms to the photometric constraints discussed earlier, we interpret it as a specularity s over the otherwise diffuse region d.

Such ambiguities arise in purely Lambertian scenes as well as those containing occlusions. In principle, one can reduce the ambiguities by analyzing multiple EPIs all at once. However, this still does not guarantee an ambiguity-free scenario.

4.3 Surface Curvature and Specularities

Within an EPI, closer scene surfaces have a more horizontal orientation than those farther away. Also, for opaque surfaces, these more horizontal EPI-strips run across (occlude) more vertical EPI-strips.

In the presence of specularities, these depth ordering constraints are often violated. The manner in which these violations occur are cues to the local surface curvature in the scene.

For convex surfaces, the locus of virtual features resides behind the surface. The corresponding EPI-strip of specular reflection has a more vertical orientation than that of the underlying diffuse component (Fig. 9(a,b)). In contrast, concave surfaces typically form the virtual feature in front of the surface. The EPI-strip of the specular component is therefore more horizontal.

On separating the diffuse and specular component within an EPI-strip, one can estimate the local curvature of the surface. If the specular EPI-strip component is more horizontal, the underlying surface must be concave. However, if the specular EPI-strip is more vertical, the surface is most likely to be convex.

5 A Technique for Removing Specularities

We now describe a technique to remove the specular components from an image sequence and estimate the underlying diffuse colors associated with the specular regions.

The technique first extracts EPI-strips from each EPI. Each EPI-strip is then rectified so that trails within it are vertical. The rectified EPI-strip is then analyzed for specularities using a variant of [SAA00]. Our technique is more general in that it is designed to work with textured reflectors and all the three types of sources shown in the first row of Fig. 7 and is not constrained to planar surfaces. Once separation is achieved, the corresponding region within the EPI is marked and excluded from future computations. This process is repeated for all other EPI-strips within the EPI.

5.1 Extracting Good EPI-Strips

A critical stage of the technique is the extraction of EPI-strips within each EPI. Since the camera is assumed to move linearly, the EPI-strips of scene points should be straight lines. As a result, edge detection with straight line fitting is sufficient to detect the bounding sides of EPI-strips.

The slope of an edge in the EPI is directly related to the distance of the associated point from the viewer. We order the edges based on their distance from the viewer, beginning with the closest strip and ending with the farthest. Pairs of depth-ordered edges define candidate EPI-strips. Each EPI-strip is rectified (sheared) such that the slanted bounding edges are vertical (Fig. 9(b)). Every column of the rectified EPI-strip now corresponds to a stabilized scene point viewed over time.
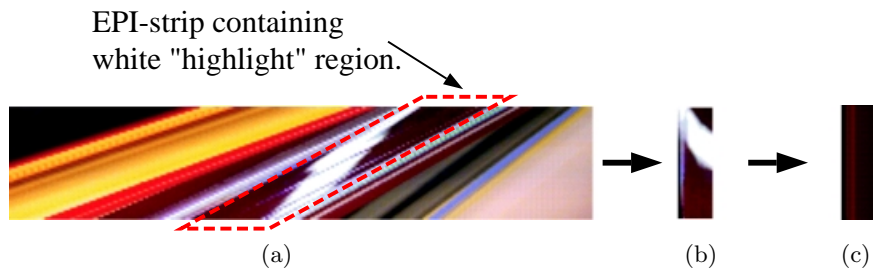
EPI-strip containing
white "highlight" region.

| | | |
| (a) | (b) | (c) |

Fig. 9. An EPI of the sequence in Fig. 10. (a) Typical highlight pattern seen on convex specular surfaces. Chromatically, the highlight region seems to occlude the underlying texture of the surface. However, the orientation of the highlight is more vertical implying a farther depth. This confirms the bright pattern to be caused by a specularity. (b) Rectified section of the marked EPI-strip. The diffuse component is now made vertical, while the specular component is oriented beyond $90°$. (c) Using photometric analysis along with geometric reasoning, the highlight is extracted to reveal the underlying diffuse component.

An EPI-strip candidate is deemed to be good if most of the edges within the rectified strip are vertical. However, specularities and occluding regions would produce slanted edges within the rectified EPI-strip. Since we order the edges from nearest to farthest, the closer EPI-strips would have already been extracted, leaving behind specular strips. Rather than defining goodness of an EPI-strip candidate using photometric constraints, a metric based on predominant orientation is far more robust in practice (details are in [CKSS02]).

5.2   Specularity Extraction

An EPI-strip that is selected is then analyzed for the presence of highlights. The scenario assumed here is that of a textured reflector with an arbitrary source. Many highlight regions tend to be saturated in parts. To simplify our process, we look for specularities in EPI-strips containing pixel intensities above a predefined minimum value.

