



Computer Science and Artificial Intelligence Laboratory
Technical Report

MIT-CSAIL-TR-2009-051

October 22, 2009

Shape from Sheen

Roland W. Fleming, Antonio Torralba, and Edward
H. Adelson

Shape from sheen

Roland W. Fleming, Antonio Torralba and Edward H. Adelson

Introduction

Glossy surfaces are everywhere. If you stand in a typical bathroom and look around, you'll see that most of the surfaces surrounding you are lustrous or glossy. Glazed tiles, gleaming faucets, polished basins and plastic shampoo bottles are all peppered with specular highlights.

Given that highlights are so common it seems plausible that the visual system might somehow be able to make use of them in the estimation of three-dimensional shape. Indeed, it is now well established that specular reflections aid shape perception in the presence of other cues (Blake and Bühlhoff, 1990, 1991; Mingolla and Todd, 1986; Norman, Todd and Orban, 2004; Todd and Mingolla, 1983; Todd, Norman, Koenderink and Kappers, 1997), and in fact are sufficient on their own to confer a vivid sense of shape (Fleming, Torralba, Dror and Adelson, 2003; Fleming, Torralba and Adelson, 2004). This is demonstrated in Figure 1, which shows a computer-generated image of a perfectly specular (*i.e.* mirrored) surface. Despite the fact that the image contains no shading, texture or other cues to 3D shape, we nevertheless have a vivid impression of its 3D shape.

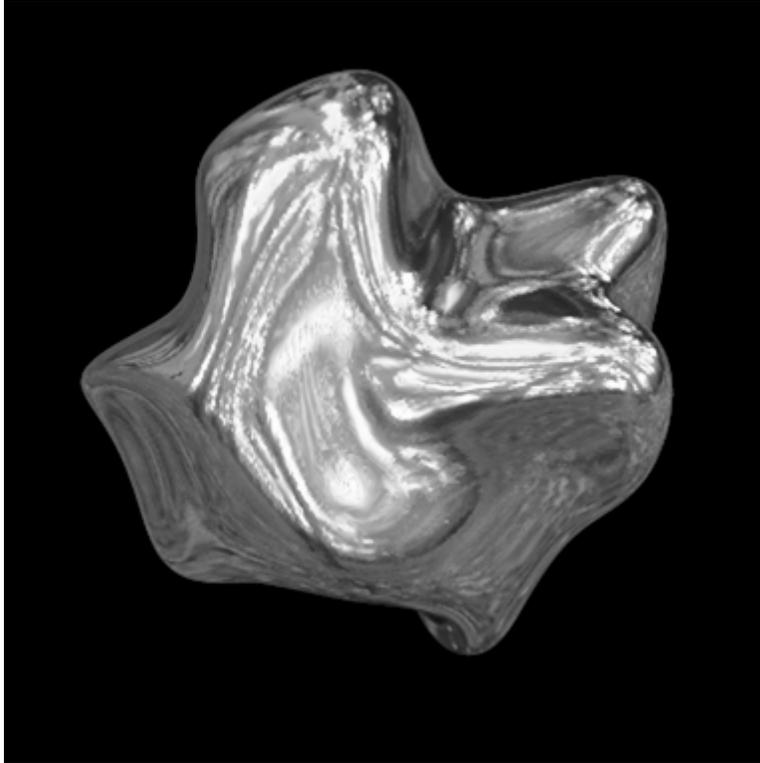


Figure 1. A computer-generated image of a perfectly specular surface reflecting a natural environment. Even though the image contains no stereo, motion, shading or texture cues, we still have a vivid sense of the object's 3D shape. This chapter is about the image information that can be used to recover 3D shape from specular reflections.

This chapter is about how the visual system could recover 3D shape from specularities, from a theoretical point of view. We start by discussing some of the problems that specular reflections seem to pose to the visual system. We then argue that the visual system can use the way that reflections are distorted across the surface to recover some important constraints on 3D shape. We show that the pattern of distorted reflections remains surprisingly even when different scenes are reflected in the surface of the object. The key idea presented in this chapter is that relatively simple image measurements can sometimes carry unexpectedly powerful information about the 3D world.

The optical misbehaviour of mirrors

Glossy surfaces exhibit a number of behaviours that make them potentially highly problematic for human vision. Specular reflections are quite paradoxical. On the one hand they are one of the simplest interactions that light can have with matter. On the other hand they create distorted, impenetrable images of the world on the surface of an object, which writhe and flow over the surface as the object is moved. We will now discuss some of the qualities and quirks that make mirror reflections unique.

According to the optical equations of Fresnel (1821/1866), whenever a light-ray encounters an interface between two media of different optical density, some of the light will be reflected in a specular way. For specular reflection, the angle of the reflected ray is equal to the angle of the incident ray relative to the surface normal, as shown in Figure 2. This is the paradigmatic concept of reflection that we encounter as schoolchildren.

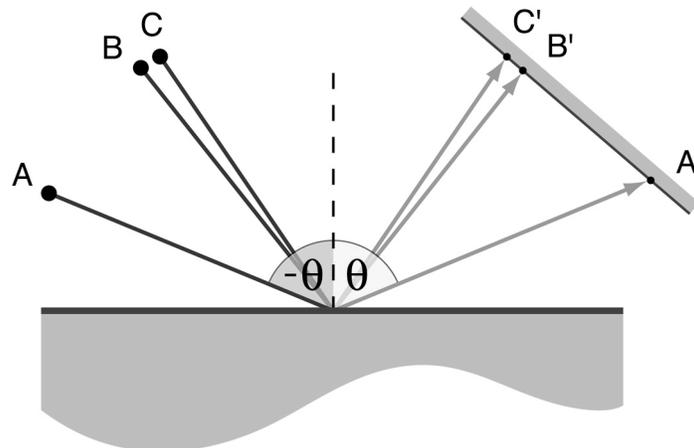


Figure 2. The geometry of specular reflection. A light ray arriving from point **A** strikes the surface at angle $-\theta$ relative to the surface normal (dashed line). The reflected ray exits the surface at angle θ on the other side of the surface normal, projecting in the direction of **A'**. Note that neighbouring points in the world, **B** and **C**, are reflected to neighbouring points, **B'** and **C'** in the virtual image.

Although the geometry of specular reflection is extremely simple, it has a number of consequences that are important for the following discussion. The first is simply that a specular surface tends to generate a virtual image of the world that is facing the surface. To anyone who is familiar with mirrors this, of course, comes as little surprise. But it is worth contrasting this with matte surfaces, which also reflect light but do not form images of the world on their surfaces, because the reflected light is diffusely sprayed out of the surface in random directions, as shown in Figure 3. By contrast, as can be seen in Figure 2, for smooth specular surfaces, *neighbouring points in the world project to neighbouring points in the reflected image*. This becomes rather important in our later analysis.

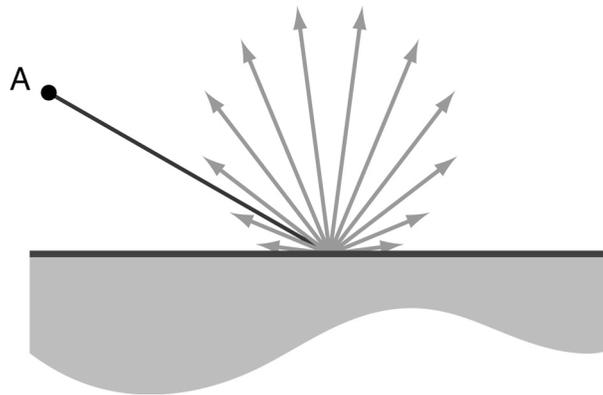


Figure 3. Ideal diffuse (Lambertian) reflection. A light ray arriving from point **A** strikes the surface and is sprayed out in random directions. Consequently, the material appears matte and does not form an image of the world on its surface.

Second, mirror reflection is *angle doubling*. The reflected ray does not exit along the surface normal. If this were the case then the angle of the reflected ray would depend only on the orientation of the surface, and not on the position of the source. Instead, the

ray reflects at *twice* the angle between the source and the surface normal (*i.e.* at the same angle on the other side of the surface normal). This means that if you rotate a mirror through a given angle, for example, 45deg, the direction of the reflected ray changes by twice this angle, *i.e.* 90deg (hence ‘angle doubling’). One counter-intuitive consequence of this is that a hemispherical mirror contains reflections of the entire sphere of incoming directions. This is depicted in Figure 4. The fringes of the surface contain reflections of points in the world that are located *behind* the hemisphere.

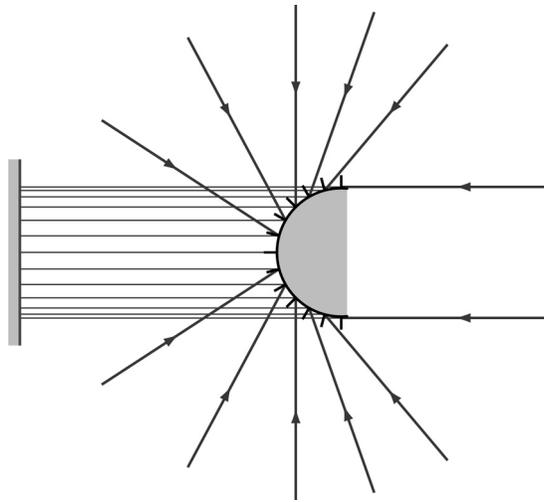
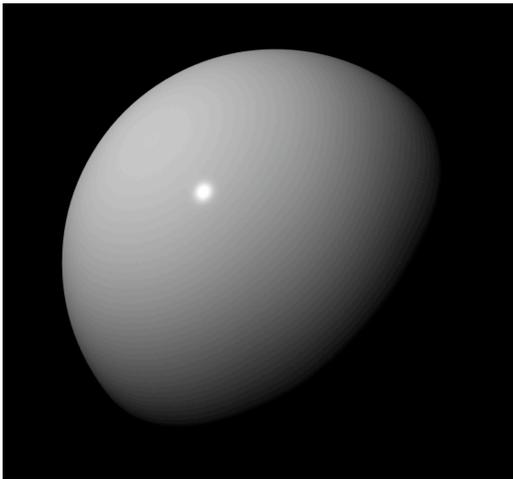


Figure 4. Because specular reflection is angle-doubling, a mirrored hemisphere projects points from the entire sphere of incident directions into the image. Incident rays that arrive from behind the hemisphere are nevertheless visible in the surface.

The angle-doubling behaviour of mirrors leads to another key difference between specular and diffuse reflections. Imagine a sphere that is illuminated by a single point light source. If the sphere has diffuse surface reflectance, then the brightest point on the surface is the one that points directly at the light source. This is the brightest point as it receives the greatest incident energy per unit surface area. By contrast, if the sphere is specular, the brightest point on the surface is the location whose surface normal points exactly *half way* between light source and the viewer’s line of sight. Except for the

special case where the light source arrives from the same position as the viewer, this will not be the same place on the surface as for the diffuse sphere. This is demonstrated in Figure 5. Here we use a single glossy sphere (*i.e.* with both diffuse and specular components). Note that the isophotes reveal that the image intensities have two distinct peaks, one for the diffuse component and one for the specular component.

(a) raw image



(b) isophotes

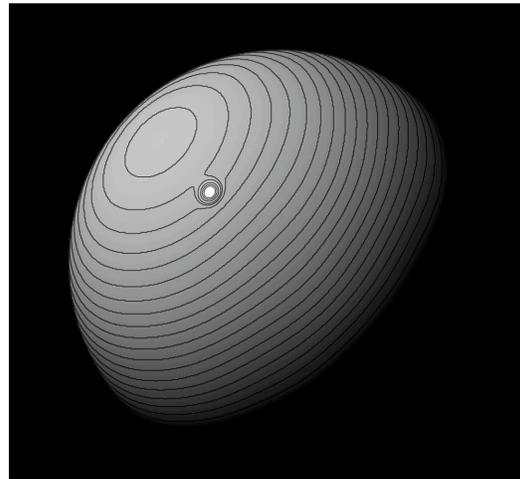


Figure 5. Image (a) shows a glossy surface illuminated by a single point light source. The isophotes in (b) reveal two distinct peaks in the image intensities, one resulting from the diffuse reflectance component, and one from the specular reflection.

A third way in which specular reflection is unusual is the fact that the position of specular highlights is extremely unstable. Unlike texture markings, specular reflections twist and slide across the surface whenever the viewer, object or light source moves. The movement of specular reflections over the surface gives mirrors their distinctive visual fluidity. In a series of extremely compelling demonstrations, Hartung and Kersten (2002, 2003) have shown that the characteristic motion of specular reflections is crucial for their lustrous, chrome-like appearance. Hartung and Kersten ‘attached’ specular reflections to the underlying surface, so that they moved with the surface instead of sliding across it.

When the patterns are fixed to the surface, they tend to look like matte surface markings, even though in any static frame from the motion sequence they appear like lustrous mirror reflections.

Konderink and van Doorn (1980) demonstrated that the rate at which the highlight moves across the surface depends on its curvature. When an object is rotated, highlights move slowly around regions of high curvature, but dart rapidly across flatter regions.

The mobility of specular highlights means that they also have unusual binocular behaviour. Under binocular viewing, the two eyes view the surface from different viewpoints. This means that the position of the highlight relative to the surface is different for the two eyes. Note that this is quite unlike texture markings or shadows, which remain attached to the surface irrespective of the position of the viewer. Because highlights appear at different locations for the two eyes, they do not fall at the same stereoscopic depth as the surface that generated them. For convex geometries, they fall behind the surface, while for concave surfaces they fall in front, as depicted in Figure 6. Blake and Bülthoff (1990, 1991) have shown that the binocular placement of highlights affects how the visual system interprets them. For example, if a highlight is placed unnaturally at the same depth as the surface, it tends to look like a stain or marking instead of a reflection.

