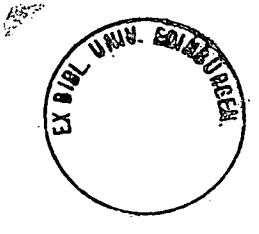
Inferring Surface Shape from Specular Reflections.

Gavin J. Brelstaff

Ph.D., University of Edinburgh, 1988.



Abstract

This thesis concerns the field of machine vision, in which computers are used to interpret pictures provided by television cameras. In particular it tackles the problems of interpreting pictures containing the glossy highlights produced by specular reflection. These specular highlights or specularities occur only where a glossy surface reflects the direct illumination of a light source almost straight into the camera's aperture - as if the surface were a mirror. Practical vision systems ought to detect specularities for two reasons: firstly so that they will not be mistaken as surface markings, and secondly so that they can be used to infer shape information about the surface upon which they appear to lie. Schemes for both detection and shape inference are discussed and developed in this thesis.

Surfaces which are not glossy - Lambertian surfaces - form images of a certain constrained class. The detection scheme described in this thesis identifies regions of an image where the Lambertian constraints break. These regions are candidate specularities. In all, the evidence provided by tests on three separate constraints is combined to detect specularities. The scheme will work even for achromatic ('black and white') cameras.

Two independent methods of inferring surface shape information are discussed. The first uses the manner in which the surface distorts the image of the light source. The second requires two cameras and uses the stereo disparity of the specularity relative to a nearby surface marking. The results obtained by the two methods can be combined to provide fuller information. Often the local surface curvature can be fully recovered.

The detection and inference schemes have been implemented in a software system coded in C, together with a user friendly interface for interrogating the results.

Acknowledgements

First and foremost, I am indebted to Dr Andrew Blake, who initiated my interest in Computer Vision, supervised this project and provided a constant barrage of invaluable suggestions that inspired me to follow many fruitful lines of research.

Excellent suggestions, accurate criticism and pertinent mathematical advice flowed freely from my friend and colleague Dr Andrew Zisserman. Much vital support was supplied by the advice criticism and humour of Constantinos Marinos—not to mention Sabine.

Thanks goes to the team at AIVRU, Sheffield for their help, especially to Stephen Pollard for assistance in acquiring images. This work was done under the umbrella of the ALVEY IKBS Image Interpretation Research Theme, Consortium II. I am grateful to IBM and SERC for their support through my CASE studentship, and again to SERC for support under grant GR/D 1439.6. The huge quantity of diligent proof-reading performed by Howard Hughes was greatly appreciated. Thank are also due to all the people that were the Department of Computer Science during my stay—particularly helpful were Duncan Ballie, Ian Nixon Alistair Sinclair and Hugh Stabler. Credit also goes to the Department itself for providing such excellent facilities and magnanimously extending my welcome. Thanks also to Liz, John, David, Cecilia, Andy and Cinzia.

Declaration

This thesis is my own composition, and the research reported herein is my own work in accordance with the University Regulations.

Table of Contents

Chapter 1

Introduction

A large variety of visual cues aid our perception of three-dimensional scenes. Those that are provided by specular reflection motivate the work in this thesis. A simple example of when specular reflection plays an important role in visual perception can be observed by viewing the hollow end of a metallic soup spoon[1]. The most prominent visual cue is the bright blob, or specularity, formed by the specular reflection of a light source on the curved surface of the spoon. Without this cue the surface shape would be relatively unconstrained. It might be convex or concave; it might be highly or gently curving. The specularity on the surface constrains the curvature. Its rounded shape suggests that a spherical surface is possible and its apparent leftward motion when the viewpoint switches from the left to the right eye, suggests a concavity. Turn the spoon over and the specularity appears to move to the right, suggesting a convex sphere. In this thesis, methods are developed that exploit both the shape and motion of specularities for the purposes of machine vision. Figure 1–1 provides an example of the results achieved: both radii of curvature of an ellipsoid are determined to a reasonable accuracy. The purpose of machine vision is the extraction of useful descriptions of the 3-D surfaces from the images supplied by t.v. cameras, such as the image pair shown in figure 1–2. Once obtained, these descriptions can

[1]Not included.

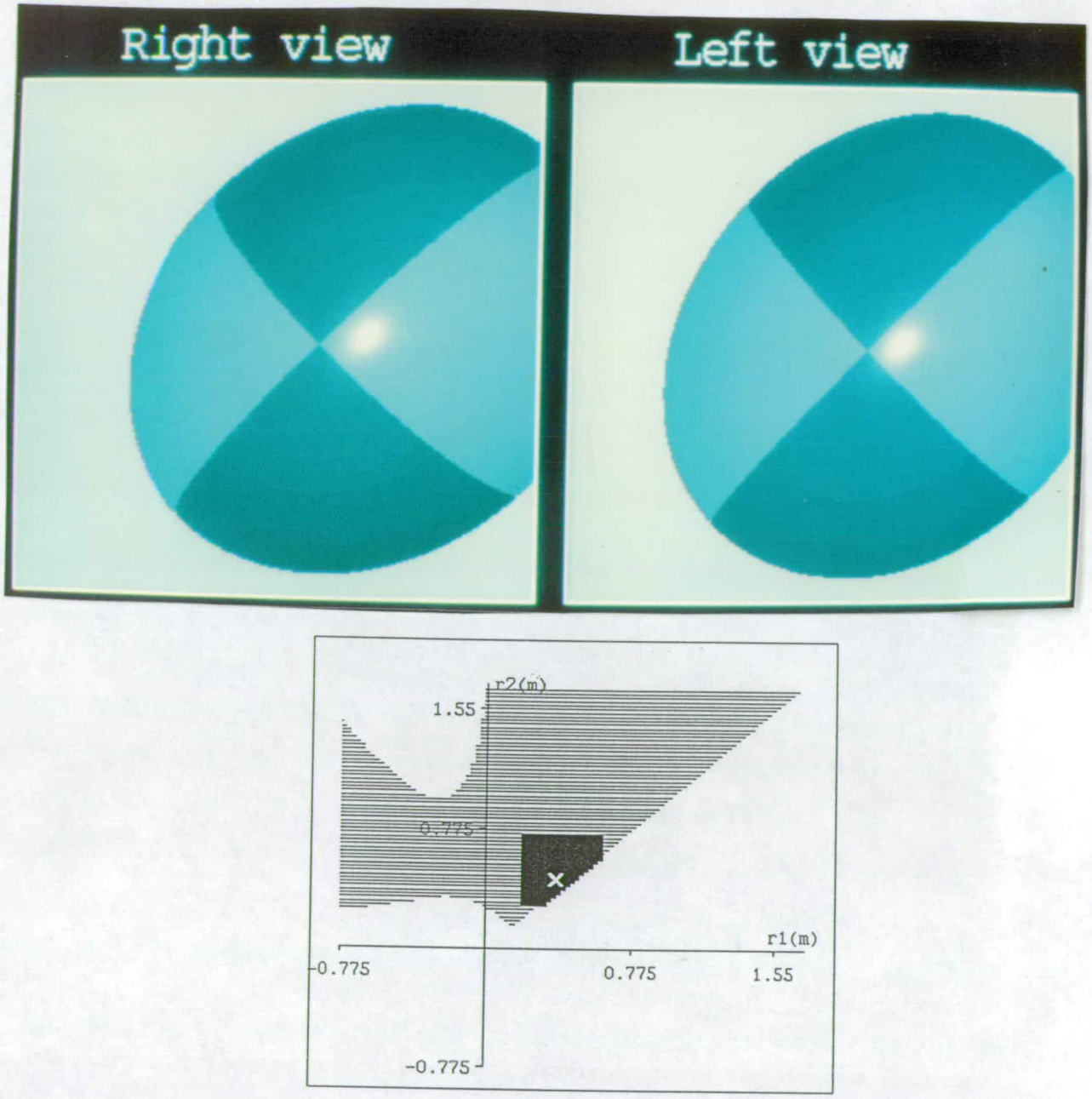


Figure 1–1: *An example of the results obtained by the shape-from-specularity scheme .* **Top***: A stereo pair of images of an ellipsoid—aligned for binocular fusion.* **Below***: The scheme localises the radii of curvature r_1 and r_2 in the vicinity of the specular reflection to lie within the black-box marked on the graph of (r_1, r_2) space. Thus providing a useful estimation of their exact values—indicated by the white cross.*

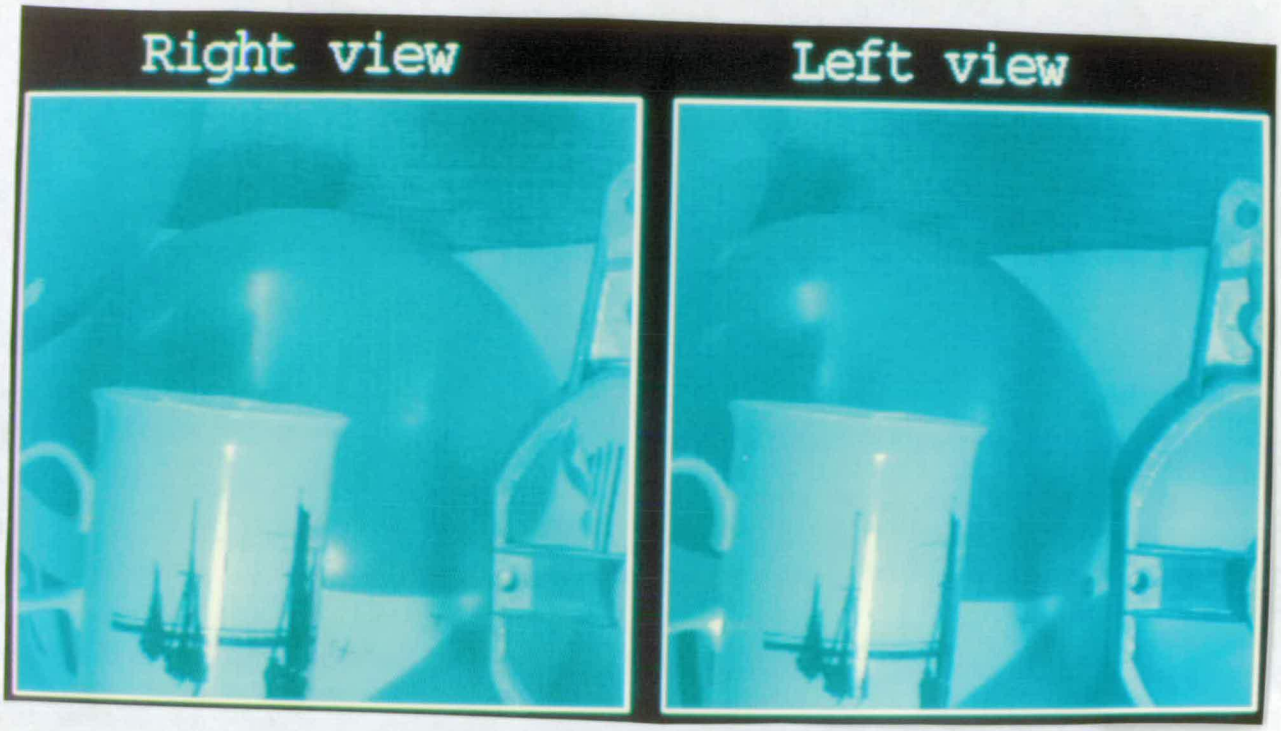


Figure 1–2: *A typical stereo-pair of images obtained by t.v. cameras, Specularities in these images have been detected and used to infer surface curvature.*

be applied to robotics tasks such as object recognition, collision avoidance and automated production. However, obtaining these descriptions involves much more than merely pointing the camera in the right direction. The "raw" images provided by cameras are two-dimensional and usually noise-ridden—while three-dimensional, noise-resistant descriptions are required for robotics applications. Workers have found some success in devising methods that exploit visual cues, in order to extract useful surface descriptions from "raw" images. However, the descriptions provided so far, are as yet incomplete: in comparison with those effortlessly provided by our eyes they are rather impoverished. By exploiting a previously untapped visual cue—specular reflection—the methods developed in this thesis enrich the surface descriptions available to machine vision systems. Although specular reflections are not guaranteed to be present in every image, they are the norm, rather than the exception, for a robot manipulating metal parts.

The methods developed in this thesis—where possible—accord with the principles of machine vision—set down by Marr in his book *Vision* [83], i.e:

- The descriptions that they provide *degrade* fairly *gracefully* in adverse conditions
- They adhere to the principle of *least commitment*: preferring, in the face of uncertain cues, to register a description as uncertain rather than risk a false one.
- They are fairly *robust* to variations in lighting conditions, viewing geometry, surface properties and noise levels.

In the past, specularities have themselves, been regarded as a variety of noise, because:

1. They create prominent features in unpredictable positions upon otherwise recognizable objects.
2. They disrupt an existing method (binocular stereo) that estimates the depth of surfaces.

The best that was hoped for was that some method of detecting and excising the "specular-noise" would be found. The specularity detector developed in this thesis creates descriptions that can be used for this purpose. They are the same descriptions as are used to infer surface shape. They can also be used to identify surfaces that are glossy: glossiness is after all a useful recognizable surface property.

The task of detecting specularities is harder than it seems at first sight: their characteristics—shape, size and "brightness" are not always the same. Physical models exist that predict their characteristics, for some surfaces of known material and known shape, lit by a known illumination. Without detailed knowledge of the surface and lighting conditions, a general-purpose specularity detector cannot make direct use of these models. In any case, physical models

do not exist for all common surface materials. The detector developed in this thesis is based on a much simpler model: that of a completely matt (non-specular) surface. By establishing constraints on the images that such a surface forms, specularities are detected at image locations where they are seen to be violated. In this way three different tests for specularities are developed: a test for the specularities that:

- are too "bright" to be matt. This test requires the image to be pre-processed through a retinex: a process that effectively removes gradual illumination variations.

- have peaks that are too "sharp" to be matt.

- have edge-contrasts that are too great to be matt.