In any column of the rectified EPI-strip, the pixel with lowest intensity gives us an upper bound on the diffuse component of that scene point. For every column, we estimate this upper bound and assume the scene point to have the associated color. The residual is then the specularity. To validate this step, we group all pixels that are strongly specular and observe their relative orientation within the EPI-strip. If they have a more vertical orientation, then they must be specularities. Note that this is only true for convex surfaces. In our current implementation, we do not consider the effect of concave reflectors.

5.3   Results Using an Image Sequence of a Real Scene

To validate our technique, we took an image sequence of a real scene that contains both specular and Lambertian objects. The camera was mounted on a linear
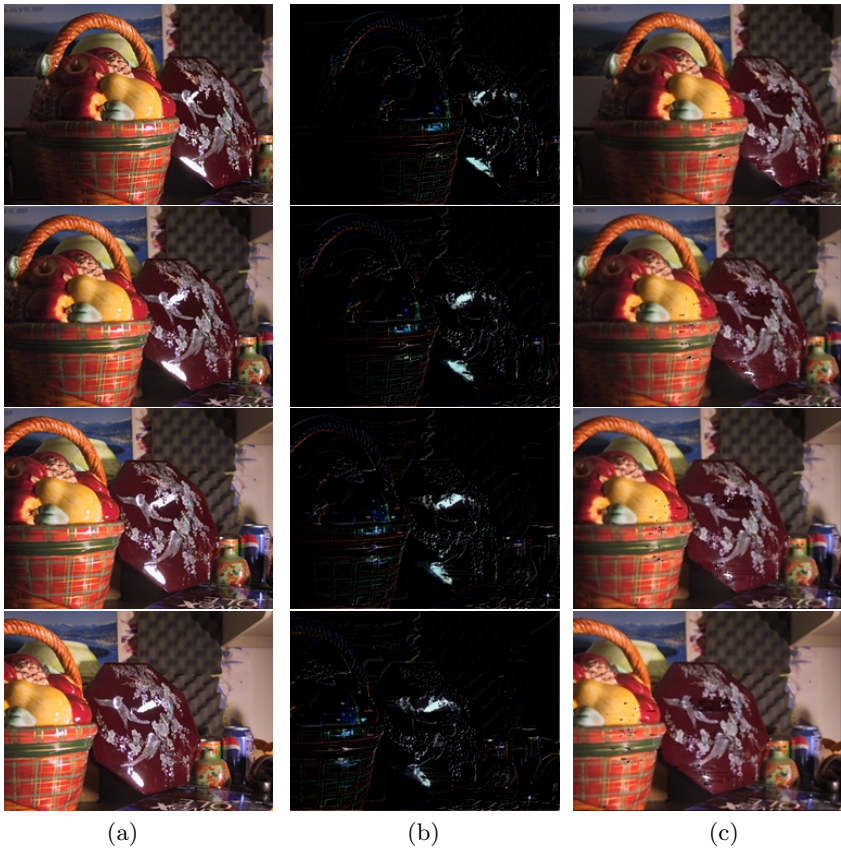
(a)   (b)   (c)

Fig. 10. (a) A subset of our images of a real scene. (b) On performing EPI analysis using a combination of geometric as well as photometric constraints, the specular components were removed from the image sequence. The two strong highlights regions in sequence (a) are robustly detected. (c) The diffuse component video stream is almost void of specular effects. Some artifacts show up in this sequence due to incorrect EPI-strip selection as well as re-sampling issues while computing the rectified EPI-strip (Fig. 9(b)).

translation stage about 3 feet away from the scene. A set of 50 images were captured at uniform intervals as the camera was translated from left to right. A subset of the acquired images can be seen in Fig. 10(a).

This sequence of images were then stacked together to form a spatio-temporal volume on which the above EPI analysis was performed. As seen from Fig. 10 (b), the specular regions were effectively segmented out from the image sequence. Also, the underlying diffuse component of the scene was recovered successfully (Fig. 10(c)).

However, inaccurate EPI-strip extraction and interpolation issues while creating the rectified EPI-strip result in some visible artifacts (black spots and residual specularities in Fig. 10(c)). Since we use a relatively simple technique

to detect and separate specular layers, the results are quite sensitive to the EPI-strip selection process.

6   Conclusions

In this paper, we first present a geometric analysis of the behavior of specularities in typical scenes. We study their image traces under linear camera motion and introduced the metrics *disparity deviation* (DD) and *epipolar deviation* (ED) to characterize specular motion. We show that these deviations depend on the surface curvature as well as the orientation of the specular surface. One might expect that reflections from curved surfaces would always produce curved EPI traces. Surprisingly, both flat and highly curved surfaces do not produce significant deviations. Instead, it is the mildly curved (especially convex) surfaces that produce the largest deviations. In addition, the closer the object, the larger the deviations tend to be.