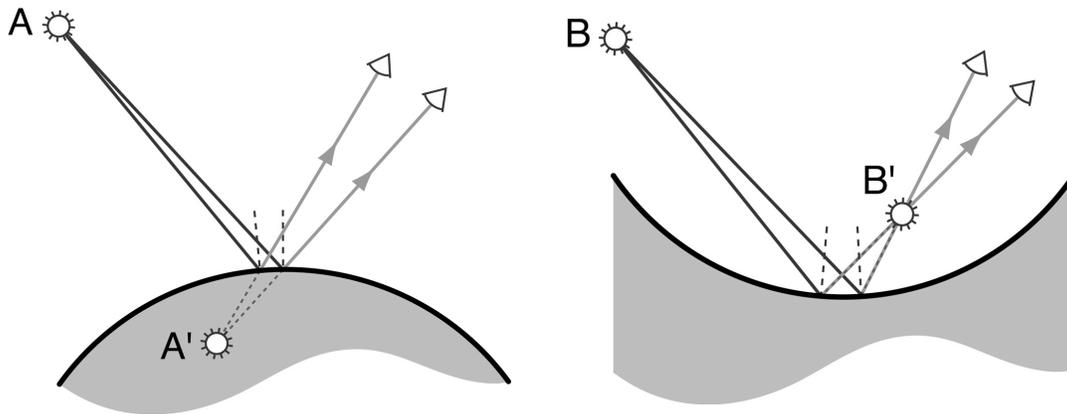


Figure 6. Stereoscopic depth placement of specular highlights. When A is reflected in a convex object, its reflection A' falls behind the surface. Conversely, when B is reflected in a concave object, its reflection B', floats in front of the surface.

For the current discussion, the most important consequence of these effects is that it is not trivial to recover shape from motion parallax or binocular stereopsis when the surface is specular. If the visual system treated reflections as if they were simply surface markings, diffuse shading, or cast shadows, then it would almost always arrive at an erroneous estimate of 3D shape.

Mirrored objects are even more recalcitrant if you consider how they appear in single, static images. If you look at a textured object, the contrasts that occur in the retinal image are caused directly by the markings on the surface. For example, where there is a change from white paint to dark paint on the surface, there will be a corresponding change from light to dark intensities in the image. This is not true of mirror reflections. If you look into a mirrored surface, all you will see is a reflection of the other objects in the scene. This is depicted beautifully in a number of works by M.C Escher, such as *Hand with Reflecting Sphere* (1935), which is shown in Figure 7. The image of the sphere contains reflections of the artist's room, his books and furniture, and

even himself. However, it does not contain any shading patterns or texture of its own. To put it fancifully, there is a sense in which a perfectly mirrored surface is ‘invisible’ because none of the features that can be seen in the surface belong to the mirror itself. Although you can look ‘through’ or ‘into’ a perfect mirror, you can never look directly at the surface itself¹.



Figure 7. *Hand with Reflecting Sphere*, Lithograph by M. C. Escher (1935). Note that the image of the sphere contains nothing other than reflections of the surrounding objects, including the artist himself. [Permissions necessary ?]

One consequence of this is that the image produced by a mirrored object changes completely depending on the world surrounding the object. As the object is moved from

¹ Indeed, in mediaeval alchemy, quicksilver was thought to be endowed with metaphysical properties (specifically it was thought to represent the mind) because the curious practitioner who gazes into a bead of mercury sees nothing but a reflection of himself. Mercury, of course, has numerous other unusual properties, such as being a dense metal that is liquid at room-temperature, and having a meniscus that curves in the opposite direction from most commonly encountered liquids. These other extraordinary attributes doubtless also contributed to its mystical reputation.

scene to scene different surroundings will be reflected in the surface, and so the image changes dramatically depending on the context. It is hard to imagine that the visual system can make sense of the infinitely variable patterns of light and dark that are generated by a mirrored object.

To make matters worse, because the image is just a reflection of the surrounding scene, it is possible to create almost any arbitrary image by carefully manipulating the environment in which the object is placed. For example, an object can be made to appear to have a bump or dent in its surface simply by distorting the pattern reflected in the surface. To the visual system, the input would be the same, so the visual system would have no way of knowing whether the environment or the shape of the object is responsible.

Because specular reflections ‘misbehave’ in so many ways, it might seem quite impossible to recover the shape of an object from the twisted patterns that are reflected in its surface. Indeed, in computer vision specularities have often been treated simply as a nuisance, and a number of schemes exist for removing them or mitigating their effects. In summary, we have seen that:

- (1) Specularities slide around relative to the surface whenever the object, viewer or light sources move, so the visual system cannot recover shape using standard structure for motion processing.
- (2) Specularities fall at the ‘wrong’ stereoscopic depth, so the visual system cannot use standard binocular processing to recover surface depth.
- (3) In the case of a perfectly mirror, the image consists of nothing more than a distorted reflection of the environment surrounding the object.

- (4) A single 3D shape can produce many different images (depending on the context), while a single image is consistent with many different 3D shapes.

Despite this, as we saw in Figure 1, a single static image of a perfectly mirrored object placed in a complex realistic setting generally yields a vivid sense of 3D shape. Somehow we appear to be solving the task. In the next section we discuss how the visual system might go about using specular reflections for 3D shape estimation. Our starting point is the following observation. Even though the image of the mirrored object only contains reflections of the other objects in the scene, nevertheless, the way that these reflections are *distorted* can furnish the visual system with information about 3D shape.

Approaching the problem of estimating 3D shape from specular reflections

If you place an object in front of a curved mirror, such as a polished kettle or the back of a spoon, you will notice that the reflection of the object is warped out of shape. An example of this is shown in Figure 8. Note that the reflection of the square is deformed: the sides are not parallel, and indeed, they are not even straight. Note also that the other objects in the surrounding scene are smeared and twisted out of shape.

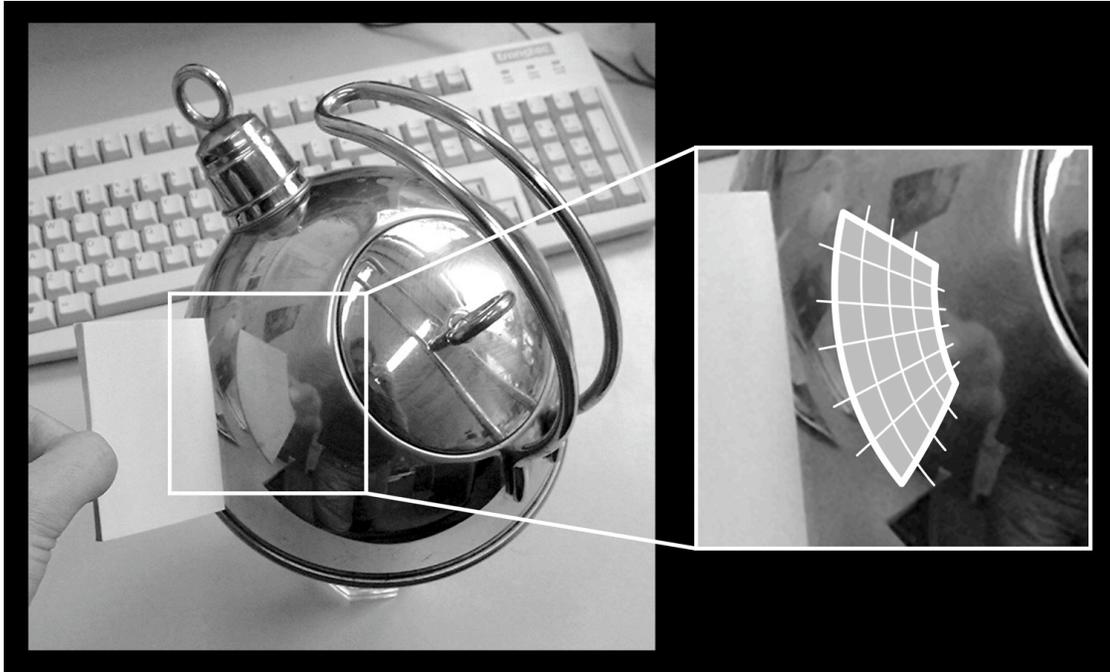


Figure 8. Distorted reflections in the surface of a metallic kettle. Note that square card is warped out of shape in the reflected image. This is emphasized in the close-up. If the visual system could estimate the deforming transformation that has been applied to the card, it could in principle use this information to infer the 3D shape of the kettle.

Importantly, the distortions introduced by the reflecting surface depend directly on its 3D shape. We will discuss the relationship between 3D shape and image distortion in more detail shortly. For the moment the important point is that the way that a reflection is distorted bears the signature of the reflecting object's geometrical form: tacit in the warps and twists of the reflection lies information about the 3D shape of the underlying surface. This means that, at least in principal, the visual system could use the way that a reflection is distorted to recover information about the object that is doing the reflecting.

This opens a strategy for recovering 3D shape. The visual system could first identify a recognizable object within the reflection of the environment. Having segmented the reflected object's boundary, the visual system could then estimate the

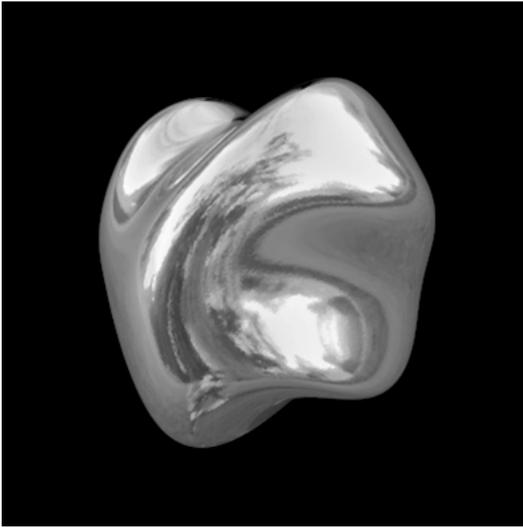
deforming transformation that has been applied to the boundary. In theory, once this transformation is known, the visual system could infer the 3D shape of the surface that is responsible for the distortion.

However, one major problem with this approach is that the visual system can only interpret the *distorted* image of an object, if it knows what the *undistorted* object looks like. This means that the visual system has to accurately locate the corresponding features in the surrounding scene, or at least make strong assumptions about the world (e.g. lines are generally straight rather than curved).

There are a number of reasons for thinking that the visual system probably does not recover shape from specular reflections in this way. The first is simply that it seems unlikely that the visual system forms such a precise estimate of the scene surrounding an object. When we want to know the shape of a highly polished object, we do not generally look around the room first to estimate the positions of all the structures in the scene; it is usually sufficient just to look directly at the object of interest.

Second, we can get a good sense of an object's 3D shape even when the scene surrounding the object is not visible. In Figure 9, an object has been cropped out of its original context and shown against a neutral black background. This means that there is no information about the surrounding scene except what can be seen in the surface itself. Despite this we have a vivid sense of the object's 3D shape.

(a) natural environment



(b) 1/f noise

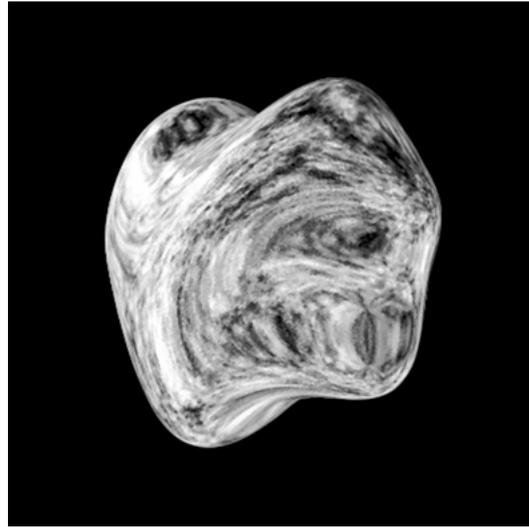


Figure 9. An irregularly shaped mirror that has been cropped out of its original contexts and placed on a neutral background. In (a) the object is rendered in a natural environment. Note that at some locations it is difficult to discern what objects are reflected in the surface. In (b) the same object is reflecting a pattern of random noise. Despite the fact that no recognizable objects are present in the reflections, we nevertheless get a fairly compelling sense of the object's 3D shape.

We have verified this impression psychophysically (Fleming *et al*, 2004). Using the gauge figure task, we have found that subjects are capable of estimating the 3D surface orientations of mirrored objects reliably and accurately even when the objects are shown against a black background. This suggests that the human visual system does not need to know the precise locations of objects in the surrounding scene in order to interpret distorted reflections.

Third, note that at many points on the surface the reflection is so smeared that the reflected object is totally unrecognizable. The reflections look more like a streaky texture than a recognizable scene. It seems highly unlikely, then, that the visual system could find correspondences between locations in the scene and their reflections in the surface.

In fact, we can get a rather good impression of the 3D shape even when no recognizable structures are reflected in the object's surface. Figure 9(b) shows the same

shape reflecting a pattern of random noise. The pattern reflected in the surface of the object does not contain recognizable objects like buildings or trees, or even extended edges that the visual system could assume are straight. Instead, the patterns look more like a texture that has been warped onto the surface.