The specularity detector is essentially a low-level autonomous process, able to extract concise 2-D descriptions of the specularities present in 'raw' images. In this respect, it fits into the lower level of the hierarchical structure, proposed by Marr [83] for the purpose of obtaining 3-D descriptions of surfaces. This hierarchy consists of a set of modular processes—each of which is self-contained and only communicates with the others via the descriptions created. 3-D surface descriptions cannot be created directly from the "raw" images. However, by applying a set of low-level autonomous processes to the "raw" images, concise 2-D descriptions—or base representations (BR's)—of image features, conveying various visual cues, can be extracted. These BR's can then be read by higher-level processes—or knowledge sources (KS's)—to create 3-D surface descriptions. The advantage of this *Marrian* approach is that individual modules can be designed independently—simplifying the substantial task of machine vision into more manageable sub-tasks. Figure 1–3 shows how both the specularity detector and the shape-from-specularity modules, developed in this thesis, fit into this structure. The specularity detector, the edge detector and the retinex are all autonomous, low-level processes that create BR's. The specularity detector reads the other two BR's in order to create its own. The binocular

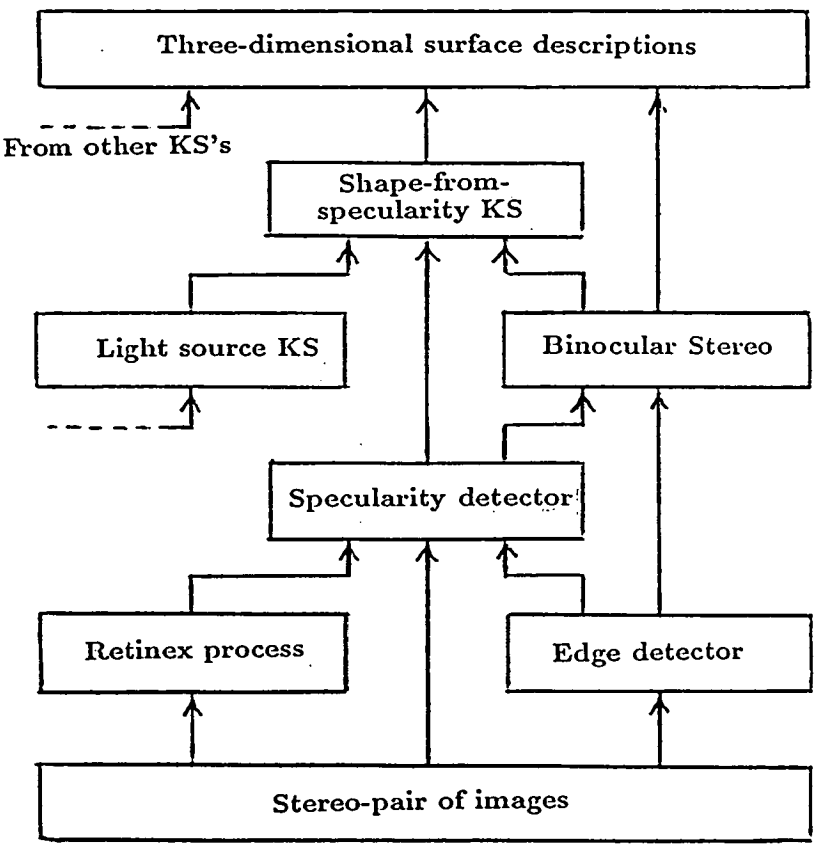


Figure 1–3: *How the specularity detector and the shape-from-specularity module fit into a hierarchical structure.*

stereo and shape-from-specularity KS's read certain BR's and create 3-D surface descriptions. Binocular stereo is integrally linked to the processes concerning specularities—it uses the detector's BR to excise unwanted specular features and creates a depth map that is used by the shape-from-specularity module. Note: shape-from-specularity is not an autonomous process—it requires access to a KS that reasons about the light sources that illuminate the visible surfaces.

A variety of KS's have been designed to exploit the different visual cues, e.g: shape-from-shading can exploit matt shading variation and shape-from-texture deals with texture variations. No KS on its own will always provide the complete 3-D surface description. Binocular stereo, for example, only estimates depths along contours visible in the image—and certainly not at specularities. The shape-from-specularity module helps fill in gaps in the description, by providing curvature information at points in regions between the contours.

A shape-from-specularity module was first proposed by Blake [13,15]. The scheme developed in this thesis represents a significant refinement of Blake's original ideas. The scheme is essentially an extended application of established geometrical techniques.

1.1 Organisation of the rest of this thesis

Chapter 2: provides an extensive review of the process of the reflection of light by surfaces. Both the geometrical aspects (e.g: the BDRF) and the physical processes are considered. Conclusions pertinent to the detection of specularities are provided.

Chapter 3: reviews previous work done into detecting specularities. It also considers the current work that is being carried out that exploits colour cues—unlike the detector in this thesis. Also described is the retinex process, that was mentioned above.

Chapter 4: presents the specularity detector: its development from the matt-world constraint, the formulation of the three tests, a description of a scheme that combines the evidence provided by each test and creates the descriptions of specularities. Finally, there is an evaluation of the descriptions provided by the scheme for a test set of images.

Chapter 5: reviews previous work done into the inference of surface shape—concentrating particularly on the role of specularities. Included is a description of the binocular stereo process, mentioned above. It ends with an extensive review of Blake's ideas for inferring shape from specularities.

Chapter 6: presents new work on shape-from-specularity. This includes the development of both monocular and stereoscopic inferences and ends by describing the scheme that combines them, in order to extract curvature information.

Chapter 7: describes the implementation of the shape-from-specularity scheme as a user-friendly interactive computer program. A step-by-step "walk-through" of the the program, analysing the specularities present in a real stereo pair of images, is provided. Readers eager to get a flavour of this scheme, might start by browsing this section. The chapter concludes with an evaluation of the results obtained by applying the scheme to a substantial set of images.

Appendices of particular interest are appendix J that details the software implementation of the detection and inference schemes and appendix K that contains a published paper [22] describing the implementation of the retinex process, as employed by the specularity detector.

Chapter 2

Surface reflection

2.1 Introduction

A substantial body of knowledge exists concerning the reflection of light by surfaces—consisting of both theoretical models and experimental data. This chapter reviews the state of this knowledge. Emphasis is placed on those aspects that are directly relevant to the tasks of this thesis: specularity detection and shape inference from specularities. The review splits into two distinct parts:

- Consideration of the geometrical aspects of surface reflection. This ends in a discussion of the highly applicable concept of the reflectance map.

- Consideration of the physical processes involved when light interacts with different surface materials.

First, a description of surface reflection in terms of ray optics is given. In this way the basic concepts and problems involved are introduced in an informal manner that is familiar to most readers. Later it proves necessary to introduce some formal definitions of the concepts.

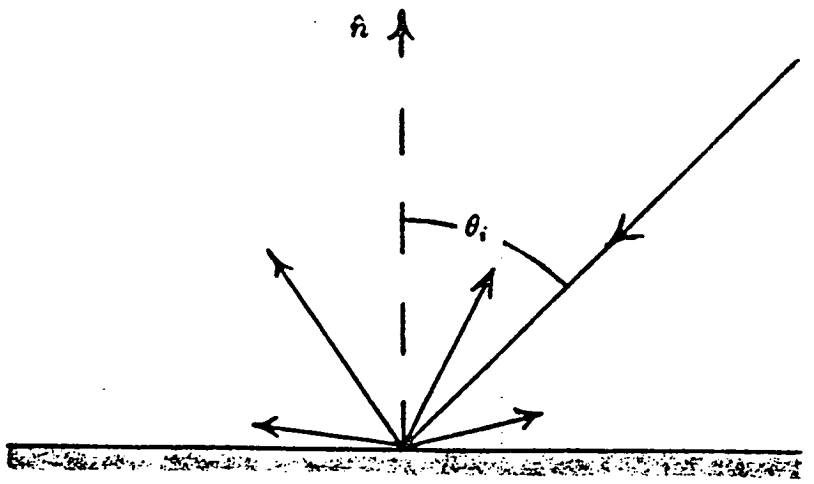


Figure 2–1: *When an incident ray impinges upon a surface, rays can be reflected in many different directions.* θ_i *is the angle of incidence and* $\hat{\mathbf{n}}$ *is the local surface normal direction*

2.2 Description of reflection in terms of ray optics

2.2.1 Introduction

The light incident on and reflected by a surface can be considered in terms of light rays—i.e. beams of infinitesimal width. Consider the case shown in figure 2–1. An incident ray impinges upon a surface along a direction at angle θ_i with the local surface normal vector $\hat{\mathbf{n}}$. As a result of reflection processes on or just below the surface new rays are reflected back into the environment. The reflection of light by a surface can be separated into two different processes:

1. specular reflection—in which the incident light is reflected back into the environment without entering the surface.

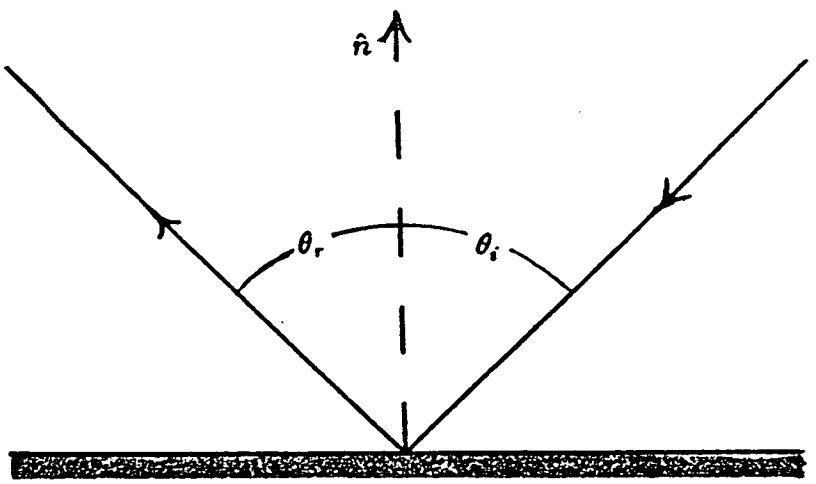


Figure 2–2: *In the specular case an incident ray is reflected as a single ray—in the direction of mirror reflection. The angle of incidence, θ_i is equal to the angle of reflection, θ_r.*

2. diffuse[1] reflection—in which the light enters the surface material and is scattered so as to reemerge back into the environment.

The case of completely specular and completely diffuse reflection can be considered in terms of rays. The specular case is described as a mirror-like reflection in which a single ray is reflected—as shown in figure 2–2. The reflected ray lies in the same plane as both the incident ray and the local surface normal $\hat{n}$. The angle of reflectance θ_r is the same as the angle of incidence θ_i. In the diffuse case rays are reemitted in many directions contained within the hemisphere above the surface point—as depicted in figure 2–3. In both the specular and diffuse cases the irradiance[2] that is measured along the reflected rays depends upon

[1]Later it is shown that diffuse reflection can occur when rough surfaces reflect in a specular manner. So this type of reflection is better termed subsurface scattering or volume reflection.

[2]Irradiance is a precise term for brightness or intensity—it is defined in appendix A.

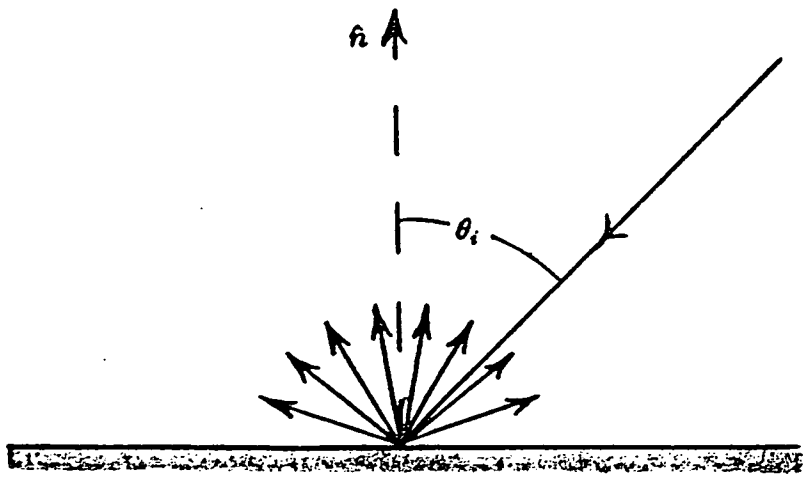


Figure 2–3: *In the diffuse case an incident ray is reflected in many different directions.*

the angle of incidence θ_i. Below simple examples of both cases are reviewed, before considering the problems introduced when they both occur at once.

2.2.2 A simple specular case

In the most simple specular case—that of an optically smooth surface—the angular dependence is in accordance with the Fresnel relations. This results in an attenuation of the reflected light by a factor F which depends explicitly on the angle of incidence θ_i. Figure 2–4 shows typical $F(\theta_i)$ distributions. So for a smooth surface illuminated by a single ray the irradiance[3] E_S is:

$$E_S(\theta_i) = \begin{array}{ll} E_o F(\theta_i) & \text{along the direction of mirror reflection } \theta_r \\ 0 & \text{for any other direction.} \end{array} \tag{2.1}$$

E_o is the irradiance of the incident ray.

---

[3]In fact an infinitesimal ray strictly contains no energy and thus has irradiance zero [62]. Irradiance is defined over finite solid angles. For present purpose it is reasonable to assume that the ray occupies a small finite solid angle and has a non-zero cross-section.

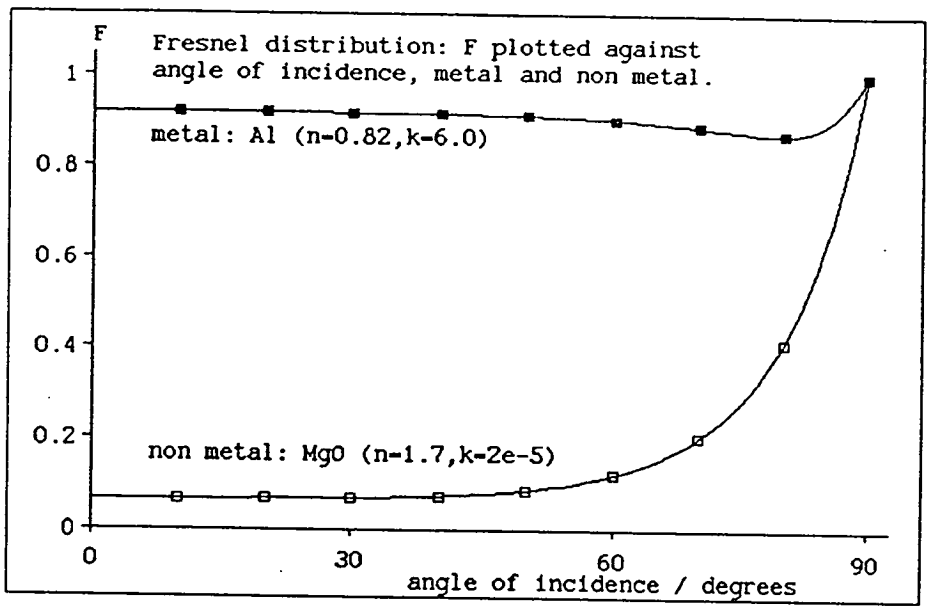


Figure 2–4: *Typical $F(\theta_i)$ distributions for a metal and a non-metal.*

Reflection by rough specular surfaces is more complicated. However for these surfaces the θ_i dependence is still present—see section 2.4.

2.2.3 A simple diffuse case

For the most simple case of diffuse reflection the irradiance measured along all directions of reflection is proportional to $\cos\theta_i$—in accordance with Lambert's Law (see section 2.4). In this case the image irradiance along all directions of reflection is uniform and given by:

$$E_d(\theta_i) = E_o \rho_d \cos(\theta_i), \tag{2.2}$$

where E_o is the irradiance or the incident ray and ρ_d is a constant of the material, known as albedo (or the diffuse reflection coefficient). It represents the proportion of incident light energy that is reflected by a Lambertian surface. In theory it can range from zero to one. When albedo is small a surface appears dark, and when it is large a surface appears light. In many practical circumstances albedo has a limited dynamic range. This fact is important in

the development of one of the tests designed to detect specularities, described in chapter 4.

2.2.4 Composite surfaces

A composite surface is one that reflects light in both a specular and a diffuse manner. A question of immediate interest in this thesis is whether a composite surface can be conveniently characterised by the ratio of the amount of light flux it reflects specularly to that which it reflects diffusely. If this were the case, this ratio would provide an attractive constant of the surface material—which a specularity detector might attempt to estimate. In short the answer is that no such characterisation is possible. The bidirectional reflectance distribution function (BRDF) in general is required to properly define the ratio. It describes how the ratio varies with direction of incidence and reflection and is discussed further in section 2.3.2.