Our findings suggest that it might be difficult to differentiate diffuse from specular components using geometric constraints alone. As a result, we supplement our geometric analysis with photometric considerations, which make up the second part of our paper. We develop a taxonomy of specular reflections to aid in the design of hybrid algorithms that use both geometric and photometric constraints. Finally, we present results on a real image sequence, using our hybrid algorithm to separate the two components into different layers.

In the future, we would like to move from the "local" edge-based approach for selecting EPI-strips to a more global approach. This should significantly help the photometric analysis phase. Our ultimate goal is to able to separate diffuse and specular components under general camera motion.

References

[BB88]   A. Blake and G. Brelstaff. Geometry from specularity. *Proc. International Conference on Computer Vision*, pages 394–403, 1988.

[BB89]   H. H. Baker and R. C. Bolles. Generalizing epipolar-plane image analysis on the spatiotemporal surface. *International Journal of Computer Vision*, 3(1):33–49, 1989.

[BG84]   J. W. Bruce and P. J. Giblin. *Curves and Singularities*. Cambridge University Press, Cambridge, 1984.

[BLL96]   R.K. Bajcsy, S.W. Lee, and A. Leonardis. Detection of diffuse and specular interface reflections and inter-reflections by color image segmentation. *International Journal of Computer Vision*, 17(3):241–272, March 1996.

[BN98]   D. N. Bhat and S. K. Nayar. Stereo and specular reflection. *International Journal of Computer Vision*, 26(2):91–106, February 1998.

[BS73]   D. G. Burkhard and D. L. Shealy. Flux Density for Ray Propagation in Geometrical Optics. *Journal of the Optical Society of America*, 63(3):299–304, March 1973.

[CKSS02] A. Criminisi, S.B. Kang, R. Swaminathan, and R. Szeliski. Separating Diffuse and Specular Layers using EPI Analysis, MSR-TR(in preparation). Technical report, Microsoft Research, 2002.

[FA99]   H. Farid and E.H. Adelson. Separating reflections and lighting using independent components analysis. In *Proc. Computer Vision and Pattern Recognition*, pages I:262–267, 1999.

[Ham28]   W. R. Hamilton. Theory of Systems of Rays. *Transactions of the Royal Irish Academy*, 15:69–174, 1828.

[HB88]   G. Healey and T. O. Binford. Local shape from specularity. *CVGIP*, pages 62–86, 1988.

[Ike81]   K. Ikeuchi. Determining surface orientation of specular surfaces by using the photometric stereo method. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 3(6):661–669, 1981.

[IRP94]   M. Irani, B. Rousso, and S. Peleg. Computing occluding and transparent motions. *International Journal of Computer Vision*, 12(1):5–16, January 1994.

[Kem01]   M. Kemp. *Visualizations: The Nature Book of Art and Science.* University of California Press, 2001.

[KSK88]   G. J. Klinker, S. A. Shafer, and T. Kanade. The measurement of highlights in color images. *International Journal of Computer Vision*, 2(1):7–32, June 1988.

[LB92]   S.W. Lee and R. Bajcsy. Detection of specularity using colour and multiple views. In *Proc. European Conference on Computer Vision*, pages 99–114, 1992.

[LS97]   J. Lengyel and J. Snyder. Rendering with coherent layers. In *Computer Graphics Proceedings, Annual Conference Series*, pages 233–242, Proc. SIG-GRAPH'97 (Los Angeles), August 1997. ACM SIGGRAPH.

[NFB97]   S.K. Nayar, X. S. Fang, and T. Boult. Separation of reflectance components using color and polarization. *International Journal of Computer Vision*, 21(3):163–186, February 1997.

[ON97]   M. Oren and S. Nayar. A theory of specular surface geometry. *International Journal of Computer Vision*, 24(2):105–124, September 1997.

[SAA00]   R. Szeliski, S. Avidan, and P. Anandan. Layer extraction from multiple images containing reflections and transparency. In *Proc. Conference on Computer Vision and Pattern Recognition*, volume 1, pages 246–253, June 2000.

[Sch94]   H. Schultz. Retrieving shape information from multiple images of a specular surface. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(2):195–201, February 1994.

[SGN01]   R. Swaminathan, M. D. Grossberg, and S. K. Nayar. Caustics of Catadioptric Cameras. In *Proc. International Conference on Computer Vision*, pages II:2–9, July 2001.

[SHW92]   S. A. Shafer, G. Healey, and L. Wolff. *Physics-Based Vision: Principles and Practice.* Jones & Bartlett, Cambridge, MA, 1992.

[SSK98]   Y. Schechner, J. Shamir, and N. Kiryati. Polarization-based decorrelation of transparent layers: The inclination angle of an invisible surface. In *Proc. International Conference on Computer Vision*, pages 814–819, 1998.

[Wei01]   Y. Weiss. Deriving intrinsic images from image sequences. In *Proc. International Conference on Computer Vision*, pages II: 68–75, 2001.

[Wol90]   L.B. Wolff. Polarization-based material classification from specular reflection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(11):1059–1071, November 1990.