The very fact that the reflections look more like a smeared texture than a recognizable scene suggests that perhaps we should change our conception of the patterns that are reflected in the surface. Instead of conceiving of the reflected world as an aggregation of discrete physical bodies, perhaps we should think of it instead as a sort of continuous texture that is stretched over the surface. If this were the case then the visual system might be able to recover shape from specular reflections by analogy to the way that it recovers shape from texture. In the next section, we elaborate this way of thinking about the patterns that impinge on specular surfaces.

The environmental patterns that illuminate a mirrored object.

In the real world, light is emitted from luminous sources, such as the sun, and bounces many times before it reaches our eyes. Light reverberates between surfaces, reflecting back and forth from object to object, and spraying out in all directions to fill the world with light. The reverberation of light creates a dense and highly structured field of luminous flux throughout the environment, which Gibson (1979) called the *ambient optic array*. In practice, the ambient optic array means that a point in the real world typically receives light from almost every direction.

Consider what the world looks like from this isolated point in space. Rays converge on this point from every direction, bearing light from all the surfaces in the

surrounding scene. These rays form *an image* of the surrounding world. In particular, we can represent the set of all converging rays as a *spherical* image. Each location on the sphere represents a different incident direction, and the intensity at that location is the amount of light arriving from that direction. We refer to these spherical images as ‘illumination maps’.

Paul Debevec (1998; Debevec, Hawkins, Tcho, Duiker, Sarokin and Sagar, 2000) has pioneered photographic techniques for acquiring illumination maps from locations in the real world. These illumination maps are extremely useful for computer graphics as they allow us to illuminate a synthetic object with patterns of light from the real world. The result is an image of the synthetic object that almost perfectly simulates how the object would have looked were it actually placed at that location in the world. We saw examples of such images in Figures 1 and 9.

Note that in an illumination map, all incident rays are treated equally. The illumination map is simply an image — like any other image — which contains raw intensity values at each location: it does not explicitly encode any information about the physical origin of the intensity. It makes no difference whether light arrives directly from a luminous source, or indirectly, reflected from surfaces in the surrounding scene, the result is simply a particular incident intensity value at the corresponding pixel. Thus, it does not make sense to think of an object as being illuminated by distinct light sources.

Furthermore, neighbouring pixels in the illumination map may originate from separate objects. The same light-dark transition in the illumination map could be caused by events that occur on a single surface, or by the boundary between two objects. As far as the structure of the illumination map is concerned, this is irrelevant.

This changes our conception of the world surrounding a specular object. We should not think of the scene that is reflected in a surface as a complex 3D arrangement of discrete physical bodies. Instead we should concentrate on the overall pattern generated by this complex scene. The illumination map can be thought of a sort of *texture* that is projected onto the surface of the object. This is important because it means that the visual system does not have to interpret the distorted outlines of particular environmental objects. Instead, it can interpret the way that the texture-like patterns are distorted by 3D shape. In the next section will use the concept of the illumination map to explain the relationship between 3D shape and image distortions.

Sampling the illumination map: how shape distorts reflections

Suppose we place a mirror inside an illumination map. The mirror projects points from the illumination map into the retinal image. Which point is projected into the image depends solely on the orientation of the mirror's surface, as shown in Figure 10. As the orientation of the mirror changes, different points in the illumination map are projected into the image.

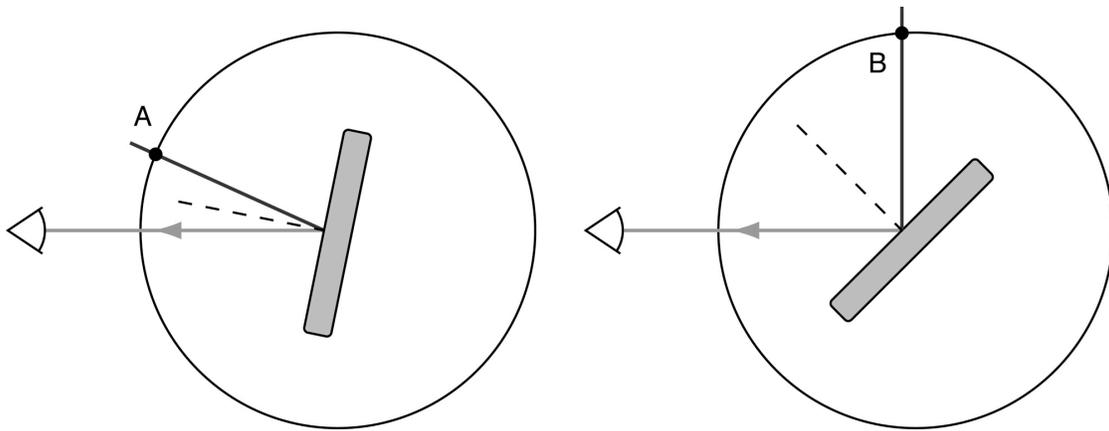


Figure 10. A mirror projects points from the illumination map into the image. Which point is seen in the image (A or B) depends solely on the orientation of the mirror.

This means that the image of an irregularly shaped 3D mirror *is* the image contained in the illumination map: it is just a distorted version of the image. Another way of putting this is that we can think of a 3D mirror as a ‘look-up-table’ for the contents of the illumination map. This look-up-table is ‘indexed’ by the object’s surface normal. As the surface normal varies continuously across the surface, the index sweeps smoothly through the illumination map, projecting different portions of the illumination map onto the retinal image.

The following observation is crucial to the subsequent arguments. Note that *the rate at which the surface samples the illumination map depends solely on the rate at which the surface normal varies across the surface*. This is depicted in Figure 11. When the normal varies slowly across the object (i.e. the surface is slightly curved), the surface sweeps through only a small portion of the illumination map. Consequently, the patterns that are reflected in the mirror will be stretched out over the surface. By contrast, when the surface normal varies rapidly across the object (i.e. the surface is highly curved), the

surface sweeps through a large swath of the illumination map. Thus, a large number of features from the illumination map are compressed into a small region of the retinal image.

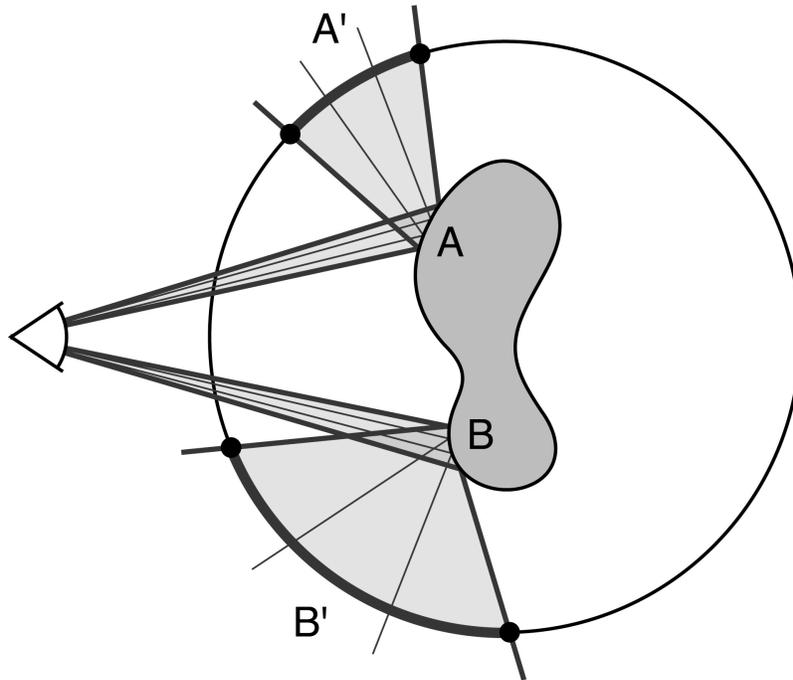


Figure 11. An irregular 3D mirror projects portions of the illumination map into the image. Around point **A**, the surface normal varies slowly. Thus, only a small portion of the illumination map, **A'** is projected into the image. By contrast, around point **B**, the surface normal varies quickly. This region of the mirror compresses a large portion of the illumination map, **B'** into the image.

In general, when the illumination map is reflected in an irregular 3D surface, it is effectively subjected to a complex coordinate transformation that stretches and compresses the patterns it contains. Crucially, however, the coordinate transformation depends directly on a property of the 3D shape: *the rate at which the surface normal varies across the surface*. This systematic relationship between surface curvature and image compression is basis of the approach to recovering shape from specular reflections that we elaborate in more detail throughout the rest of this chapter.

There are a couple of important points to note. The first is that the coordinate transform depends solely on the 3D geometry of the reflecting surface, it does not depend on the contents of the illumination map. In other words, the distortions are independent of the scene that is being reflected in the surface. This is useful because it means that if we can measure the local stretches and compressions of the image, we can recover the local surface curvatures without having to interpret specific features in the reflection. We are not interested in the contents of the reflection, simply how it is distorted.

This second important point is that stretching or compressing an image dramatically affects the local image statistics. In the next section will discuss this in more detail. The important point is that with minimal assumptions about the content of a pattern, it is possible to estimate how it has been stretched, just by looking at the local image statistics.

Effects of non-uniform scaling transformations on local image statistics.

As we have already mentioned, when the surface normal changes rapidly across the surface the reflection is compressed in the image. Conversely, when the surface normal changes slowly across the surface, the reflection is stretched out, by comparison. What effects do these stretches and compressions have on the local image statistics of the reflected pattern?

Suppose we take a natural texture of some pebbles, as shown in Figure 12(a), and compress it along one direction. The result is shown in Figure 12(b). We can see that the features of the texture are compressed into parallel streaks. The effects of this distortion

are visible in the Fourier spectrum of the pattern. Before the distortion, the pattern has an approximately uniform distribution of orientations. After the distortion, the pattern is distinctly anisotropic: power is concentrated along the direction orthogonal to the compression. In fact, any affine transformation applied to the image results in an equivalent affine transformation of the Fourier spectrum.

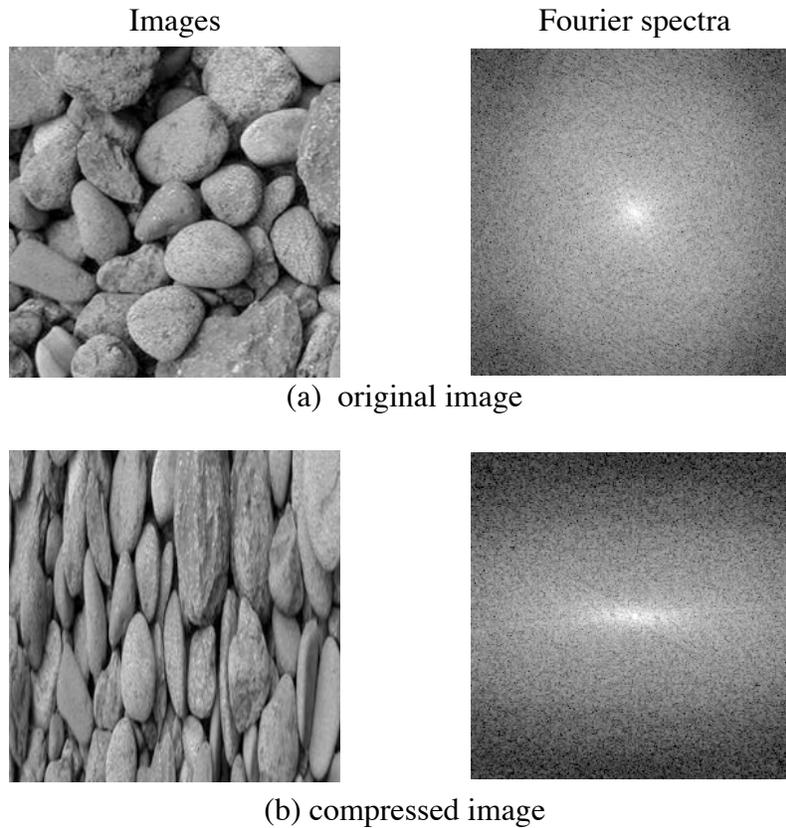


Figure 12. Effects of compression along one axis on the Fourier spectra of images. The original image (a) has an approximately isotropic spectrum. After compression (b), the Fourier spectrum is anisotropic. Note that this anisotropy is the same across scales.

This way of modifying the Fourier spectrum is quite different from standard image filtering. We are accustomed to thinking about the effects of simple filters on images, but distortions do not behave the same way. We will now discuss three critical differences.

First, note that distorting a Fourier plot shifts individual components through Fourier space. Thus, the transformed image can contain frequencies and orientations that were not present in the original. For example, consider the plaid shown in Figure x. After the stretching transformation, the components have shifted to different orientations, and higher spatial frequencies than in the original.

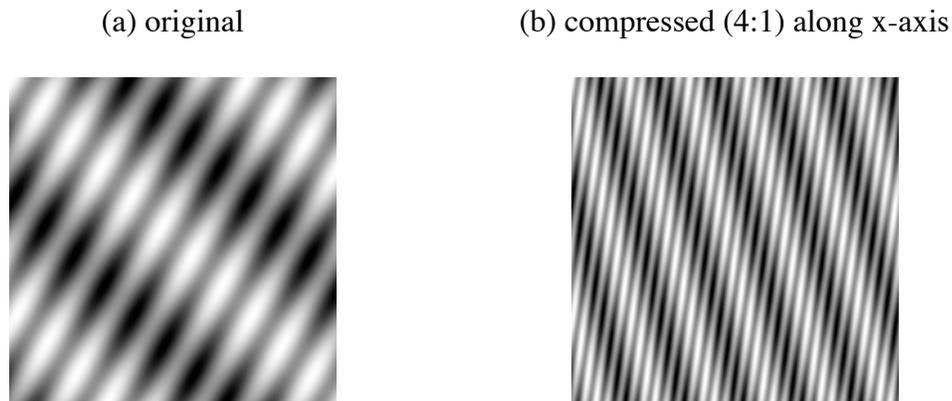


Figure 13. Compressing a plaid along the x-axis shifts the component gratings to different orientations and frequencies. Unlike filtering, the total power of the pattern remains constant.