However it is useful to continue the discussion in terms of ray optics, and consider a simple composite surface. It soon becomes clear why the BRDF must be defined. Consider again the single ray incident upon a surface.

This time a ray of irradiance E_o is incident upon the surface. At the surface a ray is immediately specularly reflected. Its irradiance is $E_o F(\theta_i)$ in accordance with equation (2.1). The remainder $E_o(1 - F(\theta_i))$ penetrates into the surface layer and may be diffusely scattered—in accordance with Lambert's law. In this case the overall irradiance[4] distribution is:

$$E(\theta_i) = E_o F(\theta_i) + [E_o - E_o F(\theta_i)]\rho_d \cos\theta_i, \tag{2.3}$$

along the direction of mirror reflection, and

$$E(\theta_i) = [E_o - E_o F(\theta_i)]\rho_d \cos\theta_i, \tag{2.4}$$

---

[4]See previous footnote.

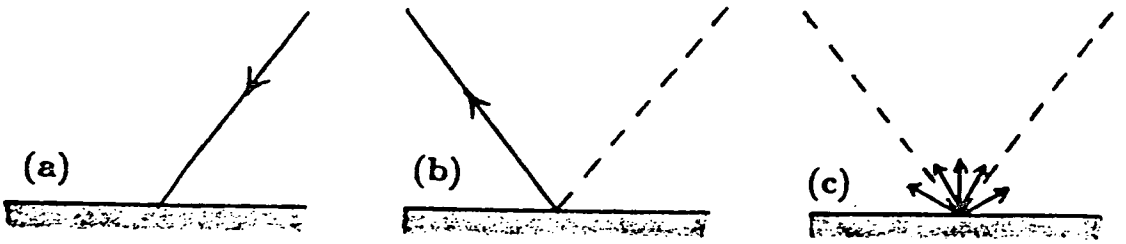


Figure 2–5: *For a composite surface an incident ray, (a), is partially specularly reflected, (b) and partially diffusely reflected (c).*

along any other direction. Figure 2–5 depicts how the two different processes combine. The effect of the term $F(\theta_i)$ in these two expressions ensures that the ratio of reflected to incident flux depends upon the angle of incidence. So even for this relatively elementary composite surface, this ratio is not a simple constant of the surface material—it also depends both upon the beam geometry and the material constants. It should be noted that this ratio has been extensively used as a parameter for reflectance models used in computer graphics [31,20]. These models have been used to synthesise some very realistic looking images of composite surfaces. However two points must be made about these images:

1. An image made by chosing an arbitrary 'ratio-parameter' is not guaranteed to correspond to any real world surface.

2. Only a small subset of the real images of composite surfaces is available by variation of the 'ratio-parameter'. Many real surfaces are not represented by this model. It is thus unsuited for use in detecting real specularities.

The next section discusses the geometrical and radiometrical concepts including the BRDF. The BRDF provides a much more precise definition of the invariant surface properties than the 'ratio-parameter' discussed above.

2.3 Geometrical considerations

2.3.1 Introduction

The equations (2.1), (2.2), (2.3) and (2.4) provide some geometrical characterisations of surface reflection. They are of limited use for two reasons:

1. they concern very special cases.

2. they apply only to narrow rays.

Equation (2.5), below, provides a generalisation of these descriptions which also encompasses the reflection of incident light beams (many rays). As such, this expression provides a suitably complete specification for surface reflection, upon which the specularity detector, developed in chapter 4, can be based. Although the expression has a simple form, its derivation [85,61] is not trivial. The beam geometries must be precisely defined and many Radiometrical concepts must be introduced. The beam geometries are defined using polar coordinates solid angles and infinitesimal directional quantities. The Radiometrical concepts are precisely defined measures of light—such as radiant flux, irradiance, radiance, reflected radiance, incident radiance, scene radiance and the bidirectional reflectance distribution function. Rather than delaying the reader by stopping here to introduce so many new terms, their precise definitions are given in appendix A. From the informal definitions given in the main text the reader should be able to glean their meanings. Wherever a precise definition is in doubt the appendix should be consulted.

2.3.2 The image irradiance equation

The precise radiometrical term for the brightness of the light recorded by an imaging device is image irradiance. This quantity is of prime importance in machine vision as it is the only information that is directly available from a camera.

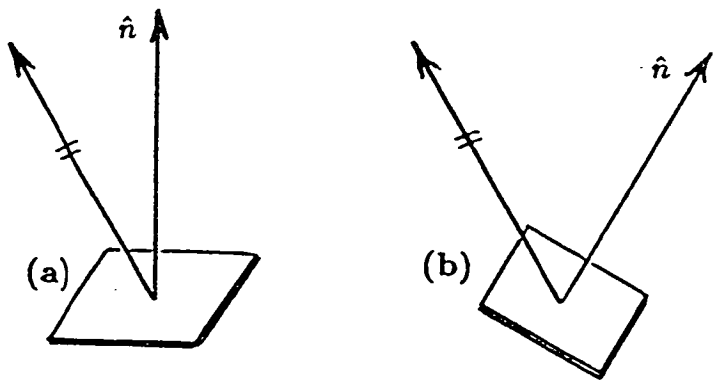


Figure 2–6: *(a) The normal direction $\hat{\mathbf{n}}$ to the local tangent plane of the surface can used be to specify the local surface orientation. (b) As orientation varies so does image irradiance. The local reflectance map specifies this variation.*

The image irradiance, E_P measured at a point P in an image is determined by the image irradiance equation [85,61]:

$$E_P = a_c R(\hat{\mathbf{n}}), \tag{2.5}$$

where a_c is a known constant of the camera—see (A.5). R is a function known as the reflectance map. Its parameter $\hat{\mathbf{n}}$ denotes a dependence on the local surface orientation. This orientation is conveniently specified by the direction of the local surface normal vector $\hat{\mathbf{n}}$ with respect to the view vector—see figure 2–6 (a). The reflectance map describes the reflection at the surface point within the scene that is imaged at P. The image irradiance E_P varies as the orientation of the surface at this point is changed—as in figure 2–6. The reflectance map specifies this variation. It is a function that combines two separate kinds of information:

1. information that quantifies the local reflectance properties of the surface material, in different directions. The bidirectional reflectance distribution function (BRDF), denoted f_r, allows a complete specification of these properties. Basically the BRDF says how bright the surface appears in

an arbitrary viewing direction when illuminated along another arbitrary direction. See section A.1.2 for more details.

2. information about the local distribution of illumination over all the different directions of incidence. The incident radiance distribution L_i provides the appropriate specification in this case. See section A.1.2 for more details.

Each value, R, in the reflectance map, the BRDF and the incident radiance distribution are related. The precise relationship is provided by equation (A.16) in the appendix. Briefly R is determined by integrating the appropriate product $f_r L_i \cos\theta_i$ over a hemispherical solid angle[5]. Once determined, values of R, can be used to predict the image irradiance via the image irradiance equation (2.5). Machine vision has frequently used the image irradiance equation to provide a constraint for computational schemes. These schemes often rely on explicit knowledge of the reflectance map. The shape from shading schemes proposed by Horn [67,60] are typical examples. However, from the discussion above, it is clear that the reflectance map is only known if both the BRDF and the incident irradiance distribution, L_i, are known. A practical vision system must be able to deal with a variety of different surfaces under a variety of illuminations. So it is unrealistic to expect such a system to know the BRDF and L_i for the surfaces that it views.

Horn's schemes have usually been tested using simple models of reflectance maps. Below some of these simple models are considered. They provide some useful insights into the forms that reflectance maps assume. For example figure 2–7 shows that specular and diffuse processes can produce characteristic

[5]Again θ_i is an angle of incidence. L_i is incident radiance, which roughly describes the amount of light incident along a given direction. More details of radiance are given in appendix A.

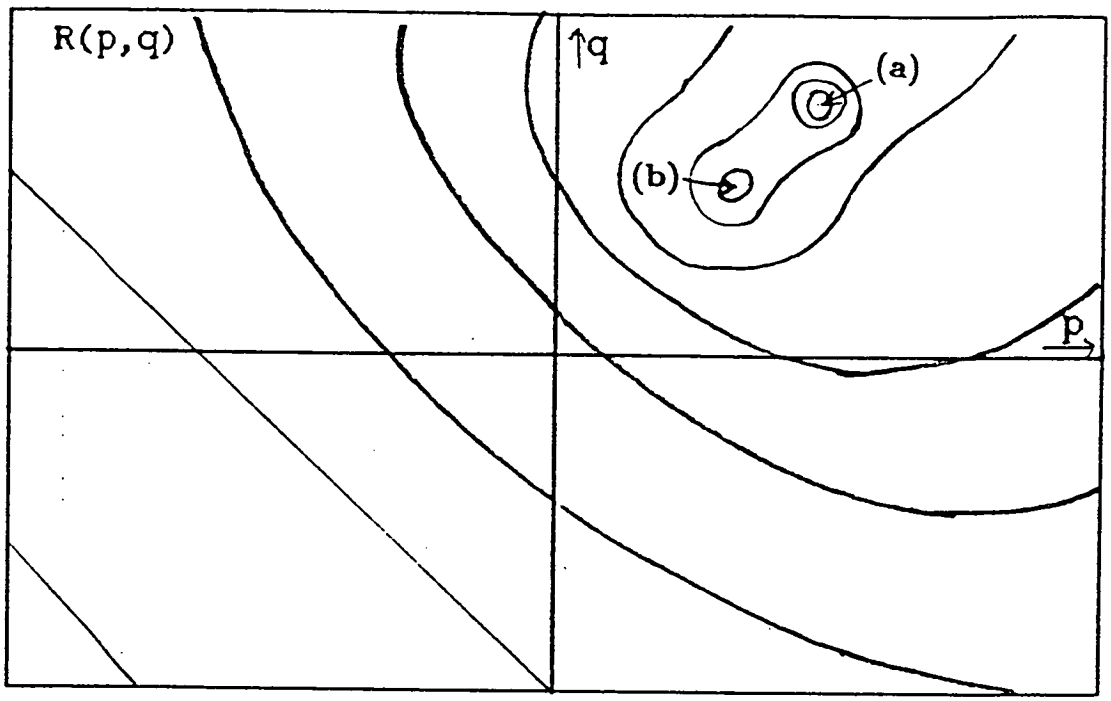


Figure 2–7: *A reflectance map made from a simple model of a directionally lit glossy surface. The contours link orientations of equal values of R. (a) and (b) mark the orientations corresponding to peaks in specular and diffuse reflection respectively.*

peaks in the reflectance map[6] . Lambertian reflectance maps prove to be particularly simple to analyse. Under many realistic illuminations it transpires that they are restricted in their dynamic range. This restriction provides a useful constraint which is later used to detect specularities.

2.3.3 Constraints on reflectance maps

The form of a reflectance map—like that depicted in figure 2–7— depends on both the reflecting properties of the surface (through the BRDF) and the local

[6]Figure 2–7 is depicted using the (p, q) coordinates of orientation defined in section A.2.3.

illumination distribution (through L_i). Theoretically the BRDF and L_i might take on any distribution. However not all the possible distributions are realised in reality. The physical surface properties constrain the form of real BRDFs, while many of the possible distributions of L_i are very unlikely to occur in natural circumstances. So the reflectance maps encountered in the real world are constrained in their form. For Lambertian surfaces the BRDF is particularly simple—it is constant (see equation (A.10))—and so the constraints are determined by the L_i distribution. The next sections examine constraints on the reflectance maps of Lambertian surfaces. These constraints are useful when developing the specularity detector in chapter 4. Reflectance maps for non-Lambertian surfaces involve some specular reflection and are more difficult to analyse—see section 2.4.

2.3.4 Lambertian reflectance maps

Lambertian reflectance maps describe surfaces that reflect no light specularly and reflect in accordance with the cosine law. Under realistic illumination conditions their forms are constrained—as examples show below. These constraints for Lambertian surfaces do not necessarily hold for other surfaces that reflect some light specularly. The specularity detector developed in chapter 4, looks for image features that break constraints that hold for Lambertian surfaces.

For a given Lambertian surface the BRDF, f_r has a constant value, ρ_d/π, where ρ_d is the surface albedo. This makes the reflectance map, R_L particularly simple:

$$R_L = \frac{\rho_d}{\pi} \int_{\omega_i} L_i \cos\theta_i d\omega. \tag{2.6}$$

The derivation of this equation is discussed in appendix A. This expression shows that R_L, at a given point on the surface, only varies with the incident radiance distribution, L_i. The value of this integral has been found for a few relatively simple L_i distributions. As a result a variety of Lambertian reflectance maps have been derived. These maps provide a useful insight into the kind

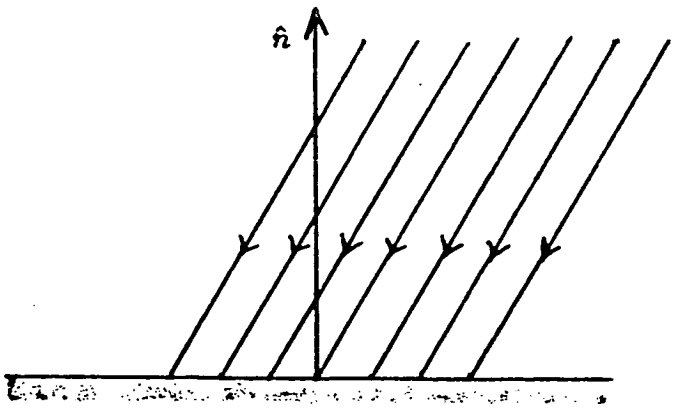


Figure 2–8: *For collimated illumination, light is incident along only a single direction.*

of constraints on Lambertian reflectance maps that are likely to hold under realistic illumination conditions. The next sections review these maps.

2.3.5 Collimated sources

A collimated beam propagates light energy along a single direction. When it illuminates a point on a surface the light is incident only along that direction—see figure 2–8. This is analogous to the single incident ray discussed in section 2.2. In this case the reflectance map can be shown [61] to be:

$$R(\hat{\mathbf{n}}) = \frac{E_o \rho_d}{\pi} \max(0, \cos\theta_i(\hat{\mathbf{n}})) \tag{2.7}$$

where E_o is the irradiance measured perpendicular to the collimated beam, and ρ_d the albedo. The dependence of R on the local surface orientation ($\hat{\mathbf{n}}$) lies entirely within the last term. The angle of incidence θ_i implicitly depends on the normal direction $\hat{\mathbf{n}}$. The 'max' term ensures that R is zero along all normal directions pointing more than 90 degrees away from the incident beam, i.e. where the surface shadows itself.

The expression (2.7) introduces two important quantities—the maximum, R_{max} and minimum, R_{min} values of the reflectance map. For the collimated

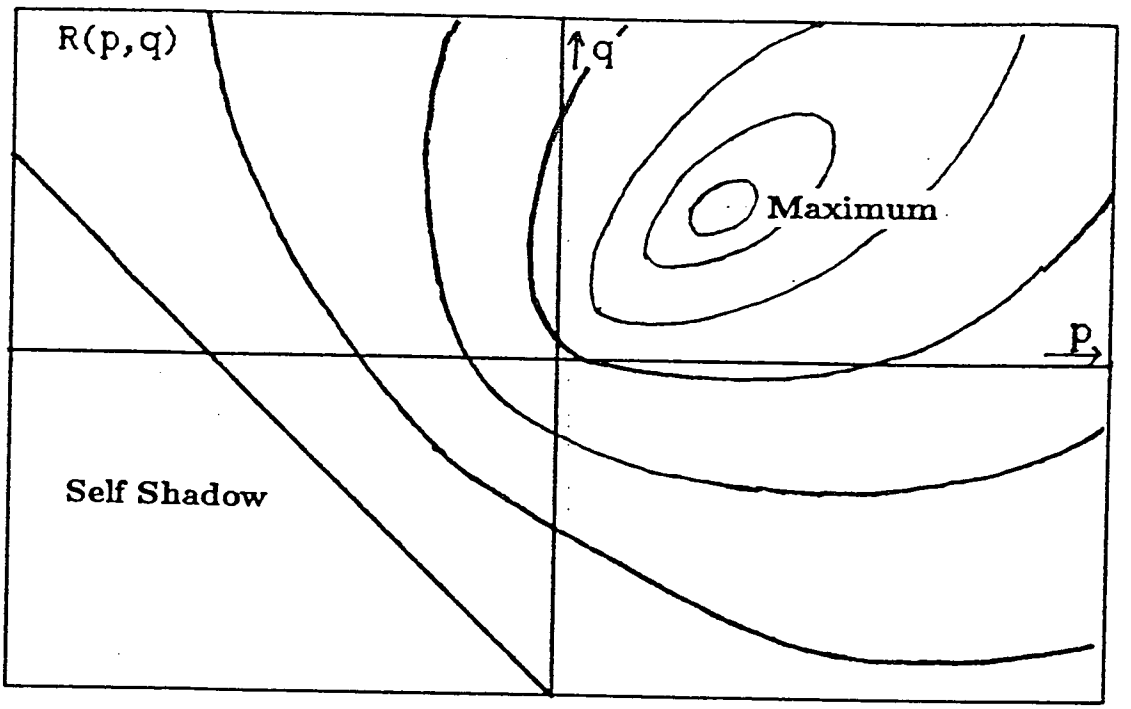


Figure 2–9: *The reflectance map for a collimated source illuminating a Lambertian surface. Its maximum is along the direction of the incident light. It is minimum in the self shadow region.*

beam R_{max} corresponds to the surface normal being oriented along the direction of incidence and R_{min} corresponds to the value in the self-shadow region. Figure 2–9 shows the reflectance map and marks these two features. That R_{min} is zero for (A.9) is an indication of the deficiency of a collimated source model. In real circumstances self shadow regions, although dark, are never completely black. Interreflection between surfaces in a real environment acts to scatter light into the self-shadow regions. This scattered light is often called ambient—because it falls on to surfaces from all directions. In reality interreflection gives rise to very complex effects and it usually accounts for 30 percent of the total illumination [86]. Workers in computer graphics [31] have improved the degree of realism of the images that they synthesise by modifying equations like (2.7) to include a term for the ambient light:

$$R(\hat{\mathbf{n}}) = \frac{E_o \rho_d}{\pi}[A + (1 - A)\max(0, \cos\theta_i(\hat{\mathbf{n}}))]. \qquad (2.8)$$

The effect of ambient light is approximated by introducing the constant term A. In theory A may range between zero and unity, in practice a value of $A = 0.3$ achieves realistic results.

For this modified reflectance map, R_{min} need no longer be zero. In fact for the typical case, when $A = 0.3$, the dynamic range of R, i.e. R_{max}/R_{min}, is roughly $3r_\rho$, where r_ρ is the dynamic range of ρ_d which is considered later. This constitutes a simple constraint on this particular Lambertian reflectance map.

The limitations of the reflectance map described by (2.8) are twofold:

1. complex interreflections are not modelled.

2. the distribution of illumination corresponds to a single distant point light source. A real light source extends over a finite solid angle. It is reasonable to expect the whole sky to contribute to the illumination.

In the next sections both of these limitations are discussed.

2.3.6 Complex interreflection

The so called Radiosity Method, adapted from techniques of thermal engineering [42], has recently been used in computer graphics to account for both extended light sources and complex interreflections [29,86]. The needs of architects and designers, to realistically simulate the lighting conditions within specified environments, both indoors and outdoors, motivated this approach. The method imposes the Law of Conservation of Energy for the light interreflected within a closed system, using BRDF and other Radiometrical concepts. The method is capable of simulating images containing very realistic lighting effects [29,86]. However to do this the BRDF of each surface must be specified as well as the shape size and location of each light source. This thesis seeks a low-level process able to detect specularities in a variety of different circumstances. It is unrealistic to expect this process to have a comprehensive knowledge of the BRDF and lighting conditions within an arbitrary view. So the Radiosity

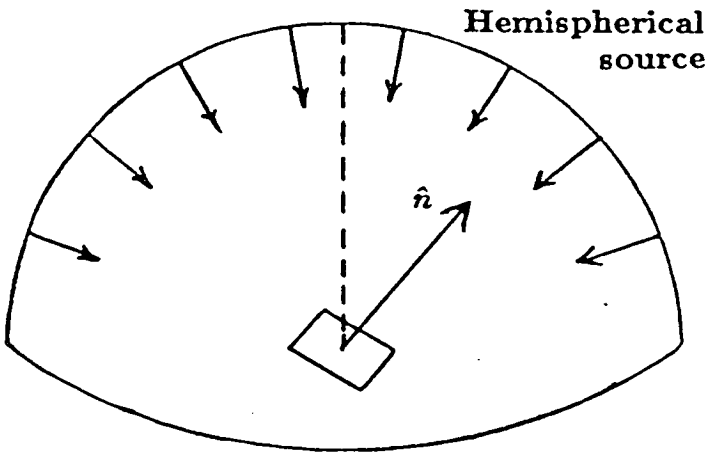


Figure 2–10: *For a uniform hemispherical source the incident radiance is equal along all directions within the hemisphere.*

method is not particularly useful when developing such a process. In this thesis interreflection between surfaces is approximated using a constant isotropic ambient illumination. Section 2.3.5 proposed that this approximation was reasonable when considering a highly directional illumination, i.e. a collimated source. Below it is proposed that it is also reasonable when considering more widely distributed illuminations.

2.3.7 Extended illumination distributions

The illumination from the sky provides a commonly encountered example of an extended illumination distribution. This provides a complementary example to the highly directional collimated source discussed in section 2.3.5. A uniform hemispherical source is a first approximation to the illumination from the sky. It can be shown [61] that for such a source the Lambertian reflectance map is:

$$R(\hat{\mathbf{n}}) = \frac{\rho_d L_o}{2}(1 + \hat{\mathbf{n}} \cdot \hat{\mathbf{Z}}), \tag{2.9}$$

where L_o is the uniform incident radiance provided along each direction within the hemisphere of the source. The vector $\hat{\mathbf{Z}}$ points to the zenith of the hemisphere. The vector $\hat{\mathbf{n}}$ is, as before, the local surface normal.

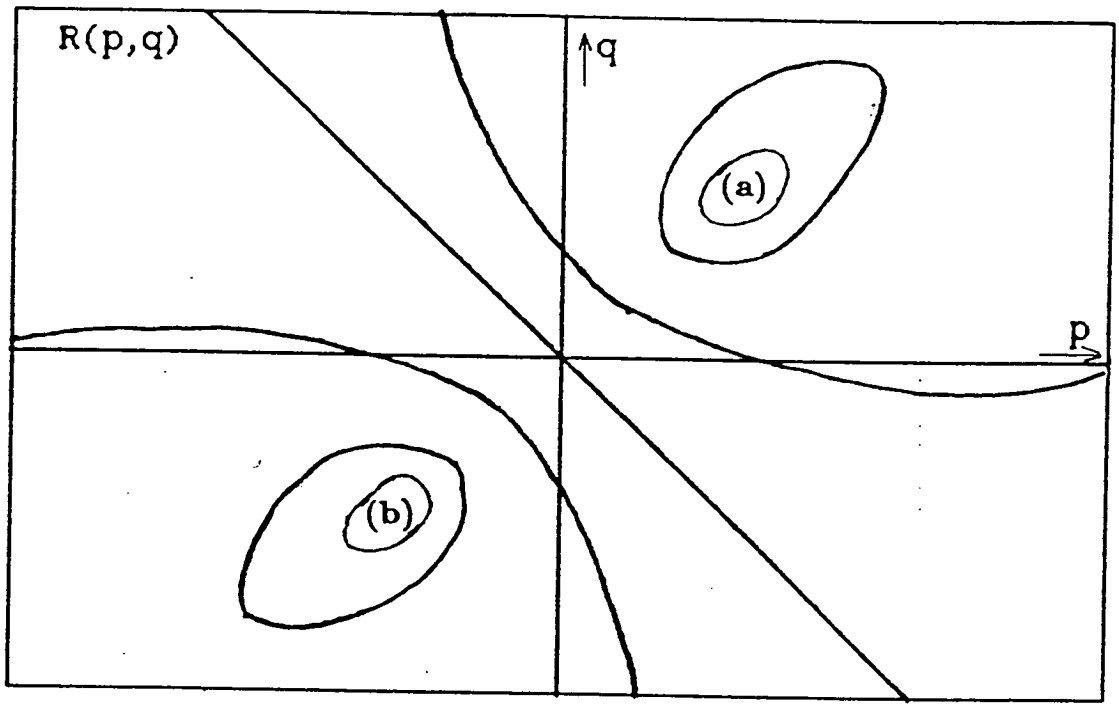


Figure 2–11: *The Lambertian reflectance map for a uniform hemispherical source. (a) marks it maximum and (b) marks its minimum.*

This reflectance map is depicted in figure 2–11. Near the maximum, the map is similar in form to (2.8)—the map for the collimated source plus ambient light. Away from the maximum this is not true. In fact R falls to zero when the surface normal points directly away from the zenith of the hemispherical source—i.e. again $R_{min} = 0$. Thus this hemispherical illumination model, has the same deficiency as the collimated source model—the ambient light due to interreflection unrealistically drops to zero. This is the case for any purely hemispherical source. Indeed, illumination engineers [56,87] who use sophisticated non-uniform hemispherical models of both overcast and clear skies also find it necessary to take into account the light scattered from below the horizon. They add a term that they call 'ground reflectance'. Figure 2–12 shows how this 'ground reflectance' arises. This term ensures that R_{min} is non-zero and again in typical cases the dynamic range of R is $3r_\rho$.

To summarise the image irradiance measured by a camera is determined by

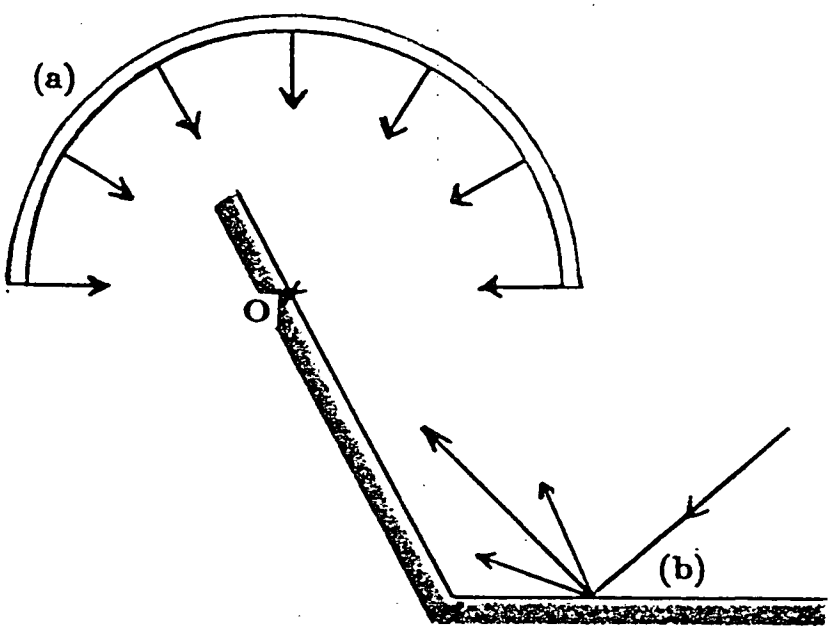


Figure 2–12: *A illumination distribution consisting of both (a) a non-uniform hemispherical source and (b) light reflected from the ground ('ground reflectance').*

local surface orientation and the reflectance map. Reflectance maps are in general difficult to analyse. Lambertian surfaces have a particularly simple BRDF and as a result Lambertian reflectance maps are easier to analyse. Consideration of the Lambertian reflectance map for both a highly directional and a widely distributed illumination indicates that it typically has a dynamic range of roughly $3r_\rho$. Where r_ρ is the dynamic range of albedo—which is discussed in section 2.4.

2.4 Physical processes of surface reflection

2.4.1 Introduction

In the sections above the reflectance map R was introduced and it was noted that R determines image irradiance and is itself determined by the illumination and the BRDF of the surface material. Although a range of realistic illumination distributions was considered, the discussion was restricted to a BRDF corresponding to a Lambertian surface: $f_{r,d}$. In this section the BRDF of real materials are examined. Of particular interest is the extent with which they depart from the form of $f_{r,d}$. Any significant departures ought to correspond to significant specular reflections if the Lambertian constraints used to detect specularities in chapter 4 are to be widely applicable. Only a partial assessment of their applicability is possible because the interaction of light with matter is only partially understood. Table 2–1 provides a summary of the various physical processes involved along with their associated BRDFs (when known). In order to make any assessment it is not sufficient to know the BRDF of each component process, in addition the following must be known:

- the surface roughness and the material constants.

- the way that the component processes combine for real materials.

Reflection	BRDF	Examples
Perfect mirror surface	$f_{r,is} = \delta(\theta_i - \theta_r)\delta(\phi_i - \phi_r \pm \pi)/[\cos\theta_i \sin\theta_i]$	polished metals (approx)
Smooth surface	$F(\theta_i, \eta) f_{r,is}$	smooth glass metals & plastic
Rough surface	$GD(\alpha)F(\psi, \eta)/[\cos\theta_i \sin\theta_i]$ or Unknown	Rough metals, MgO ceramic, paper & wood
Lambertian sub-surface	ρ_d/π	MgO powder (approx)
Real sub-surface	Unknown	fibres, wood, paints plastics & oxides

Table 2–1: *Symbols are defined in the text below.*

Before considering these aspects a summary is given for each component process. Readers not interested in these details may skip to the summaries provided in section 2.4.7 and the end of section 2.4.8.

2.4.2 Perfect mirror surface

A perfect mirror reflects each incident ray as in figure 2–2 (page 11), so that $\theta_i = \theta_r$ and so that the energy in the reflected ray is that of the incident ray. This implies [85] a BRDF of

$$f_{r,is} = \delta(\theta_i - \theta_r)\delta(\phi_i - \phi_r \pm \pi)/[\cos\theta_i \sin\theta_i]. \tag{2.10}$$

The Dirac delta functions ensure each incident ray produces a single pulse in the direction for mirror reflection. The term $[\cos\theta_i \sin\theta_i]$ stems from the definition of the BRDF in terms of solid angle. The smooth silvered surface of a mirror approximates this theoretical special case.

2.4.3 Smooth surfaces

Surface materials fall into two types: metallic (conducting) and dielectric (non-conducting). The surface reflection processes for the two classes differ but are covered by the same laws as long as the surfaces are smooth.