Second, filtering an image necessarily reduces its power. This is not true of non-uniform scaling transformations. Compressing an image along one axis changes the orientation and frequency content without affecting the overall power of the image. Thus, after the transformation, even a highly anisotropic image patch will remain high contrast. This has the secondary effect that compressed images have unnaturally high power at high frequencies.

Third, filters are generally designed to operate at specific orientations and scales. For example, a filter may permit one orientation-band to pass at low spatial frequencies and a different orientation-band to pass at high frequencies. This is not possible with non-uniform scaling transformations. Compressions and stretches apply an affine deformation to the *entire* Fourier spectrum. Therefore, they are constrained to have

consistent effects on orientation across scales. This has the important consequence that stretched images tend exhibit similar anisotropies at all scales.

The fact that distorting transformations are more constrained than standard image filtering means that they have dramatic and easily detectable effects on the local statistics of images. This is very important for the recovery of shape from specular reflections. It means that the visual system can measure the way that a reflection is distorted directly from a local image patch, without having to segment the boundaries of reflected objects, and without having to locate the corresponding features in the surrounding environment. Instead, simple image measurements can provide a direct estimate of how the surface normal varies across the surface. For example, in Figure 14, we show how the Fourier spectra of local patches of a mirrored object are systematically related to properties of the object's 3D shape.

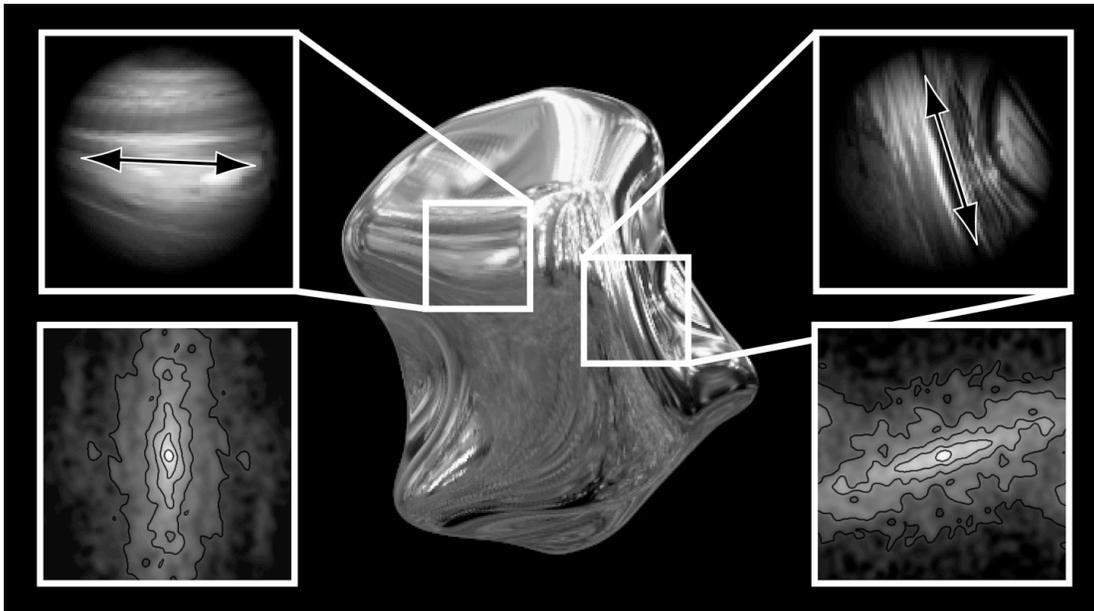


Figure 14. How the Fourier spectrum varies across the image of a mirrored object. The direction in which the surface normal changes most slowly is shown with the arrows. Note how the reflection is stretched into parallel streaks along these directions. This ‘streakiness’ has a clear effect on the Fourier spectra of the local patches.

Along directions of low curvature, the reflections are stretched out into parallel streaks. By contrast, along the orthogonal directions, the reflections are bunched up. The streakiness of the reflections is easily visible in the local Fourier spectra, which are pronouncedly anisotropic. This means that the visual system can use the local distribution of image orientations at each location in the image to infer properties of 3D shape. We discuss this idea in more detail in the next section.

Image orientations in the perception of 3D shape

We have argued that local image statistics are directly related to properties of 3D shape. In this section, we will explain how these measurements can be carried out with a population of simple oriented filters, and how shape properties are systematically related to the population response.

Imagine placing a spherical mirror in a world of random noise with a $1/f$ amplitude spectrum. This is shown in Figure 15(a).

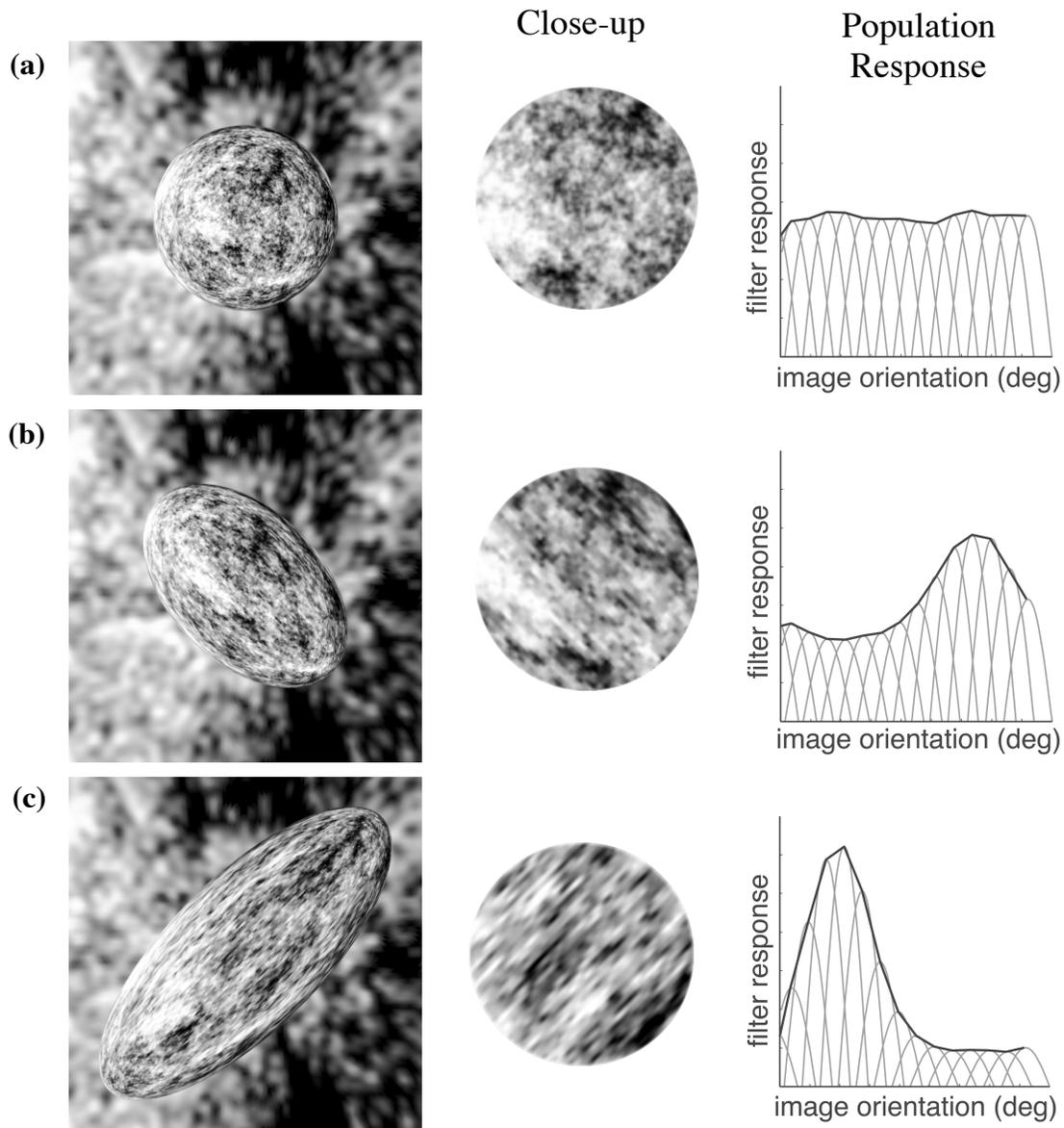


Figure 15. Mirrored surfaces in a world of $1/f$ noise, with responses of a population of oriented filters to the reflections. (a) A spherical mirror. (b) and (c) ellipsoid mirrors. Note that the population response exhibits a peak that is aligned with direction of minimum surface curvature. Peak size increases with surface anisotropy. [Figure reproduced from JOV article --- obtain copyright permission]

Now suppose we have a population of simple filters that are tuned to different orientations. How will this population respond to the image of the mirrored sphere? Note that at the centre of the image, the surface normal changes slowly in all directions. This means that the reflection of the random features is slightly miniaturized, but the

scaling does not have any preferential axis. Because the original noise pattern has a uniform distribution of orientations, the reflection (which is simply scaled uniformly) will also contain features at all orientations. Consequently, the filters in the population will respond roughly equally strongly, indicating that there is no preferential direction in which the surface normal changes more rapidly.

Now consider what happens if we elongate the sphere along one direction, to create an ellipsoid, as shown in Figure 15(b). Note that by doing this, we change rate at which the surface normal changes in different directions. Around the narrow axis of the ellipsoid, the surface normal changes rapidly, compressing a large number of features from the environment into a small portion of the image. Along the longer axis of the ellipsoid, the surface normal changes more slowly, sampling a smaller portion of the illumination map. Consequently, the reflections are elongated into parallel streaks that are aligned with the long axis of the ellipsoid. This is clearly visible in the response of the population of filters. Filters that are aligned with the streaks respond more strongly than the others, which results in a peak in the population response. Importantly, this peak response is directly related to a property of the 3D shape of the object: it is aligned with the direction in which the surface normal changes most slowly across the image.

We will now modify the 3D shape again. Suppose that we rotate the ellipsoid and elongate it further. This changes two relevant aspects of the 3D shape. First, the axis of least curvature is different: now, the surface normal changes most slowly in the perpendicular diagonal direction. Second, the surface is more anisotropic — the rate at which the surface normal changes along the long axes of the ellipsoid is even slower than previously, or, put another way, the ratio between the maximum and minimum curvatures

is greater. These two effects have visible consequences for the reflections. First, the reflected features are stretched out in a different direction. This means that the peak of the population response shifts to a different orientation. Second, the reflected features are more stretched than for the previous ellipsoid. This exaggerates the difference in response between the various filters: those filters that are aligned with the streaks respond more strongly than before, while those that are poorly aligned with the streaks respond less strongly before.

This is important as it means that a simple population of filters can make two key measurements that are directly related to the 3D shape of the surface. The peak of the population response indicates the direction of minimum surface curvature, and the size of the peak indicates the anisotropy of the surface. Together these provide two local constraints on 3D shape. In the Appendix, we express this basic intuition more formally.

Differences between intrinsic curvatures and second derivatives

Note that image orientations depend on the rate at which the surface normals vary across the image plane. We have used the term ‘curvature’ loosely to refer to the rate at which the surface normal changes across the image. It is important to clarify, however, that image distortion depends on the *second derivative of surface depths*, not on the *intrinsic curvature* of the surface. This is a subtle but very important distinction to make. We will now run through a quick example to make this distinction clearer.

Figure 16 shows a mirrored sphere. The intrinsic curvature of a sphere is the same at every point on its surface. This means that if you are standing on the surface itself and

walk around, the curvature under your feet will always be constant. However when we make an image of a sphere, we view it from a distant vantage point. The distance from this point to the surface varies across the image, creating a depth-map, such as the one shown in Figure 16. Here bright image values represent close points and dark values represent distant points. The iso-depth contours are intended to make the organization of the depth variations more salient.