Dielectrics—transparent

Most readers will be familiar with the laws of reflection for a transparent dielectric such as glass. Wave theory describes how light is both reflected by and transmitted through such a medium. When rays are drawn perpendicular to the incident, reflected and transmitted waves Snell's law specifies their direction:

$$\theta_i = \theta_r; \quad \eta_i \sin\theta_i = \eta_t \sin\theta_t, \tag{2.11}$$

where all the rays lie in the same plane and θ_i,θ_r and θ_t are the angles of incidence, reflection and transmission respectively. The refractive indices of the environment and the medium are η_i and η_t respectively. In practical circumstances relevant to this thesis the environment in question is air so that $\eta_i \simeq 1$ and η_t will be simply denoted η. The proportions of energy in the incident wave that are reflected and transmitted are determined by the Fresnel function F as was discussed in section 2.2. So for a transparent dielectric the BRDF is

$$f_r = F(\theta_i, \eta) f_{r,is}, \tag{2.12}$$

where $f_{r,is}$ is as defined above in (2.10). The precise form of F depends on the polarisation of the incident wave—as shown in figure 2–13. If—as is typical—the wave is unpolarised F is determined solely by η and θ_i. In other circumstances (e.g. light incident from specular reflections) the degree of polarisation may be high but unknown and so the form of F is uncertain. The two arguments of F specify two distinct variations:

- θ_i specifies only geometrical variations.
- η is a material constant. Different materials have different values of η. However η can vary with the wavelength λ of the incident light so it is

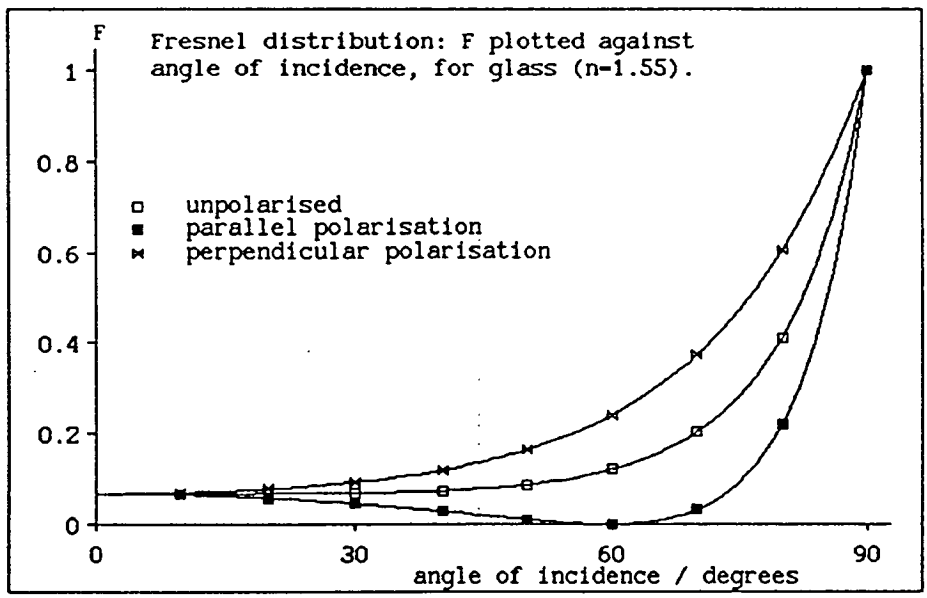


Figure 2–13: *Fresnel function for glass.*

better denoted as η_λ. For an optical glass η_λ typically has a value 1.5 to within 2% over the visible spectrum.

Later it transpires that a transparent dielectric is a simple special case: in general it is harder to isolate the geometrical and material variations of f_r.

Dielectrics—opaque

Associated with every dielectric is an absorption band—a range of wavelengths at which it absorbs electromagnetic radiation[7]. For transparent materials this absorption band lies outside the visible spectrum (330nm–770nm). Some materials (e.g. coloured glasses) have absorption bands that span only part of the visible spectrum, others are completely opaque. Equation (2.12) extends to describe absorbing dielectrics provided that a complex refractive index $\boldsymbol{\eta}$ is substituted for η: i.e.

$$f_r = F(\theta_i, \boldsymbol{\eta}) f_{r,is}, \quad \boldsymbol{\eta}_\lambda = \acute{\eta}_\lambda + \mathrm{i}k_\lambda, \tag{2.13}$$

[7] Via dipole interactions—see [54].

where $\acute{\eta}_\lambda$ and k_λ are the real and imaginary parts of η—both are wavelength dependent. k_λ is often called the extinction coefficient. The Fresnel relations in terms of $\acute{\eta}_\lambda$ and k_λ are given in appendix B. Figure 2–4 (page 13) show a typical distribution for the dielectric MgO. Note in this case two material constants are required to specify f_r. Values of k_λ are almost zero except within absorption bands where they are larger. Transparency only occurs for very small values of k_λ. For example when $k_\lambda = 0.001$ the transmissivity of a dielectric 1 mm thick is still remarkably low (0.001), [62]. Extensive tables of k_λ and $\acute{\eta}$ for real dielectric materials are given in [4].

Metals

The high conductivity of metals means that the reflection process differs from that of the dielectrics described above[8]. As a result the characteristic absorption bands of dielectrics are not present. In addition most incident light is reflected. This gives most metals a characteristic colourless silver grey appearance. Again equation (2.13) describes the BRDF incorporating a complex refractive index. Typical values of k_λ for metals are much larger than for dielectrics. This ensures that the incident light penetrates no more than a few hundred angstroms into the surface leaving little opportunity for any sub-surface scattering. In general $\acute{\eta}_\lambda$ is fairly constant over the visible spectrum. For some metals the variation is significant enough to produce a characteristic colour, e.g. copper. Tables of $\acute{\eta}_\lambda$ and k_λ for many metals are also given in [4].

2.4.4 Rough surfaces

The reflectance properties of rough surfaces (both metallic or dielectric) depend on their rms roughness σ_m : the root-mean-square deviation of the rough surface from a plane representing its mean height. For most surfaces σ_m is comparable to the wavelength λ of visible light, so that interference and diffraction occurs. When $\sigma_m/\lambda \sim 1$ there is no simple theory that predicts the form of

[8]Electrons rather than ions or atoms are involved.

the BRDF [62]. Surfaces for which $\sigma_m/\lambda \ll 1$ are known as optical surfaces. Their BRDF can occasionally be predicted by Fourier analysis of measured surface profile data [35]. However no simple generalisations can be made about these BRDFs. Surfaces where $\sigma_m/\lambda \to 0$ are effectively smooth and the laws described in the previous section apply.

For surfaces where $\sigma_m/\lambda \gg 1$, e.g. rough paper, useful predictions of the BRDF can be made. At a microscopic scale these surfaces can be modelled as many small randomly oriented micro-facets. Adjacent incident rays are reflected by adjacent facets in different directions. This surface scattering results in a BRDF that is no longer a single pulse, but is spread across a finite solid angle as depicted in figure 2–14. The precise distribution depends not only on the material constant η but also on the statistical properties of the surface. All of the various models used to predict the BRDF of such surfaces, [11,104, 105], assume that the statistical properties are directionally uniform[9] over the area of interest. Under this assumption the BRDF is characterised by a *facet distribution function* $D(\alpha)$—where α is an angle that specifies the direction of the spread relative to the direction of mirror reflection. As figure 2–14 (b) shows $\cos\alpha \equiv \hat{\mathbf{h}} \cdot \hat{\mathbf{n}}$, where $\hat{\mathbf{n}}$ is the direction of the (macroscopic) surface normal and where $\hat{\mathbf{h}}$ is the direction that bisects the vector of incidence, $\hat{\mathbf{L}}$ and the view vector, $\hat{\mathbf{V}}$, i.e.:

$$\hat{\mathbf{h}} = \frac{\hat{\mathbf{V}} + \hat{\mathbf{L}}}{|\hat{\mathbf{V}} + \hat{\mathbf{L}}|}. \tag{2.14}$$

The function D describes the relative amount of micro-facet area oriented in any given direction, or equivalently the probability density of a facet being normal to that direction. Beckmann and Spizzichino [11] show[10] if the surface height

---

[9]This is not true for surface roughness that is described by a periodic function. Such surface are not considered here.

[10]In fact [11] page 193 gives a 1-D version of D—the extension for 2-D surfaces is straight forward and Cook and Torrance [31] provide it.

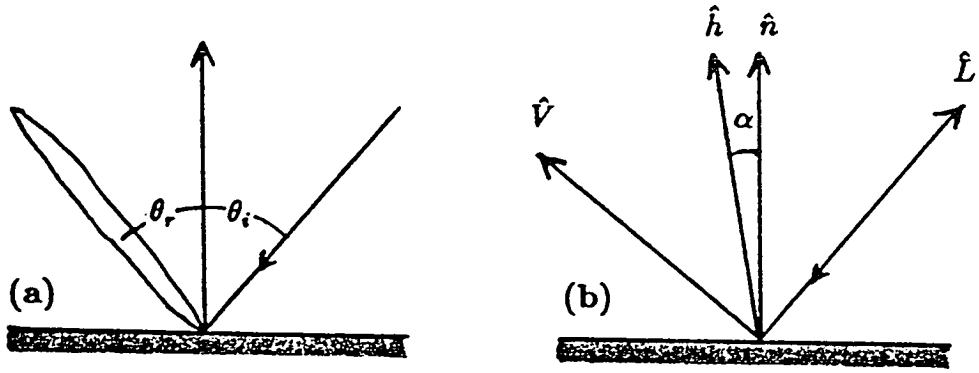


Figure 2–14: *For a rough surface specular reflected light forms a finite beam, spread around the direction of mirror reflection (a). (b) shows vector* $\hat{\mathrm{h}}$ *that bisects the view vector,* $\hat{\mathrm{V}}$ *and the vector of incidence,* $\hat{\mathrm{L}}$.

deviation from the mean plane is in a gaussian distribution then

$$D(\alpha) = \frac{1}{\pi m^2 \cos^4 \alpha} \exp\left(-[(\tan \alpha)/m]^2\right), \tag{2.15}$$

where m is the rms slope distribution defined:

$$m = \frac{T^2}{2\sigma_m^2}. \tag{2.16}$$

Here the surface roughness is characterised by two surface constants: T the autocorrelation length[11] and the rms roughness σ_m. Although the gaussian assumption seems plausible, it is not known to which materials it is applicable. Other ad-hoc functions have been used in Computer Graphics to render artificial images with varying degrees of success [20,105,90]. However no facet distribution function is capable of fully accounting for *off-specular glints*. These are specularities occurring at large angles of incidence for which the peak in their BRDFs is not along the direction of mirror reflection (i.e. at $\alpha \neq 0$), as shown in figure 2–28 on page 47. Torrance and Sparrow [104] provide a model

[11] the distance over which the correlation between surface heights drops to e^{-1}.

of the masking and shadowing processes between adjacent micro-facets that predicts off-specular glints. This model incorporates a *geometrical attenuation function* G into the expression for the BRDF:

$$G = \min\left(1.0, \frac{2(\hat{\mathbf{n}}\cdot\hat{\mathbf{h}})(\hat{\mathbf{n}}\cdot\hat{\mathbf{V}})}{\hat{\mathbf{V}}\cdot\hat{\mathbf{h}}}, \frac{2(\hat{\mathbf{n}}\cdot\hat{\mathbf{h}})(\hat{\mathbf{n}}\cdot\hat{\mathbf{L}})}{\hat{\mathbf{V}}\cdot\hat{\mathbf{n}}}\right), \tag{2.17}$$

where all vectors are as were defined for figure 2–14 (b). In this case the final form of the BRDF is:

$$f_r = \frac{GD(\alpha)F(\psi,\eta)}{\cos\theta_i \sin\theta_i}, \tag{2.18}$$

where F is again the Fresnel function and $\cos\psi = \hat{\mathbf{V}}\cdot\hat{\mathbf{h}}$.

2.4.5 Sub-surface scattering

In addition to the various types of surface reflection described above, any light that penetrates into the volume of a material can be scattered back into the environment. Figure 3–7 on page 71 depicts this process. Sub-surface scattering only occurs in significant amounts for inhomogeneous dielectrics, e.g. paint, cloth, wood and oxides. Inhomogeneities (pigment particles, fibres or grains) within such materials scatter and absorb light. Lambertian surfaces provide a simple model of the BRDF in this case:

$$f_r = \frac{\rho_d}{\pi}. \tag{2.19}$$

This distribution is constant for all orientations and only varies when surface albedo ρ_d changes. It corresponds to the theoretical situation in which a volume of material absorbs radiation and maintains radiative equilibrium by re-emitting isotropically—as a black body would. This model greatly simplifies the treatment of many problems in radiative transfer and has been widely applied both in engineering [62] and Computer Graphics [20,90,31,29,42]. However real sub-surface scattering is not perfectly Lambertian. In general it is so complicated that no adequate general model exists to describe it. Below a few of the factors involved are outlined—more details can be found in Hottel [62].

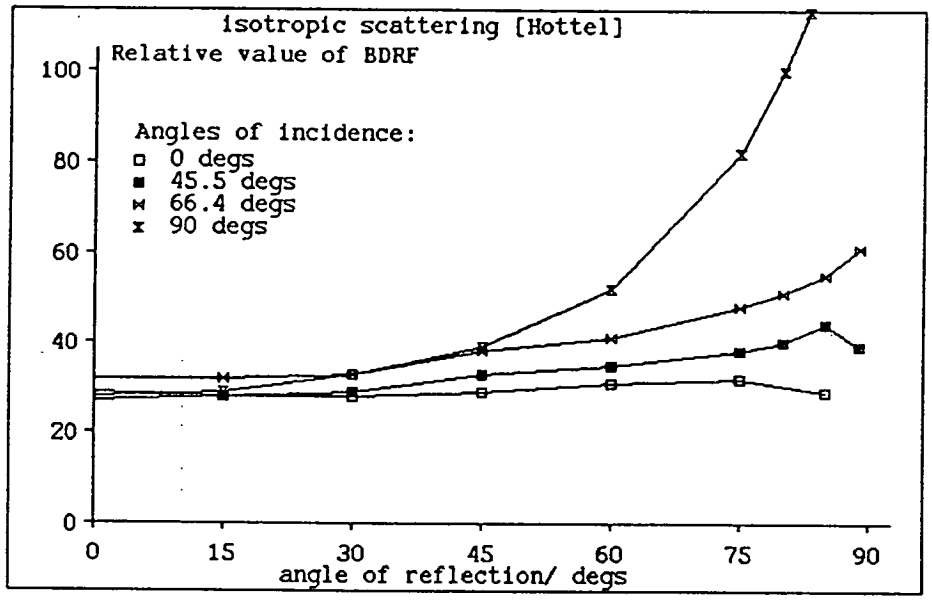


Figure 2–15: *Theoretical prediction of BRDF for an isotropic scatterer. A Lambertian surface would produce a horizontal line for all angles of incidence. The precise form of the curves depends on the value of scattering coefficients defined for the medium.*

Even for an isotropic scatter the BRDF need not be Lambertian. Typically the deviation from $f_{r,d}$ increases with angle of incidence—see figure 2–15. To obtain the BRDF of a specific isotropic scattering medium the size, shape, refractive index and cross-section of the scattering particles must be known. In addition scattering theories make use of the coefficients of absorption, scatter and extinction of the material. The form of the BRDFs depends greatly on the size r of the particles [107]. For small particles Rayleigh scattering occurs [62]. For particles in the size range $0.6 < 2\pi r/\lambda < 5$ the Mie equations [62,26,27] are applicable if their size is roughly spherical. For other shapes rigorous solutions of Maxwell's Equations are required but are usually intractable. Further complications are introduced by interference effects that occur if the particles in the medium are separated by less than about $3r$.

Anisotropic scattering is poorly understood. Radiation-transfer engineers have predicted the BRDFs for some anisotropic scatters using multi-flux meth-

ods [68,84]. These methods describe the radiation transfer process at each point within the medium in terms of two or more first order differential equations. The simultaneous solution of these equations requires specification of boundary conditions. Solutions have been obtained for volumes of scatterers bounded between parallel planes. Kubelka-Munk theory (a two-flux method) provides no prediction of directional flux distribution. The six-flux methods do—but only for specific cases. Multi-flux methods suggest no useful general forms for the BRDFs of anisotropic scatters.

2.4.6 Combining component BRDFs

From the discussion above it is clear that not all the component BRDFs listed in table 2–1 are well understood. However it is still useful to consider how those that are understood can combine in real materials. First consider a smooth surface of a dielectric—possibly porcelain-enamel or paint. In this case only two components are involved:

- for the smooth[12] surface: $f_{r,sur} = F(\theta_i,\eta)\delta(\theta_i-\theta_r)\delta(\phi_i-\phi_r\pm\pi)/[\cos\theta_i\sin\theta_i]$. This component describes the reflection at the *surface*.

- for the sub-surface scattering: $f_{r,vol}$—which for the moment will be considered Lambertian ($f_{r,vol} = \rho_d/\pi$). This component describes the reflection by the *volume* under the surface.

Making a further assumption (relaxed somewhat later) that the combined BRDF, f_r is a linear combination of $f_{r,sur}$ and $f_{r,vol}$ then

$$f_r = a\frac{F(\theta_i,\eta)\delta(\theta_i-\theta_r)\delta(\phi_i-\phi_r\pm\pi)}{\cos\theta_i\sin\theta_i} + b\frac{\rho_d}{\pi}, \qquad (2.20)$$

where a and b are constant over all orientations $(\theta_i,\phi_i;\theta_r,\phi_r)$. In this case f_r is constant for all $(\theta_i,\phi_i;\theta_r,\phi_r)$ except at directions of mirror reflection:

---

[12]Rough surfaces that produce specularities spread over larger solid angles will be considered later.

$(\theta_i, \phi_i; \theta_i, \phi_i \pm \pi)$ where specular pulses occur. The data obtained for porcelain enamel by Hunter [63] and reproduced in figure 2–17 (page 42 approximates this case. [As in all the data-graphs that follow, figure 2–17 depicts the relative value of the BRDF in the plane of incidence: $f_r(\theta_i, \phi_i; \theta_r, \phi_i \pm \pi)$ for fixed θ_i over a range of θ_r.]

In terms of the reflectance map R:

- $f_{r,sur}$ may result in sharp specular pulses or the superposition of many pulses if the light sources are extended and surfaces gently curving.

- $f_{r,vol}$—the constant component—results in slowly varying features corresponding to the slowly varying incident radiance distribution L_i—as discussed in section 2.3. Discontinuities are only created across material boundaries where ρ_d and b may change.

The basic premise on which the specularity detector in chapter 4 is based is that specular features in R are readily distinguishable by the rapid variations and large magnitudes induced by $f_{r,sur}$. However it is impractical to do this by identifying these features using specific instantiations of the model (2.20) because, among other things, the constants, a, b, ρ_d and η are unknown. Instead constraints are formulated, which when broken indicate specularities. The basic formulation is as follows:

1. R is assumed to be due solely to the Lambertian component $f_{r,vol}$ at all non-specular orientations.

2. Consequently over these orientations the magnitude and derivatives of R depend only on L_i, b and ρ_d and thus are constrained. Tests are proposed to detect specularities in unconstrained regions.

3. Where the test requires a choice of an upper bound on a physical constant such as ρ_d it is made conservatively to ensure that:

    - a wide range of different materials can be encompassed.

- the constraints may still apply when $f_{r,vol}$ deviates from a constant. As occurs even for the porcelain enamel—see figure 2–18 (page 42).

Precisely which materials are encompassed by the constraints is uncertain. Below the shortcomings of the model (2.20) are discussed. Some are not serious: so the constraints may well apply to a wider range of materials than the model strictly suggests.

Deviation from Lambertian : $f_{r,vol}$ isotropic

Sub-surface scattering for real inhomogeneous dielectrics deviates from the Lambertian distribution. Measured data (figures 2–19–2–22, pages 43–44) and theoretical models (figure 2–15, page 35) suggest that for isotropic scattering where $\theta_i < 70°$, $f_{r,vol}$ varies by less than a factor of 2. Such variations are unlikely to be confused as specular features. However for $\theta_i > 70°$ figure 2–15 shows that $f_{r,vol}$ can become very large as $\theta_r \rightarrow 90°$ and may result in specular-like features. However it is precisely at these orientations (glancing angles) that any smooth surface component ($f_{r,sur}$) becomes large. This is a consequence of the Fresnel function—see figure 2–4 (page 13). So it is uncertain that the data obtained for pressed barium sulphate—figure 2–23 (page 45)—becomes large at $\theta_r = 70°$ due to an increase of either or both of $f_{r,sur}$ and $f_{r,vol}$. In any case the specular-like feature created is likely to be detected rightly or wrongly[13].

Non-linear combinations

In fact as glancing angles are approached $f_{r,sur}$ reflects more of the incident light and thus less enters the surface. Consequently the fraction available for sub-surface scattering b is not a constant—but is better expressed as a function $b(f_{r,sur})$. So near glancing angles f_r should be considered a non-linear combination of $f_{r,sur}$ and $f_{r,vol}$—at odds with the model (2.20). However for smooth surfaces significant non-linearity only occurs at orientations where $f_{r,sur}$ dominates the behaviour of f_r and so should have negligible effect.

---

[13]The chromatic methods reviewed in chapter 3 could in principle determine whether $f_{r,sur}$ of $f_{r,vol}$ is primarily responsible.

Deviation from Lambertian: $f_{r,vol}$ anisotropic

For anisotropic sub-surface scattering the deviation of $f_{r,vol}$ from a constant cannot be easily predicted. One example where the deviation appears to be a factor of 3 is shown in figure 2–24 (page 45). If this data is typical then the constraints may well extend to cover some anisotropic scatters.

Smooth metals

The model (2.20) extends to cover smooth metals by setting $b = 0$. In terms of the constraints: the constant (zero) value of the sub-scattering component

- ensures that constraints based on the derivatives of R still apply.
- disrupts constraints based on the magnitude range of R.

Rough surfaces—metals and dielectrics

For rough surfaces the specular component $f_{r,sur}$ is distributed over a finite solid angle. Even for a fairly smooth surface like the porcelain enamel, this spread is measurable—see figure 2–17 (page 42). In fact the quality of mirrors is assessed by the narrowness of the spread of their BRDFs. However as long as $f_{r,sur}$ retains a sharp peak flanked by steep sides it will still create specular features that violate the Lambertian constraints. For a wide-spread $f_{r,sur}$—as in figure 2–25 (page 46)—the surface scattering is so diffuse that it is similar to that produced by sub-surface scattering. This occurs for the rough surfaces investigated by Torrance and Sparrow—figures 2–26–2–28 (pages 46–47) provide more examples. Notice how the off-specular glints appear as θ_i increases. These examples are for a metal where the sub-surface scattering is negligible: in effect any prominent off-specular glint is likely to violate the Lambertian constraints while the remainder of the spread is not. Essentially the same is true for a dielectric—see figures 2–29–2–32 (pages 48–49) except in this case: only at off-specular glints is it possible to distinguish $f_{r,sur}$ from $f_{r,vol}$.

The data for some other rough surfaces shown in figures 2–33–2–37 (pages 50–52) indicate some more BRDFs that the specularity detector is likely to encounter.

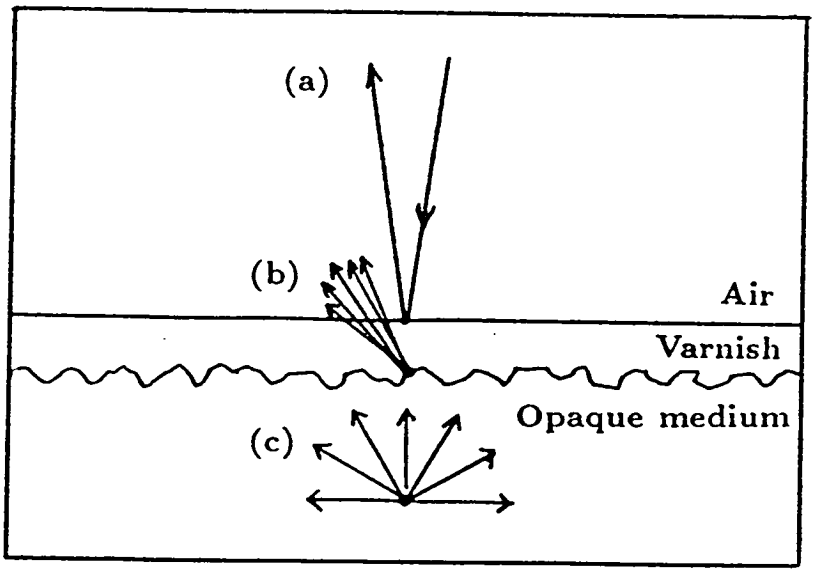


Figure 2–16: *A surface coated with a layer of varnish may have three components of reflection: (a) at the surface of the varnish, (b) at the surface of the opaque medium and (c) sub-surface scattering by the medium.*

Varnish layers

Many common materials, such as varnished wood are coated by a thin dielectric layer. As figure 2–16 shows, more than two processes may be involved in surface reflection in this case. However the model (2.20) can still extend to cover this case[14]. The specular component is considered to be a combination of two sub-components—one for the smooth reflection at the surface of the varnish and one for reflection at the top of the opaque surface. For angles of incidence up to about 70° only about 5% of the incident light is specularly reflected at the surface of the varnish[15]. In effect this superimposes a specular pulse onto the existing $f_{r,sur}$ of the lower surface. At higher angles of incidence the smooth surface reflection of the varnish begins to dominate the form of $f_{r,sur}$. In either

[14]For a detailed analysis of dielectric overlayers see Elson [34].

[15]In accordance with Fresnels law—see figure 2–13 (page 30).

case the effect of the varnish layer is to make it more likely that the surface will violate the Lambertian constraints.

2.4.7 Summary: aptness of the Lambertian model

From the discussion above the following important conclusions can be made:

- Surfaces that reflect light solely by isotropic sub-surface scattering are not Lambertian.

- Theory and experimental data indicate that in most circumstances the Lambertian model provides a good approximation to such surfaces. In all but a few special cases, e.g. at glancing angles, the model predicts image irradiance to within a factor of 2. Proviso: existing theory and data although abundant is incomplete.

- The same approximation holds for specular materials at non-specular orientations. Note figures 2–17, 2–23, 2–27, 2–28, 2–32 and 2–36.

- Off-specular glints, of the type shown in figures 2–23, 2–27, 2–28, 2–31, 2–32 and 2–34. should not be used by the shape inference method described in chapter 6.