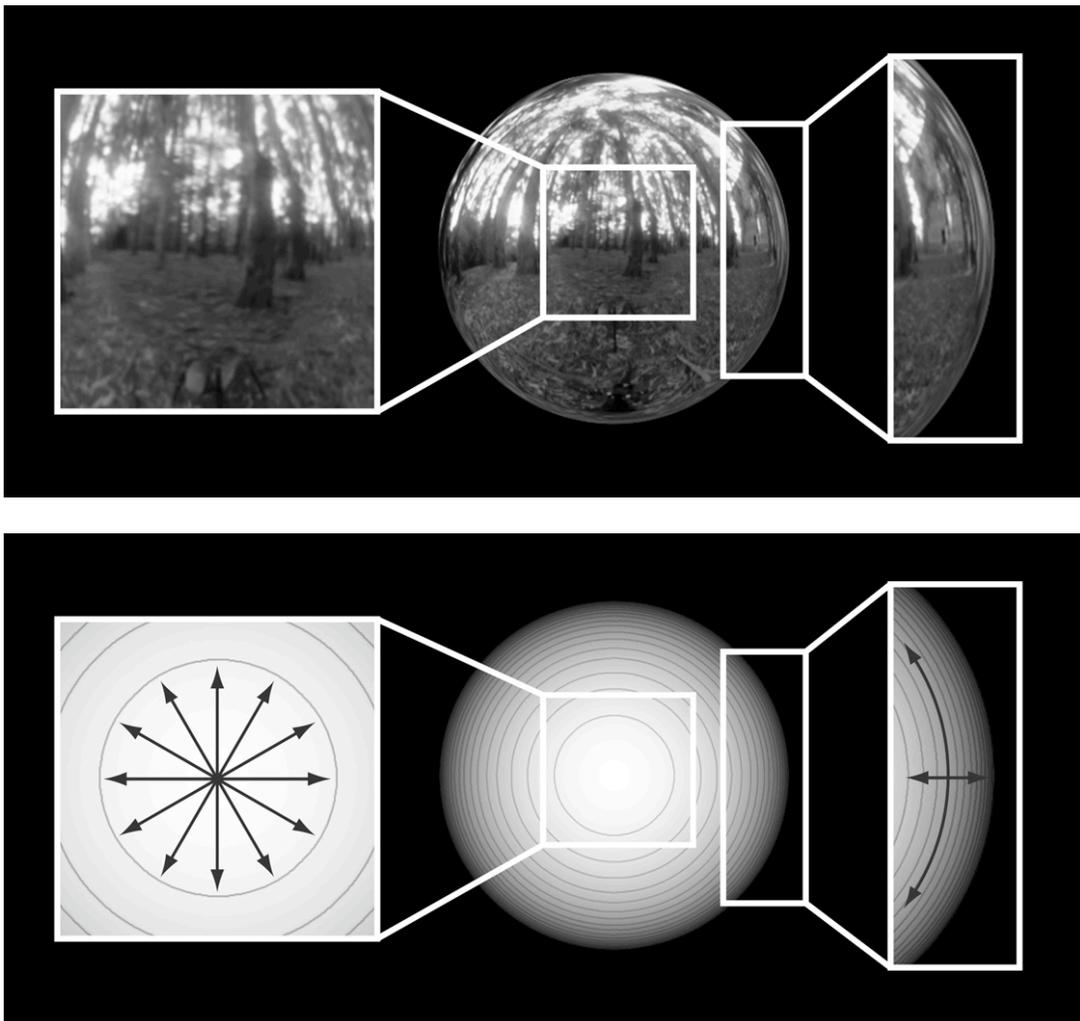


Figure 16. A specular sphere with its corresponding depth-map. In the center of the sphere, the second derivative of the depth-map is small and equal in all directions, correspondingly, the reflected image is not preferentially stretched in any direction. By contrast, at the edge of the sphere, the second derivative is large along the radial direction, but small along the circular arc parallel to the circumference. Accordingly, the reflected image is stretched along this direction.

Unlike intrinsic curvatures, the second derivatives of the depth map are not constant across the image. At the middle point of the sphere, the surface normal points directly at the viewer. If we move away from this point in any direction, the surface normal changes slowly. Thus, at the middle of the depth-map, the second derivative is small and constant in all directions. This means that the reflected image is somewhat miniaturized here, but not stretched in any given direction. This can be seen in the highlighted central region of the sphere.

By contrast, towards the edge of the depth-map, the sphere curves out of view. As we follow a radial line towards the circumference, the surface normal changes more and more rapidly. In other words, the second derivative is large in this direction. However, if we follow a circular path parallel to the circumference, the surface normal changes more slowly (in fact the slant is constant). This means along this direction the second derivative is small. The consequence of this is that the reflected image becomes bunched up as we approach the boundary of the sphere, and stretched out into concentric streaks in the orthogonal direction. This is depicted in the highlighted region at the edge of the sphere. While principle directions are ill-defined for the intrinsic curvature of a sphere, they are well-defined for second derivatives, because at all locations except the very center of the sphere, the maximum and minimum second derivatives are different from one another.

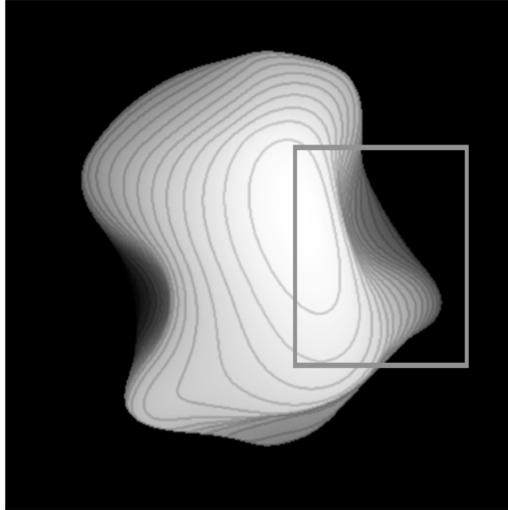
Note that the second derivative has some correspondence with our perceptual sense of curvature. We seem to have a clear visual impression that the edge of a sphere curves out of view more rapidly than the center of the sphere. This is because second

derivatives concatenate information about slant and intrinsic curvature. A slightly curved surface that is highly slanted with respect to the viewer has a large second derivative².

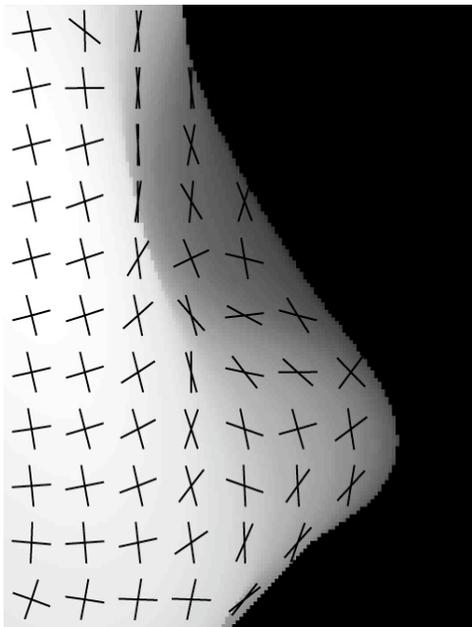
One final distinction is worthy of note. Consider the depth-map of an irregularly shaped object, such as the one shown in Figure 17. Euler (1760/1767) proved that directions of maximum and minimum curvature are orthogonal to one another at every location on such an object. For intrinsic curvatures, the principal directions are orthogonal to one another *with respect to the intrinsic coordinates of the surface*. When projected into the image plane, these are generally no longer perpendicular (because of foreshortening). This can be seen in the detail shown in Figure 17(b). By contrast, second derivatives are defined with respect to the image plane. Thus the directions of maximum and minimum curvature are always orthogonal to one another *in the image plane*. This can be seen in the detail shown in Figure 17(c).

² This has one very interesting consequence. For closed, globally convex objects, the first and second derivatives tend to be partially correlated. This means that textured and reflecting surfaces often tend to generate similar orientation fields in highly slanted regions.

(a) Depth-map



(b) Intrinsic surface curvature



(c) Second derivative

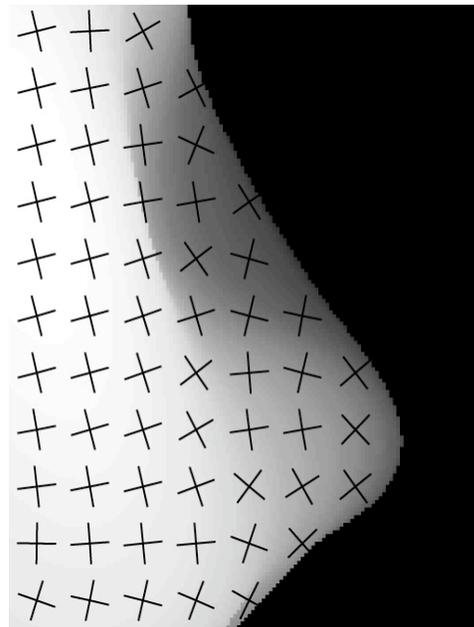


Figure 17. Principal directions of curvature are different when measured on the surface or on the projection into the image plane. (a) The depth-map of an irregularly shaped object. Grey boundary highlights region shown in (b) and (c). (b) The principal curvature directions as measured on the surface itself. Note that the directions are not orthogonal in the image. (c) The principal curvature directions derived from the second derivatives of the depth-map. Note that the directions are orthogonal in the image.

From local to global: orientation fields and the perception of 3D shape

We have outlined how simple image measurements can be informative about local 3D shape properties. However, these measurements become much more powerful when applied globally. The streaky patterns in a small patch of surface tell us relatively little about 3D shape when presented in isolation, as can be seen in Figure 18. The important information arises from the way that reflections twist and writhe across the entire surface of an object. For complex, irregular shapes, the second derivatives of the surface change continuously from location to location. This means that the distribution of orientations in the image also changes continuously, creating distinctive swirling patterns, or *orientation fields*, which seem to carry most of the important information about 3D shape.

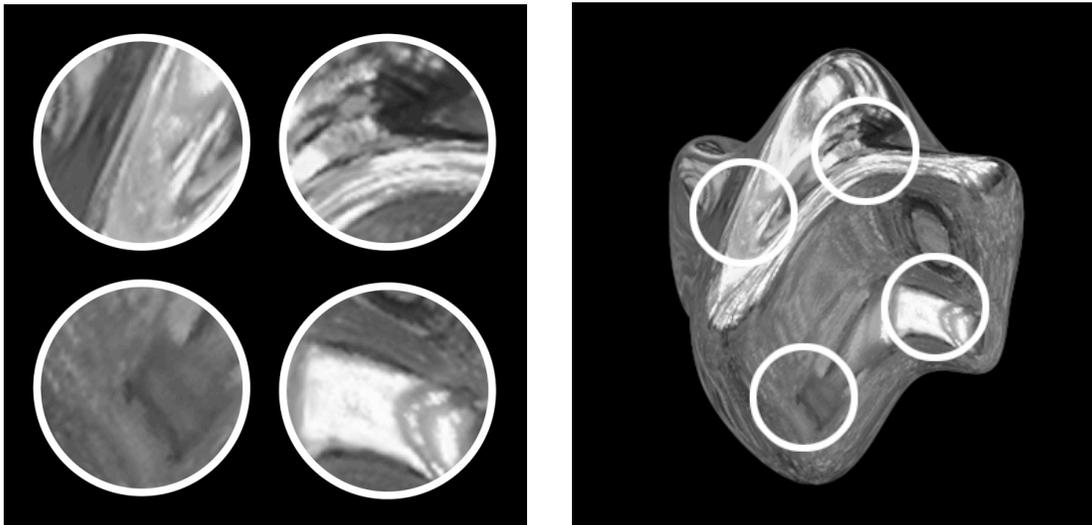


Figure 18. Local patches yield weak percepts of 3D shape. In the right panel the patches are shown in their original context, where they contribute to a strong overall sense of 3D shape.

One of the most interesting properties of orientation fields is that they remain surprisingly stable, irrespective of the patterns that are reflected in the surface. This is shown in Figure 19.

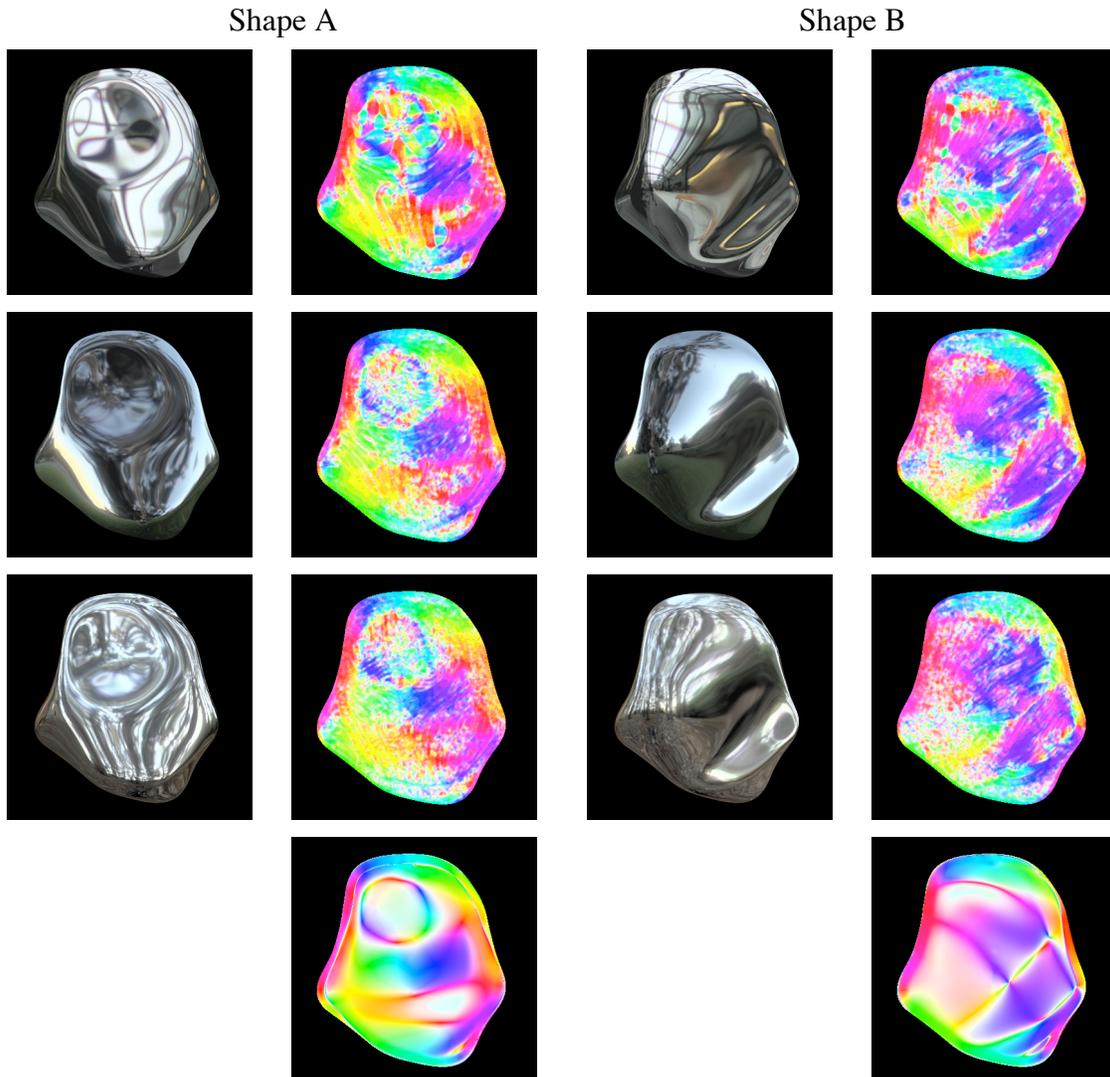


Figure 19. Shapes A and B are shown reflecting three different environments. Next to each rendered image is a colour visualization of the orientation field: hue denotes dominant image orientation, colour saturation denotes local image anisotropy. The orientation fields remain surprisingly stable across changes in the reflected world. At the bottom is shown the ‘ground-truth’ directions of minimum second derivative and surface anisotropy, which are measured directly from the 3D shape model that was used to render the images. Note the striking similarity between the image-based orientation fields and the ground truth measurements. [reproduced from JOV article --- obtain permission]

Two shapes (A and B) are shown rendered in three different environments. In different scenes, the images appear quite different from one another — for example,

locations that are dark in one scene may be bright in another scene, depending on what happens to be reflected in the surface of the object. Despite these variations, the orientation content of the image remains quite stable. The orientation fields for each image are shown using a colour visualization. Hue denotes the dominant image orientation, so, for example, red means that the image is locally dominated by vertical structures. Note that the hue changes continuously across the orientation field, indicating the local orientation of the image modulates smoothly across the image. Colour saturation indicates the local anisotropy of the image. Where the image is highly anisotropic (i.e one orientation dominates the others), the colour is highly saturated. By contrast, where the image is weakly oriented, the colour washes out to neutral white. Note that the overall pattern of the orientation fields stays quite stable across the different scenes, but is distinctly different for shape A and shape B.

At the bottom of each column in Figure 19, is shown an additional orientation field. This one represent the ‘ground truth’ data derived directly from the 3D shape model that was used to render the images. This time, hue represents the actual direction of minimum second derivative, and saturation represents actual surface anisotropy. There is quite a striking correspondence between the orientation fields derived from the rendered images and the ones derived directly from the shape model. It is almost as if orientation fields allow us to ‘look through’ the complex reflections in the image, to see the underlying properties of the 3D shape.

Why are orientation fields so stable? In the next section we discuss this question in more detail. The basic intuition is as follows. When a mirrored object is moved from scene to scene, the distribution of lights and darks can change radically because they

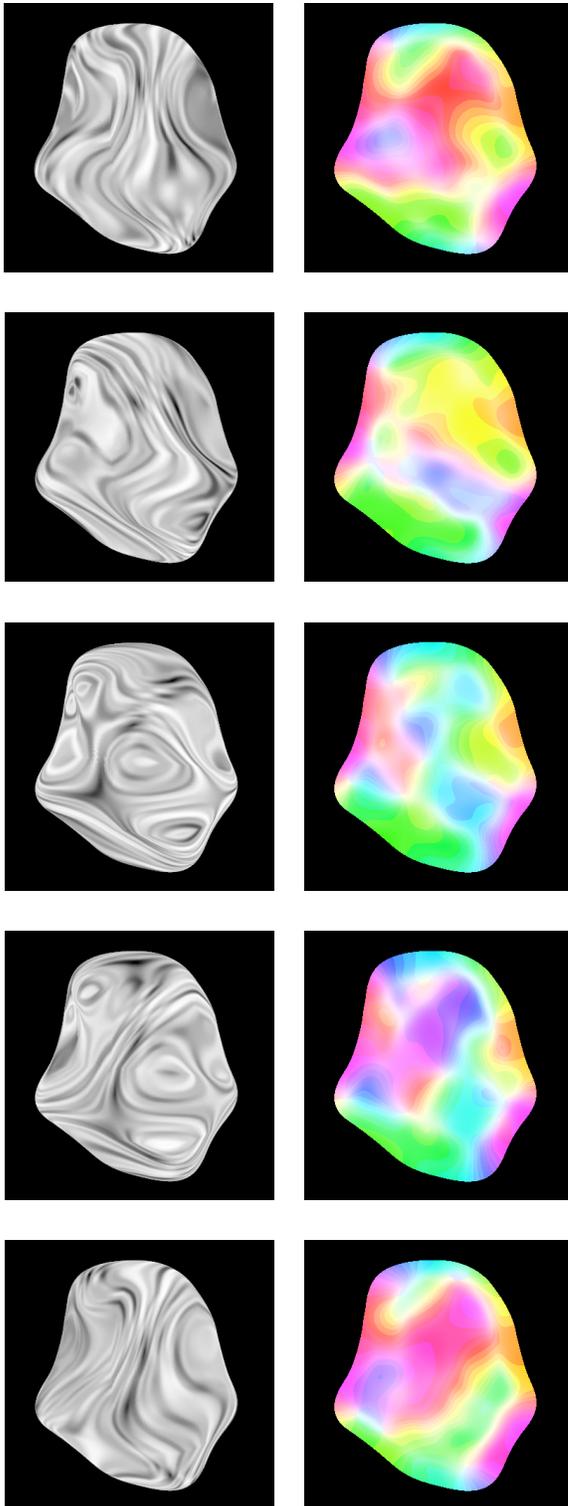
depend solely on the content of the illumination map. However, the *orientation content* of the image depends much more on the way the reflection is distorted. Image distortions depend on the 3D shape of the reflection object, not on the contents of the illumination map. Consequently, orientation fields remain more stable as the object is moved from scene to scene.

Anisotropies in the illumination map and their effects on orientation fields.

How important are the statistics of the reflected scene for the orientation fields that are generated by an object? As we have already argued, if the illumination map is roughly isotropic (*i.e.* only weakly oriented), then the orientations in the image result solely from the way that the reflection is distorted. However, what happens if the illumination map is highly anisotropic? Surely, this will create misleading patterns of orientation in the image?

In Figure 20, we show a mirrored object reflecting a series of illumination maps that consist of parallel stripes. These patterns are radically more anisotropic than we would normally encounter in the real world. In each consecutive frame, the stripes in the illumination map are aligned with a different orientation

(a) Striped illuminations



(b) Real-world illumination

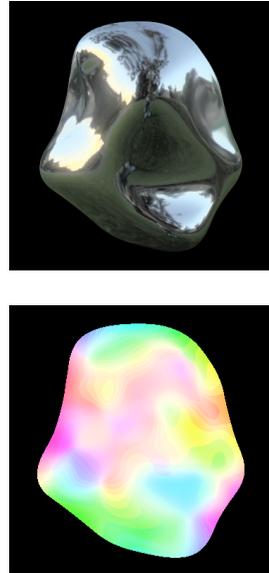


Figure 20. Mirrored objects rendered under stripy illuminations. In (a), the shapes remain constant, all that changes is the orientation of the stripes in the illumination map. As the stripes rotate, different features of the object are accentuated. Also shown are the orientation maps for each image. Hue is dominant orientation, saturation is anisotropy. For comparison (b) shows the same object under real-world illumination.

The extreme anisotropy of the illumination maps has a marked effect on our impression of 3D shape. Most observers agree that when the object is rendered under the striped illuminations (Figure 20a), the sense of 3D shape is generally poorer than under the real-world illumination (Figure 20b).

There are a couple of important points to note. First, observe that as the stripes rotate through different orientations, they selectively accentuate different features of the 3D shape. For example, when the illumination consists of vertical stripes (top panel), the horizontal shape features are de-emphasized. Intuitively, this occurs whenever the reflection is compressed along the axis of the stripes because no amount of compression along this direction can alter the orientation of the stripes. Consequently, the underlying shape leaves no measurable trace in the orientations in the image for the visual system to detect. Thus, when the stripes are orthogonal to the direction of minimum second derivative, the orientation information in the image becomes weak and inconsistent, whereas, when they are aligned with the direction of minimum curvature, they accentuate the feature.

Second, note that although the illumination map contains only a single orientation, the reflections nevertheless contain continuously varying orientations. The distortions introduced by shape change the orientations of the stripes at most locations in the image. Even though the orientation maps are less reliable than under real-world illumination, the orientations are nevertheless *biased* towards the correct directions. Indeed, it is interesting to note that the images whose apparent shapes most closely resemble (b), are those whose orientation fields most closely resemble the orientation field of (b).

Furthermore, within an image, the features of the shape that are most compelling are generally those whose orientation fields are most similar to those of (b).

The fact that even highly anisotropic illumination maps yield orientation fields that are biased toward the correct directions, is the reason why orientation fields are so stable in the real world. This can be examined more systematically by considering the relationship between orientations in the illumination map, surface anisotropy and the resulting orientations in the image. As discussed in the Appendix, there is a lawful relationship between these factors, which we will now explore in more detail.

In general the local orientation in the illumination map is at some arbitrary angle relative to the direction of minimum second derivative at the corresponding location on the surface. We will call this angle θ_R . When θ_R is 0deg, the illumination map happens, by chance, to contain a structure that is aligned with the direction of minimum second derivative. When θ_R is 90deg the structure in the illumination map is orthogonal to the direction of minimum second derivative on the surface.

The mirrored surface subjects the orientation of the illumination map to a transformation as it projects it into the image. The result is an *image* orientation that is at some new angle, θ , relative to the actual direction of minimum second derivative. θ is the error in estimating the direction of minimum second derivative based on the orientation observed in the image. When θ is small, the image orientation closely approximates the direction of minimum second derivative. When θ is large, the image orientation is a poor estimate of the direction of minimum second derivative.

When the surface is slightly anisotropic, it has a weak effect on the orientation, and thus θ will be close to θ_R . However, when the surface is highly anisotropic, it warps

the projected orientation towards the direction of minimum surface curvature, creating a small value of θ . The function relating theta to surface anisotropy is shown in Figure 21. The different lines show how theta is affected by surface anisotropy (α) for different values of θ_R .

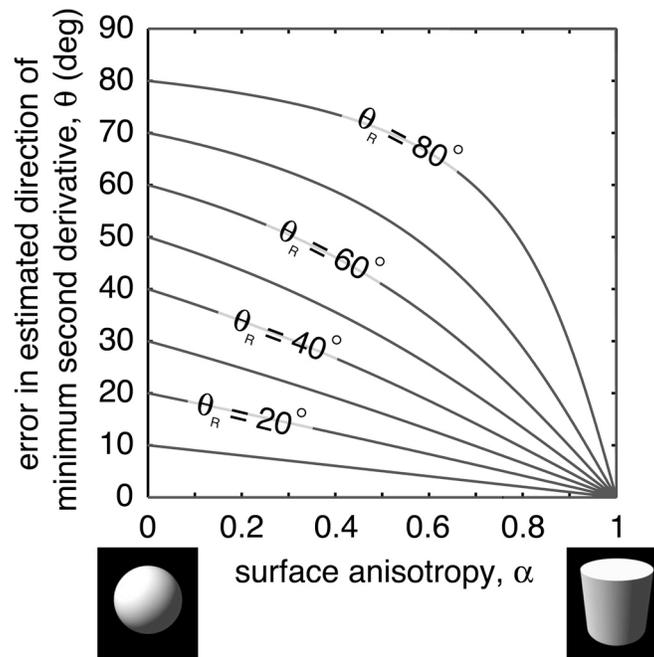


Figure x. Error in estimated direction of minimum second derivative as a function of surface anisotropy for different values of θ_R . Note that when the surface is highly anisotropic, estimation error tends to zero, irrespective of θ_R .

Note that when the surface is approximately isotropic (*i.e.* small values of anisotropy), the direction of minimum second derivative is poorly-defined, and theta tends to θ_R , as expected. By contrast, when the surface is perfectly anisotropic (*i.e.* the value of the minimum second derivative is zero), theta tends to zero, irrespective of the angle θ_R . For the special case when $\theta_R = 90^\circ$, theta is a step-function, which is 90° for all values of alpha, except 1, when theta plummets to zero. This is the case discussed

above in which the stripes in the illumination map are perfectly orthogonal to the direction of minimum second derivative. In the real world, surface curvatures are likely to be poorly correlated with the orientations of structures in the surrounding scene. Thus, on average, θ_r will be close to 45deg.