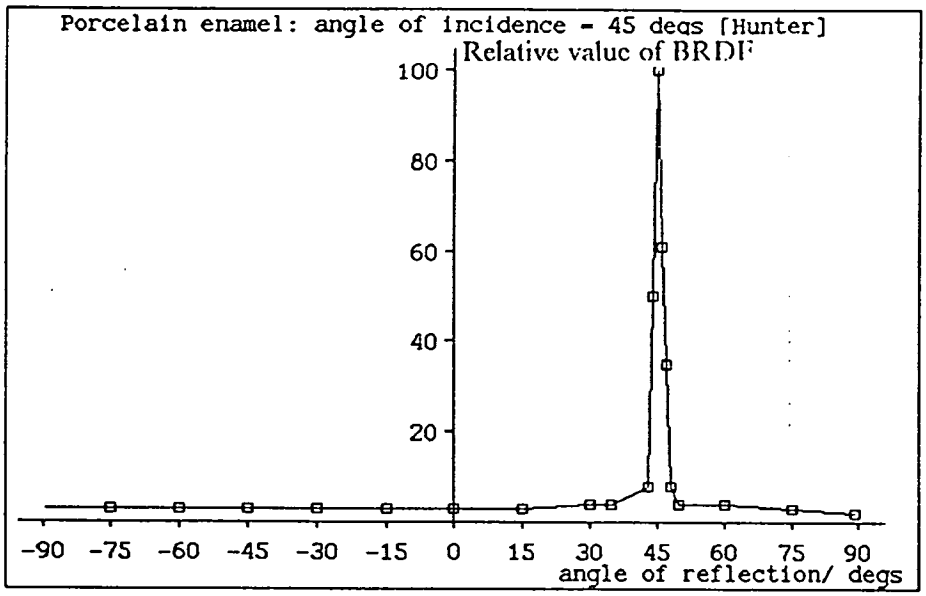


Figure 2–17: *Porcelain enamel:* $\theta_i = 45°$ *[63].*

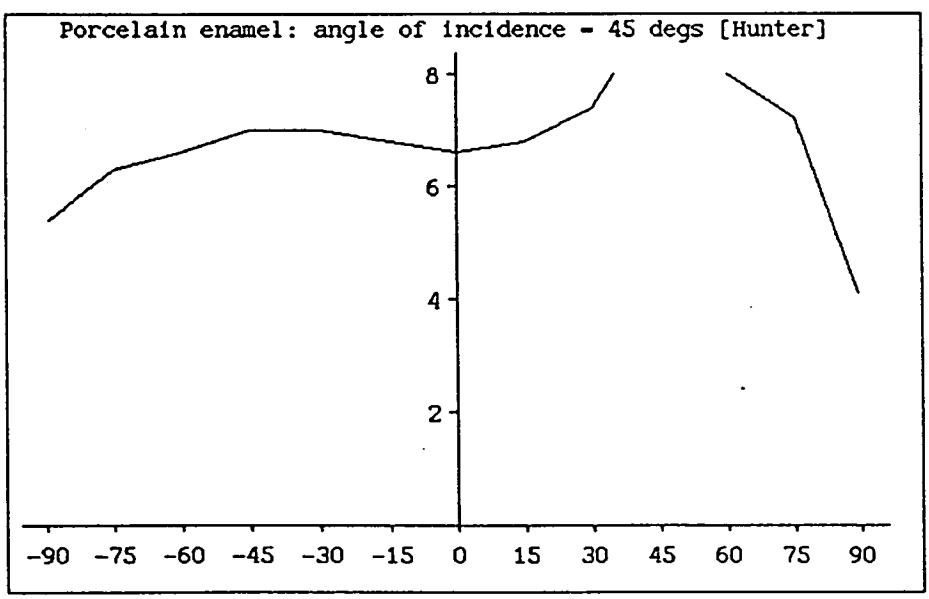


Figure 2–18: *Detail of porcelain enamel:* $\theta_i = 45°$ *[63].*

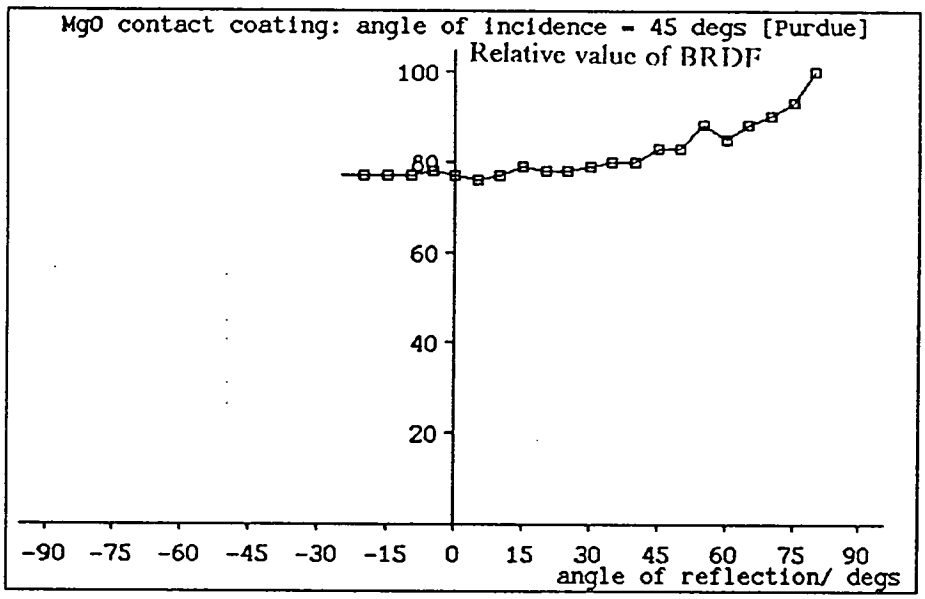


Figure 2–19: *MgO contact coating:* $\theta_i = 45°$ *[94].*

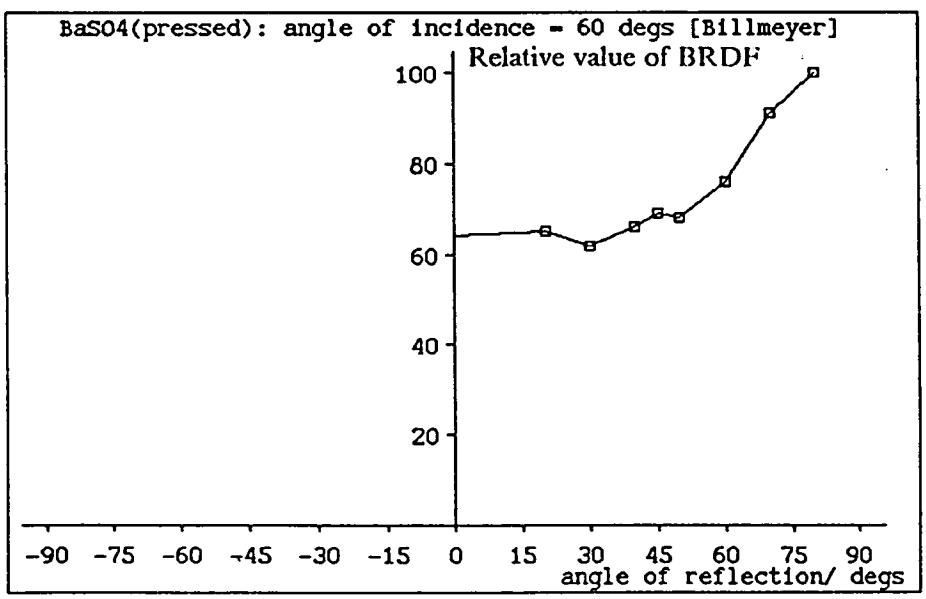


Figure 2–20: *Barium sulphate (pressed):* $\theta_i = 60°$ *[12].*

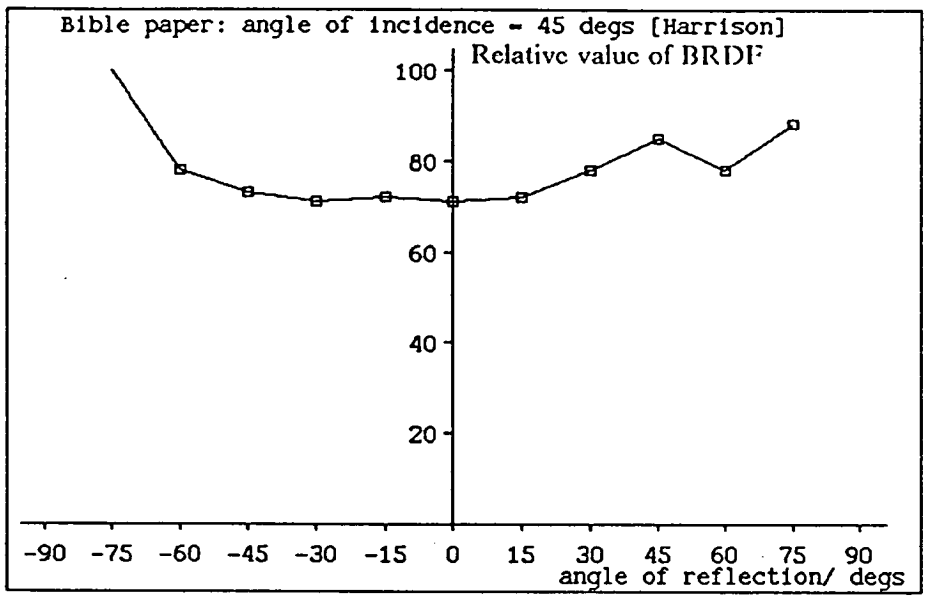


Figure 2–21: *Bible paper:* $\theta_i = 45°$ *[52].*

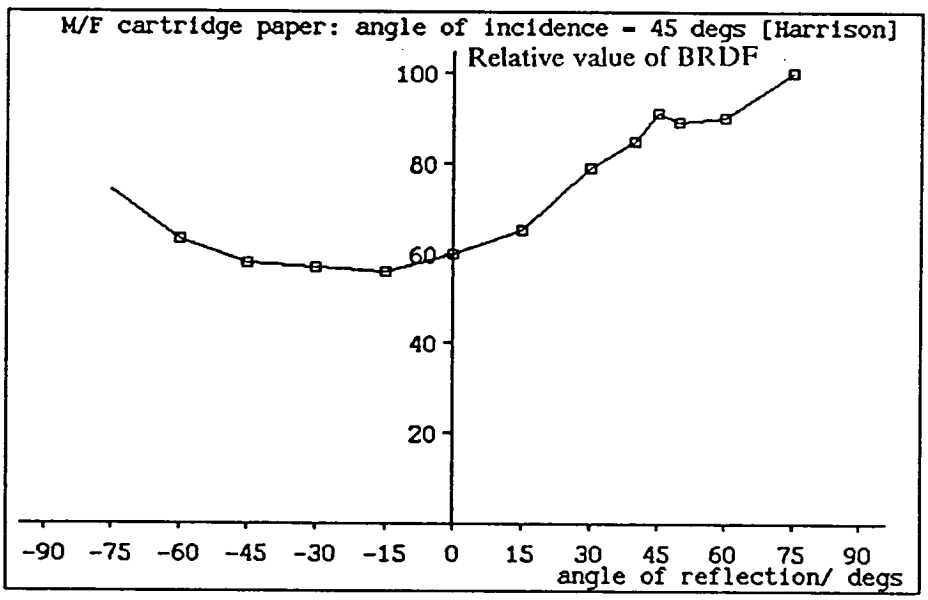


Figure 2–22: *M/F cartridge paper:* $\theta_i = 45°$ *[52].*

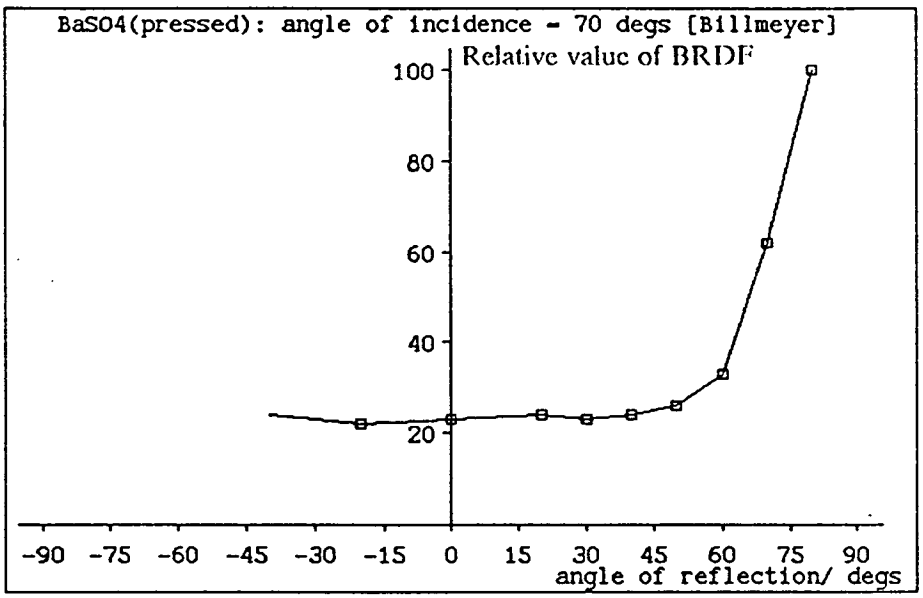


Figure 2–23: *Barium sulphate (pressed):* $\theta_i = 70°$ *[12].*

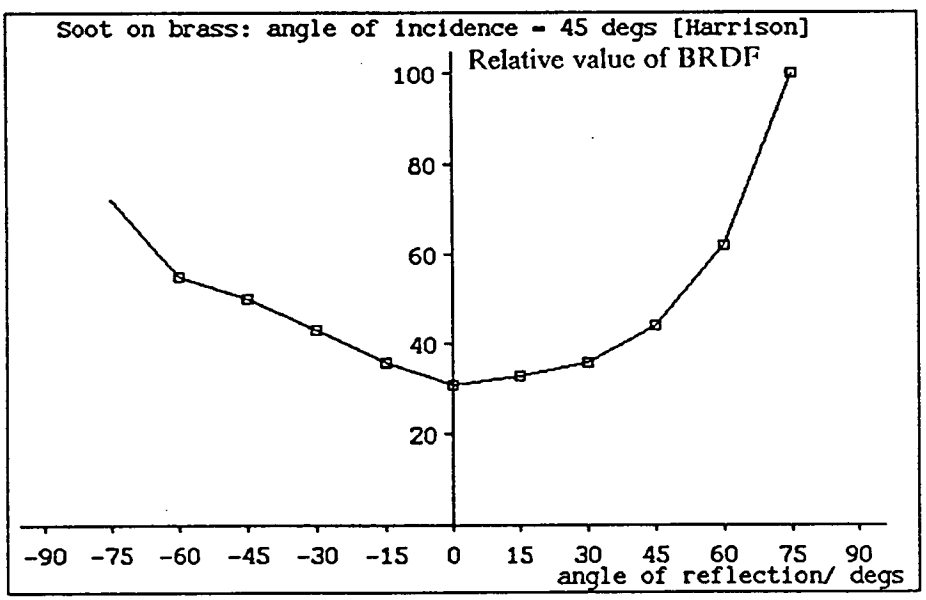


Figure 2–24: *Soot on brass:* $\theta_i = 45°$ *[52].*

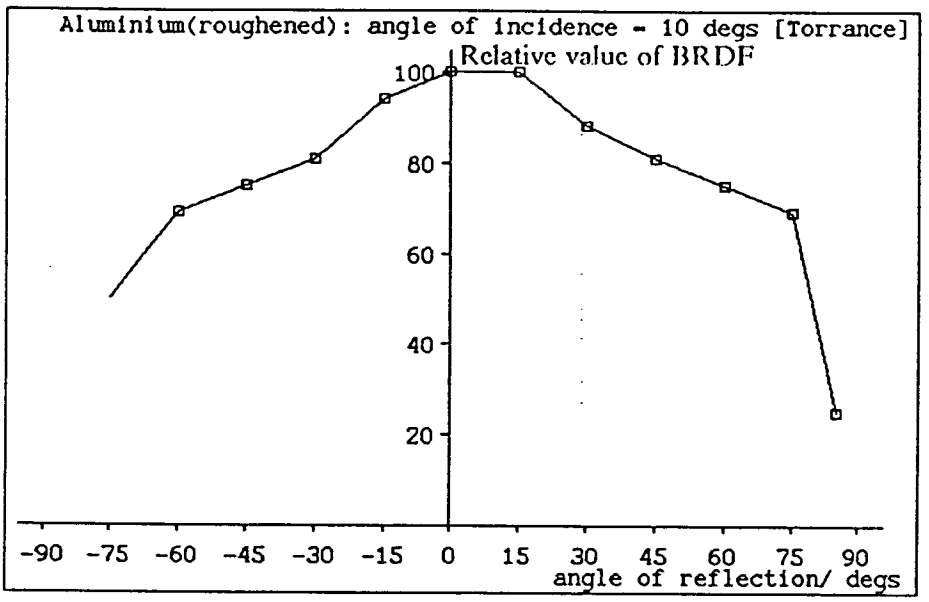


Figure 2–25: *Aluminium (roughened):* $\theta_i = 10°$ *[104].*

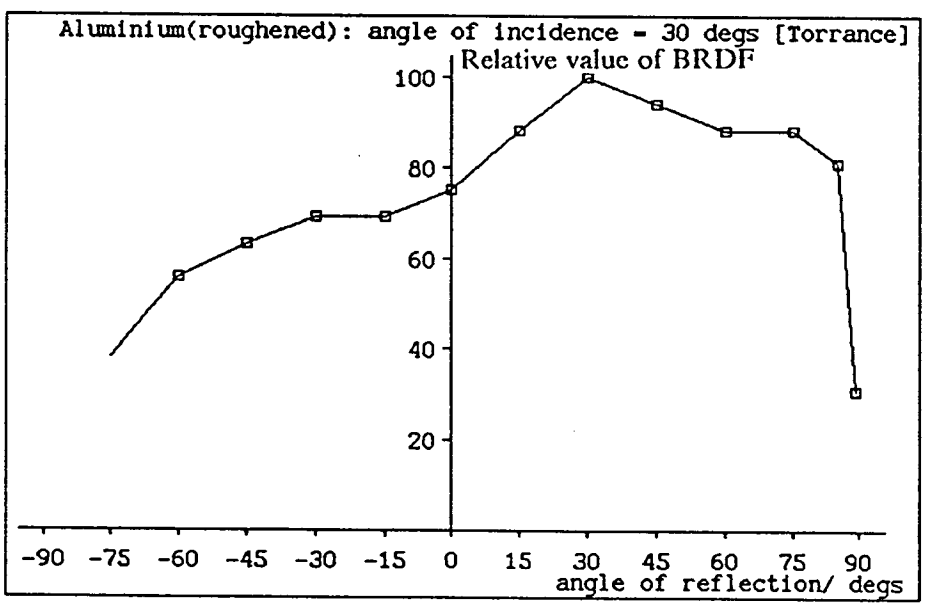