We can re-plot this family of functions as a two-dimensional ‘*generative space*’, as shown in Figure 22. Here, theta is plotted as a function of surface anisotropy and θ_r . This space covers all possible generative configurations; that is, all combinations of surface curvatures and orientations of the illumination map. For example, in the bottom right hand corner, we have occasions when the surface is highly anisotropic, and, by chance, the illumination map is well aligned with the direction of minimum second derivative. Naturally, this combination leads to image orientations that are very closely aligned with the direction of minimum second derivative (i.e. low θ). By contrast, in the top left of the generative space, the surface is only slightly anisotropic and, by bad luck, the illumination map contains structures that are almost orthogonal to the direction of minimum second derivatives. Under these circumstances, θ is large.

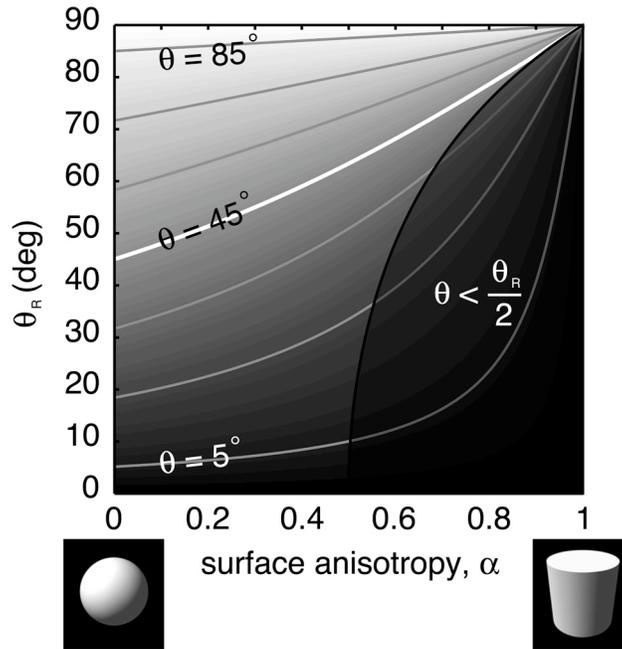


Figure 22. The generative space. θ is plotted as a function of θ_r and surface anisotropy. This space includes all possible combinations of illumination orientation and surface curvatures. Grey contours show equal values of θ . White contour shows $\theta = 45\text{deg}$. Note that most of the space is below this line. Dark area shows region of space for which $\theta < \theta_r/2$.

A number of points are worthy of comment. First, note that the majority of the space is relatively dark (*i.e.* θ is small). This means that for most possible scene configurations, the estimated direction of minimum second derivative is quite accurate. Second, nearly three-quarters of the generative space yields θ values that are less than 45deg (the area below the white contour). So, for most possible scene configurations, the image orientation is closer to the correct direction than to the orthogonal direction. Third, at every point in the space θ is smaller than θ_r , which means that the image orientation is *always* biased towards the correct direction. And finally, for nearly a third of the space—the darkened region—the image orientation is closer to the direction of

minimum second derivative than it is to the orientation in the illumination map. In other words, for a considerable portion of possible scene configurations, image orientation depends more on 3D shape than on the scene reflected in the surface.

In Figure 23, we summarize how much of the generative space yields reliable estimates of the direction of minimum second derivative. Note that the curve is positively bowed. This again shows that most of the generative space produces small values of theta. In particular, note that half of all possible scene configurations yield image orientations that are within about 23deg of the actual direction of minimum surface curvature.

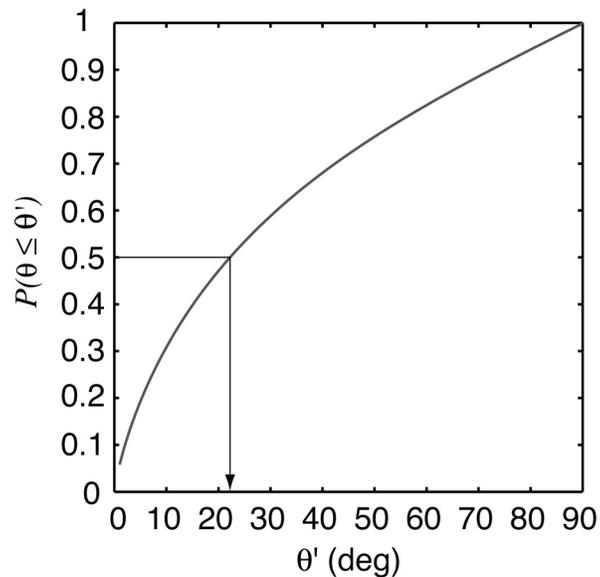


Figure 23. Cumulative proportion of the generative space that yields a given value of θ . Note that the curve is positively bowed. Thus, for example, half of the generative space yields theta values that are less than about 23deg.

This systematic relationship between the orientations in the illumination map, surface anisotropy and the orientations in the image reveals why orientation fields are generally so reliable and accurate in the real world. Most *possible* worlds yield quite reliable orientation fields, and completely random illumination maps yield the least

biased orientation fields of all. Thus, the visual system may not have to explicitly estimate the specific statistics of our natural environment in order to recover shape from specular reflections. Instead, it can implicitly rely on the geometry of specular reflection to warp the world into alignment with the curvatures of the 3D shape.

Orientations fields as a common currency for 3D shape perception.

The orientation fields that are generated by specular surfaces are simple to extract from the image, and provide reliable and accurate constraints on 3D shape. However, most surfaces in the world are not *purely* specular. Although glossiness is common, so is diffuse shading and patterns of surface texture. As Figure 24 shows, specular, diffuse and textured versions of an object can all lead to a compelling percept of its 3D shape. How do we recover 3D shape from these different optical properties of objects?

As discussed at the beginning of this chapter, there are many respects in which shading, highlights and texture behave differently from one another. Diffuse reflectance leads to smooth patterns of shading, while both textures and specularities lead to dense, complex patterns across the shape. Texture patterns depend solely on the way that the surface is pigmented, while both specularities and shading patterns depend on the position of light sources and other objects in the environment. Specularities move around when the observer changes position, while both shading and texture remain effectively attached to the surface irrespective of the viewpoint.

Given that shading, highlights and texture behave differently from an optical point of view, it might seem likely that the visual system has to apply fundamentally different

strategies to estimate 3D shape from these cues. Shape-from-shading has traditionally been thought to be quite separate from shape-from-texture, and those few authors who discuss specular reflections as a source of information about 3D shape (e.g. Norman *et al.*, 2004; Oren and Nayer, 1996, Savarese Li and Perona, 2003) generally treat them as a complimentary source of information, which is quite independent from classical shape-from-shading.

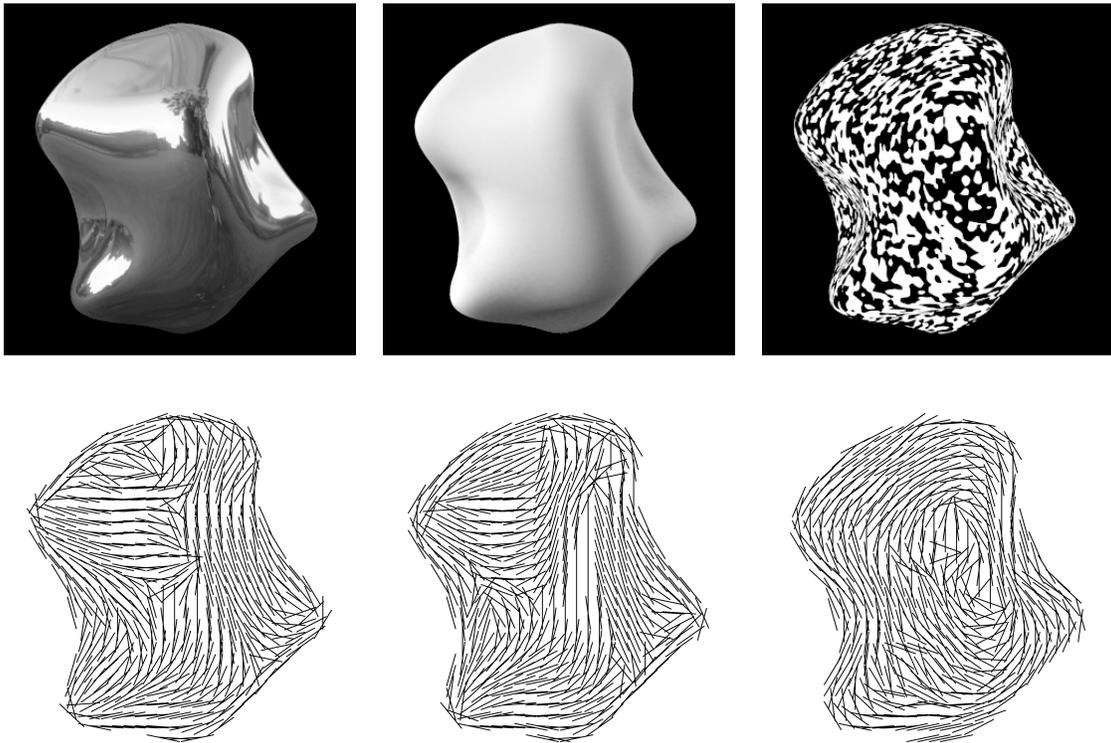


Figure 24. Three renderings of an object: mirrored, diffuse and textured. The three images look very different, and have traditionally thought to provide the visual system with distinct visual cues to 3D shape. However, in all three cases, the orientation content of the image (bottom row) seems to carry important information that could be used for 3D shape estimation.

However, if we look at the orientation content of the images, we see that these seemingly different cues may have more in common than previously thought. Figure 24 (bottom row) shows the dominant image orientation at each location in the image of specular, diffuse and textured versions of an object. The striking observation is that in all

three cases, the images contain smooth, continuously varying patterns of image orientation that seem to vary in a way that is directly related to the 3D shape of the object.

This is particularly surprising in the case of the diffuse object: the image is so smooth, that we do not readily perceive the fact that it contains continuously varying orientations. Nevertheless, simple filters are capable of detecting this latent orientation structure. Note, indeed, that the orientation field is very similar to the one produced by the specular surface. This is because the same basic logic is responsible for the orientation structure of the image. The intensity of the surface depends on the direction in which it is pointing. Thus, when the surface normal changes slowly, the image intensity will tend to vary slowly. By contrast, when the surface normal changes rapidly, the image intensity will tend to change rapidly too. This leads to a systematic relationship between the second derivatives of the surface and the orientations in the image. Although orientation fields are less stable for diffuse surfaces than for specular surfaces, the fact that shaded and textured objects produce orientation fields that are related to 3D shape suggests that the visual system could use some of the same processing strategies for these seemingly disparate cues.

A few researchers have previously argued that the orientation content of shaded images could be used to recover 3D shape (*e.g.*, Ben-Shahar and Zucker, 2001; Breton and Zucker, 1996; Huggins, Chen, Belhumeur, and Zucker, 2001), and a number of algorithms exist for estimating shape-from-texture using simple image statistics (Malik and Rosenholtz, 1997; Super and Bovik 1991, 1995). Thus, we suggest that orientation

fields could act as the basic ‘common currency’ that the visual system uses for recovering 3D shape from a range of different cues.

Conclusions

In this chapter, we have laid out some theoretical observations on the recovery of 3D shape from specular reflections. We started by reviewing some of the eccentric properties of specular reflections that make them potentially problematic for the visual system. Specular reflections can make a surface very bright or very dark, depending on what happens to be facing the object. Indeed, a specular object could in principle be made to take on any arbitrary appearance by carefully modifying the world that is reflected in its surface. This makes specular reflections unstable and ambiguous, and yet somehow, the visual system is capable of recovering the 3D shape of an object from the distorted reflections in its surface.

We outlined one approach to solving this problem based on identifying recognizable objects in the reflection and estimating the way that these reflections are distorted. However, we argued that the human visual system probably does not go about the problem in this way. Instead, we suggested that the visual system can estimate the way the reflection is distorted directly from its local image statistics.

We showed that the rate at which the surface normal changes across the surface determines the way the reflection is distorted. Specifically, when the surface normal changes rapidly, the reflection is compressed, and when the surface normal changes slowly, the reflection is stretched out by comparison. Put another way, surface curvatures

apply non-uniform scaling transformations to the reflected image. These transformations have dramatic effects on the local statistics of the reflection because they introduce reliable anisotropies across scales.

We showed that populations of simple oriented filters are capable of measuring these image properties and can thereby provide the visual system with estimates of the direction of minimum second derivative and surface anisotropy. If you look at the output of these filters at every point in the image, you will find smoothly varying patterns, which we call *orientation fields*. Orientation fields provide accurate and surprisingly stable constraints on 3D shape, which suggests that it is possible to recover the shape of a mirrored surface without knowing the organization of the scene surrounding the object.

Orientation fields are stable in the real world because the geometry of specular reflection always warps orientations in the world closer to the direction of minimum second derivative. We analyzed the *generative space* that describes all possible combinations of surface curvatures and orientations in the world, and showed that in most possible worlds, image orientation provides a reliable indication of the direction of minimum curvature.

Finally, we argued that image orientations could play a more general role in the perception of 3D shape. It is well-known that early stages of visual processing involve extraction of image orientations. Here we have argued that these measurements are not useful solely for filtering the image into orientation bands; rather, they could, in principle, be put directly to good use in the estimation of 3D shape.

Appendix: *Relationship between image derivatives and second derivatives of surface height.*

We will now elaborate more formally the basic intuition that describe in the main body of the text. To anticipate, we will show how the local orientation structure of the image is closely related to the rate at which the surface normal changes across the surface. Consequently, the visual system can use simple orientation-detection mechanisms to recover some important constraints on 3D shape.

The starting point is the process of rendering an image from an object, as depicted in Figure 25.

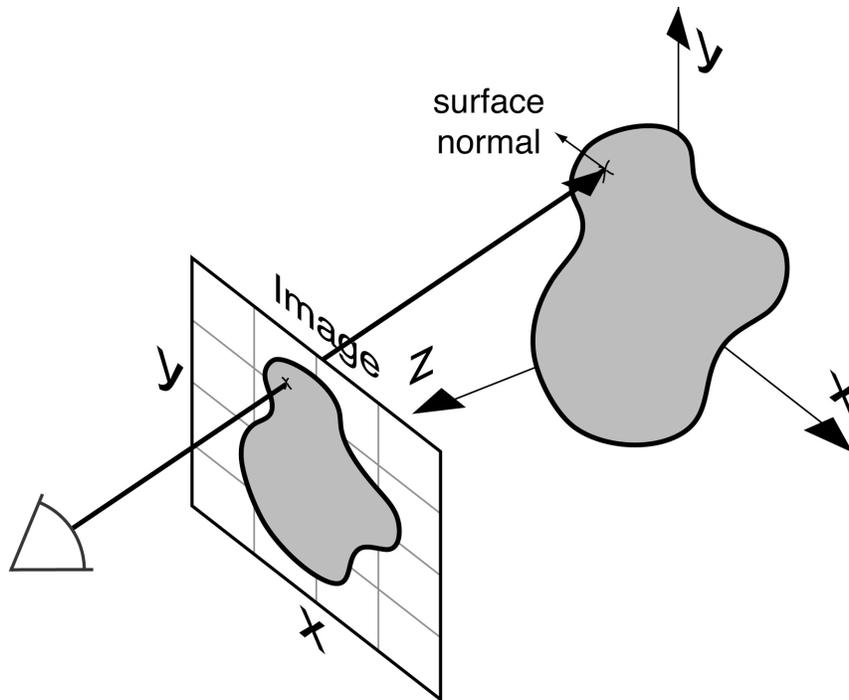


Figure 25. Rendering an image of an object. Rays are cast through an image plane towards a 3D surface. Image intensity, $I(x, y)$ is determined by the surface normal at the corresponding surface location $S(x, y, z)$. Under orthographic projection the observer is infinitely far away and the image is simply projection along the z -axis.

Each point on the object's surface has a unique location, $S(x, y, z)$ in 3D space. In order to create an image of the object, rays are cast through an image plane until they strike the surface. The intensity at a given point in the image, $I(x, y)$ is simply the intensity of the surface at the corresponding point $S(x, y, z)$. Note that for a distant viewer, the image is effectively a projection of $S(x, y, z)$ along the z -dimension. Thus, we can represent the shape of the object, as seen from the image plane, as a 2D function: a map of how the z -dimension varies across the image, which we will call $Z(x, y)$.

What determines $I(x, y)$? Suppose we want to render an image $I(x, y)$ of a mirrored surface whose shape (surface height) across the image is $Z(x, y)$. The intensity of a point on the surface depends solely on the direction that it is pointing. Thus, the image irradiance equation is:

$$I(x, y) = R(p(x,y), q(x,y))$$

Where R is the reflectance map, which is indexed by the surface derivatives $p = Z_x$ and $q = Z_y$. For convenience, we will use the notation $Z_x = \partial Z / \partial x$ and $Z_y = \partial Z / \partial y$. The function R is unknown and can be quite complicated. Our goal is to find properties of the image that reflect properties of the shape independent of the reflectance function.

The gradient of the image depends on the gradient of the intensities in illumination map, and on the rate at which the surface sweeps through the illumination map (i.e. on the rate at which the surface normal changes). Thus:

$$I_x = R_p \cdot Z_{xx} + R_q \cdot Z_{yx}$$

$$I_y = R_p \cdot Z_{xy} + R_q \cdot Z_{yy}$$

Using matrix formulation and factoring out the contributions of shape and the illumination map to the image derivatives, we can write:

$$\nabla I = \begin{bmatrix} Z_{xx} & Z_{yx} \\ Z_{xy} & Z_{yy} \end{bmatrix} \begin{bmatrix} Rp \\ Rq \end{bmatrix} = H \nabla R$$

Where H is the Hessian matrix containing the second order derivatives of the surface $Z(x, y)$. This is the basic equation that states the dependency of the image derivatives on both the rate of change of the surface normal and on the derivatives of the illumination map.

Importantly, we can re-express the surface in terms of its principal curvatures by decomposing H into eigenvectors:

$$H = k_1 (e_1 e_1^T) + k_2 (e_2 e_2^T)$$

Where e_1 and e_2 are eigenvectors (column vectors) and k_1 and k_2 are the corresponding eigenvalues, *i.e.* the magnitudes of the principal curvatures. By convention $k_1 < k_2$.

Thus, we can re-express the image gradient as:

$$\nabla I = k_1 (e_1^T \nabla R) e_1 + k_2 (e_2^T \nabla R) e_2$$

$$\nabla I = k_1' e_1 + k_2' e_2$$

Where $k_1' = k_1 (e_1^T \nabla R)$ and $k_2' = k_2 (e_2^T \nabla R)$. Note that when the surface has one direction with null curvature, $k_1 = 0$, then the image gradient is parallel to e_2 : $\nabla I = k_2' e_2$ irrespective of ∇R . Now, let's see how the orientation of the reflectance map (∇R) affects the correlation between the image orientation and the direction of minimum curvature. If θ_R is the angle between ∇R and the direction of maximum curvature, e_1 , then the error angle between ∇I and e_2 will be:

$$\cos(\theta) = \frac{1}{\sqrt{1 + \frac{k_1^2}{k_2^2} \tan^2(\theta_R)}}$$

This is the equation used for constructing the ‘generative space’ in the section entitled *Anisotropies in the illumination map and their effects on orientation fields*.

Surface anisotropy is defined as:

$$\alpha = 1 - \frac{k_1}{k_2}$$

In the main text, we argue that local anisotropy of the image can be estimated using a population of filters tuned to different orientations. A more stable measurement of image orientation and estimate of the local image anisotropy can be obtained using the structure tensor (Knutsson, 1989). The structure tensor is a matrix obtained as:

$$J_i(x, y) = (\nabla I \nabla I^T) * K_\rho = \begin{bmatrix} (Ix)^2 & IxIy \\ IxIy & (Iy)^2 \end{bmatrix} * K_\rho$$

Where $*$ is the convolution operation and K_ρ is a spatial gaussian of variance ρ^2 . $J_i(x, y)$ is a 2x2 matrix at each image location. Without spatial smoothing, the matrix J_i is rank one. After smoothing by integrating over a small neighborhood, the eigenvalues and eigenvectors describe the local image structure. The matrix J_i can be decomposed as $J_i = \mu_1 w_1 w_1^T + \mu_2 w_2 w_2^T$, where w_1 and w_2 are the eigenvectors of J_i , and the eigenvalues are $\mu_1 \geq \mu_2$. The streakiness (anisotropy) of the image is captured by the eigenvalues (μ_1 and μ_2) of J_i . Image regions with linear structure (highly streaky) have one vanishing eigenvalue. In an isotropic neighborhood both eigenvalues are equal. If $\mu_1 > \mu_2$ then the image has oriented structure.

From the rendering equation, the image structure is related to the structure of the reflectance map:

$$J_i = (H \nabla R \nabla R^T H) * K_\rho$$

If the second order derivatives of the shape change slowly we can approximate the structure tensor by treating the matrix H as constant in the region of support of the gaussian kernel K_ρ (this approximation will be correct if we can approximate locally the surface Z by a second order polynomial):

$$J_i \approx H J_r H$$

This equation describes how the image structure (orientation and anisotropy) relates to the second order derivatives of the surface and to the local structure of the reflectance map.

As discussed in the main text, there are some cases in which the image structure is oriented exactly along the principal direction of the surface curvature. If the surface has one direction with null curvature (e.g. a point on a cylinder), then eigenvectors of J_i are also the eigenvectors of H and this is independent of the illumination. If J_r is the identity (this approximately obtained for a specular object reflecting a $1/f$ noise illumination map), then the eigenvectors of J_i are the eigenvectors of H : image flow lines are parallel to direction of minimum curvature. If J_r is the identity, then $\mu_1 / \mu_2 = (k_1 / k_2)^2$. This gives an estimation of which image regions might have a large difference between image orientation and the directions of principal curvature of the shape. The smaller the ratio μ_1 / μ_2 the closer is the image orientation to the direction parallel to the direction of maximal surface curvature.

Acknowledgements:

This work was funded by Thanks also to Ed Sandifer of the Euler Archive for digging up the correct reference to Euler's seminal proof that the principal curvature directions are always orthogonal.

References

- Ben-Shahar, O., & Zucker, S. (2001). On the perceptual organization of texture and shading flows: From a geometrical model to coherence computation. In *Proceedings of CVPR*, pp. 1048—1055.
- Blake, A. & Bülthoff, H.H. (1990) Does the brain know the physics of specular reflection? *Nature*, **343**: 165—168.
- Blake, A. & Bülthoff, H.H. (1991). Shape from Specularities: Computation and Psychophysics. *Philosophical Transactions of the Royal Society Series B* **331**: 237—252.
- Breton, P. & Zucker, S. W. (1996). Shadows and shading flow fields. In *Computer Vision and Pattern Recognition*, 782—789.
- Debevec, P. E. (1998). Rendering synthetic objects into real scenes: Bridging traditional and image-based graphics with global illumination and high dynamic range photography. *Computer Graphics (SIGGRAPH)*.
- Debevec, P. E., Hawkins, T., Tchou, C., Duiker, H.-P., Sarokin, W. and Sagar, M. (2000). Acquiring the reflectance field of a human face. *Computer Graphics (SIGGRAPH)*.

- Euler, L. (1760/1767). "Recherches sur la courbure des surfaces", (Enestrøm number 333) *Mémoires de l'académie des sciences de Berlin* [16] 1767, p. 119-143, reprinted in the *Opera Omnia*, Series I, vol. 23 p. 1-22.
- Fleming, R. W., Torralba, A., & Adelson, E. H. (2004). Specular reflections and the perception of shape. *Journal of Vision*, 4(9), 798-820, <http://journalofvision.org/4/9/10/>, doi:10.1167/4.9.10
- Fleming, R. W., Torralba, A., Dror, R. O., & Adelson, E. H. (2003). How image statistics drive shape-from-texture and shape-from-specularity [Abstract]. *Journal of Vision*, 3(9), 73a, <http://journalofvision.org/3/9/73/>, doi:10.1167/3.9.73.
- Fresnel, A. J. (1821/1866). *Oeuvres completes d'Augustin Fresnel*. Ed. de Senarmont, H., Verdet, E. & Fresnel, L.. Paris, France: Imprimerie imperiale.
- Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin.
- Hartung, B., & Kersten, D. (2002). Distinguishing shiny from matte [Abstract]. *Journal of Vision*, 2(7), 551a, <http://journalofvision.org/2/7/551/>, doi:10.1167/2.7.551.
- Hartung, B., & Kersten, D. (2003). How does the perception of shape interact with the perception of shiny material? [Abstract]. *Journal of Vision*, 3(9), 59a, <http://journalofvision.org/3/9/59/>, doi:10.1167/3.9.59.
- Huggins, P.S., Chen, H.F., Belhumeur, P.N., and Zucker, S.W., (2001). Finding Folds: On the Appearance and Identification of Occlusion, in CVPR'01 (*Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, Kauai, HI, December 2001), Vol. 2: 718-725, IEEE Computer Society.

- Knutsson, H. (1989). "Representing local structure using tensors," in *The 6th Scandinavian Conference on Image Analysis*, Oulu, Finland, June 1989, pp. 19–22.
- Koenderink, J. J., & van Doorn, A. J. (1980). Photometric invariants related to solid shape. *Optica Acta*, **27**(7): 981—996.
- Malik, J., & Rosenholtz, R. (1997). Computing local surface orientation and shape from texture for curved surfaces. *International Journal of Computer Vision*, **23**, 149–168.
- Mingolla, E. & Todd, J. T. (1986). Perception of solid shape from shading. *Biological Cybernetics*, **53**: 137—151.
- Norman, J. F., Todd, J. T. & Orban, G. A. (2004). Perception of 3D Shape from Specular Highlights, Deformations of Shading, and other Types of Visual Information. *Psychological Science*.
- Oren, M. and Nayer, S. K. (1996). A Theory of Specular Surface Geometry. *International Journal of Computer Vision*, **24**: 105—124.
- Savarese, S., Li, F. F., & Perona, P. (2003). Can we see the shape of a mirror? *Journal of Vision*, **3**(9), 74a. <http://journalofvision.org/3/9/74/>, doi:10.1167/3.9.74.
- Super, B. J. and Bovik, A. C. (1991). "Three-dimensional orientation from texture using Gabor wavelets," in *Proc. SPIE Conf. of Visual Communications and Image Proc.*, vol. 1606, pp. 574-586.
- Super, B., & Bovik, A. C. (1995). Shape from texture using local spectral moments. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **17**, 333-343.

Todd, J. T. & Mingolla, E. (1983). Perception of surface curvature and direction of illuminant from patterns of shading. *Journal of Experimental Psychology: Human Perception and Performance*, **9**: 583—595.

Todd, J. T., Norman, J. F., Koenderink, J. J., Kappers, A. M. L. (1997). Effects of texture, illumination, and surface reflectance on stereoscopic shape perception. *Perception*, **26**: 807—822.