Figure 2–26: *Aluminium (roughened):* $\theta_i = 30°$ *[104].*

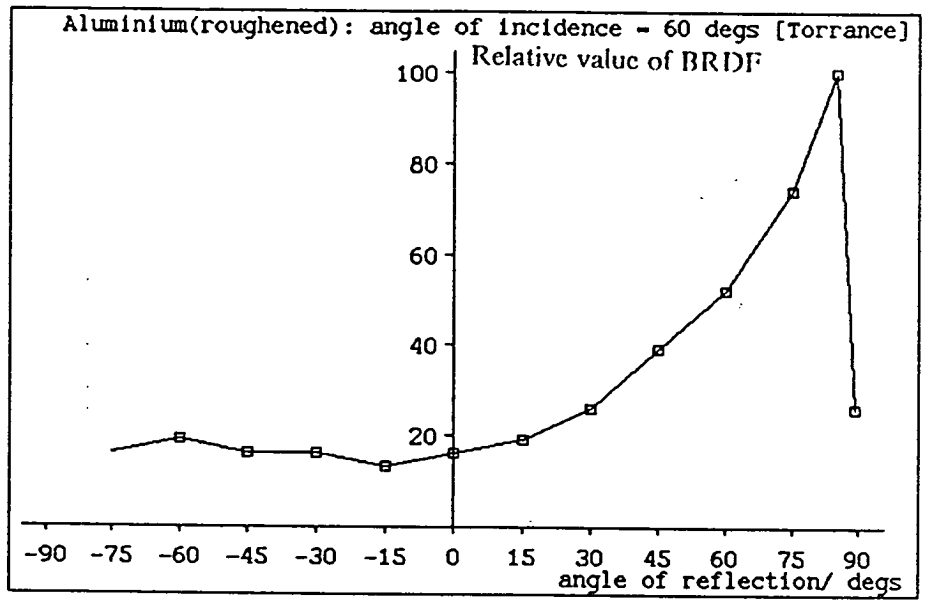


Figure 2–27: *Aluminium (roughened):* $\theta_i = 60°$ *[104].*

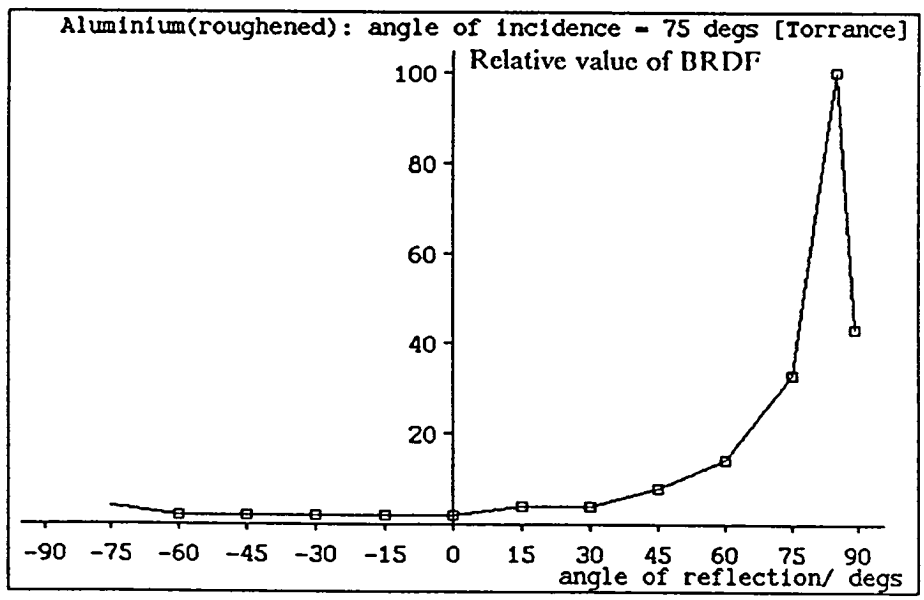


Figure 2–28: *Aluminium (roughened):* $\theta_i = 75°$ *[104].*

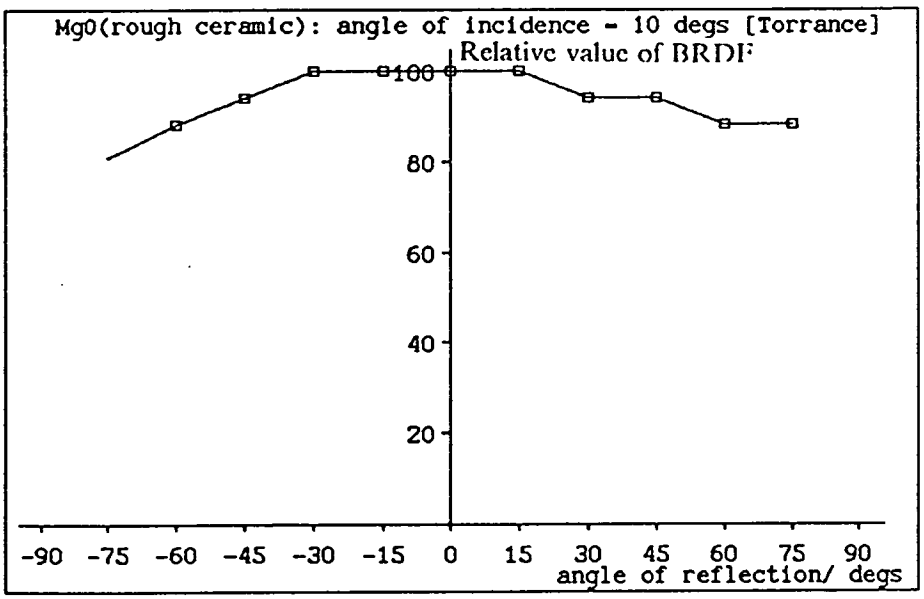


Figure 2–29: *Magnesium oxide(roughened):* $\theta_i = 10^o$ *[104].*

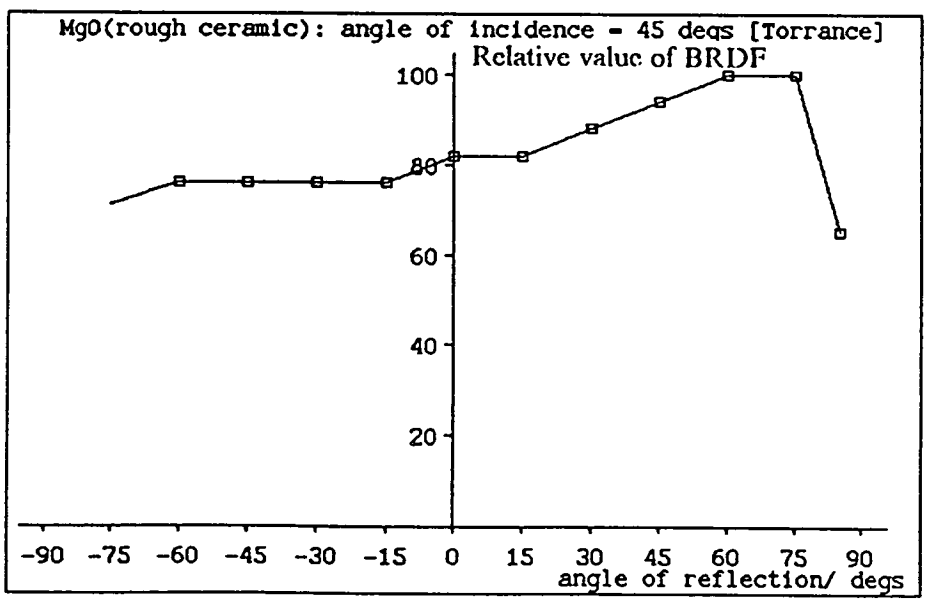


Figure 2–30: *Magnesium oxide(roughened):* $\theta_i = 45^o$ *[104].*

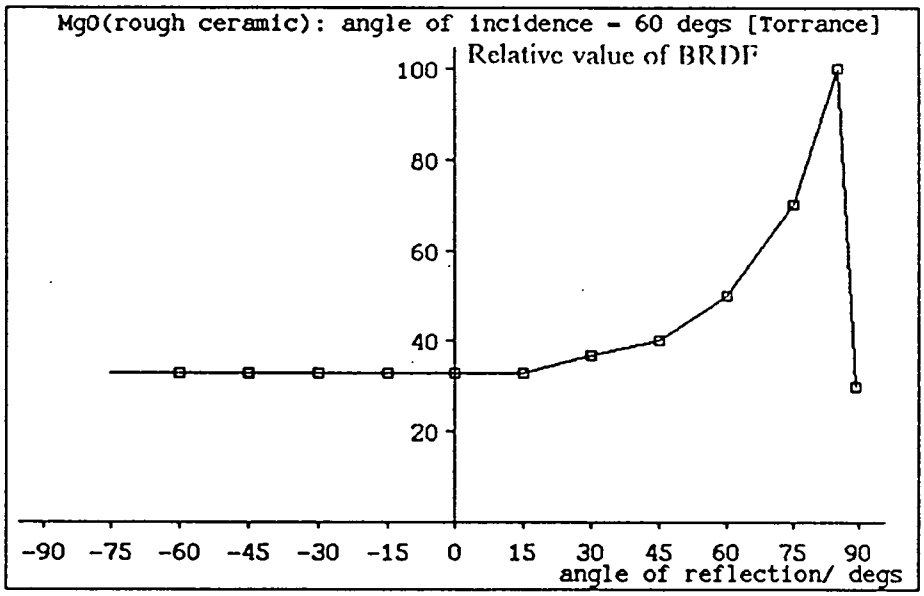


Figure 2–31: *Magnesium oxide(roughened): $\theta_i = 60°$ [104].*

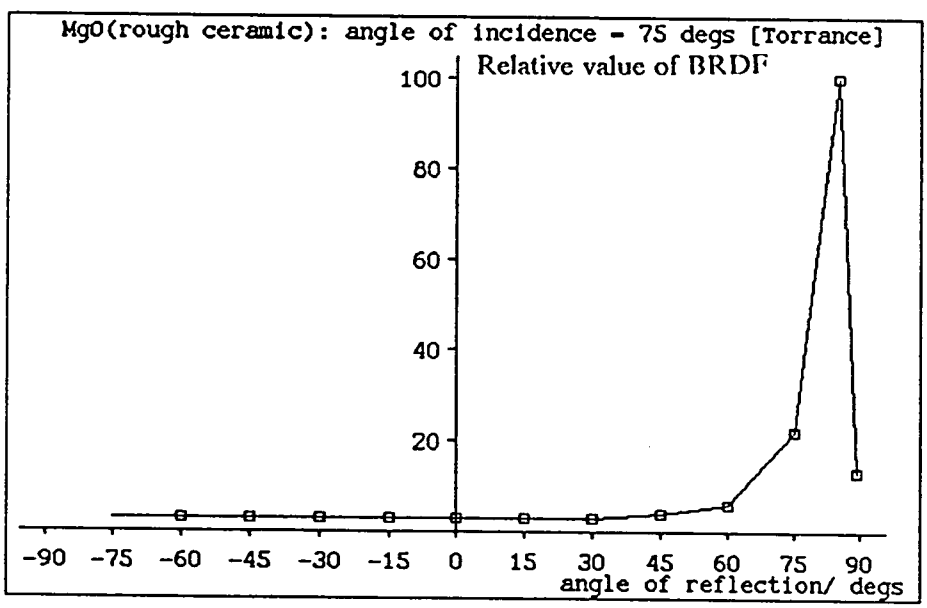


Figure 2–32: *Magnesium oxide(roughened): $\theta_i = 75°$ [104].*

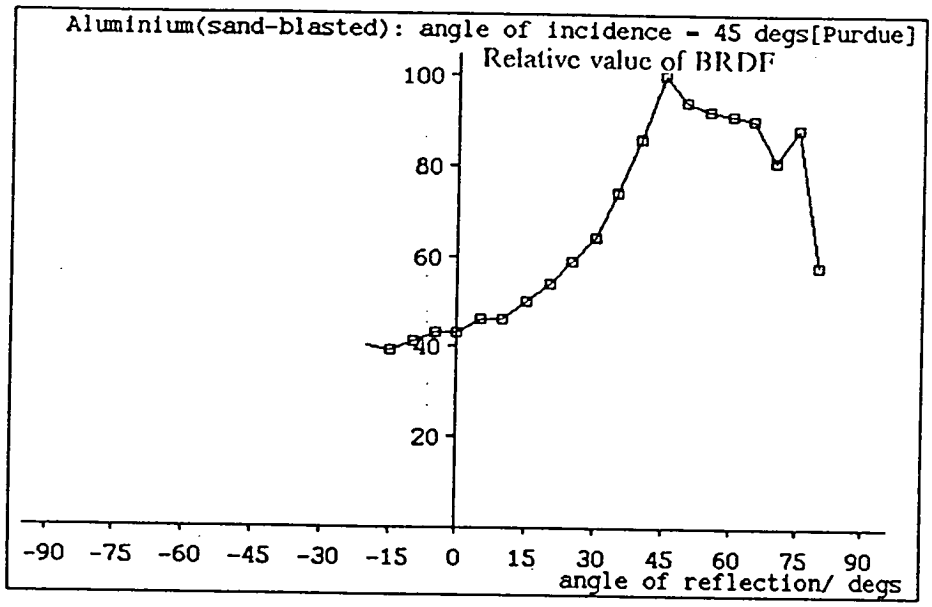


Figure 2–33: *Aluminium (sand-blasted):* $\theta_i = 45°$ *[94].*

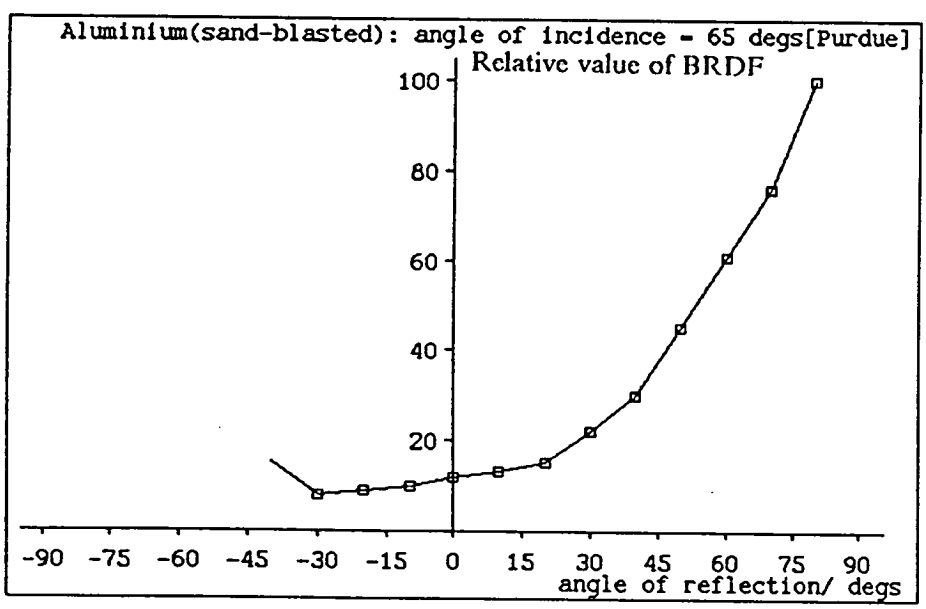


Figure 2–34: *Aluminium (sand-blasted):* $\theta_i = 65°$ *[94].*

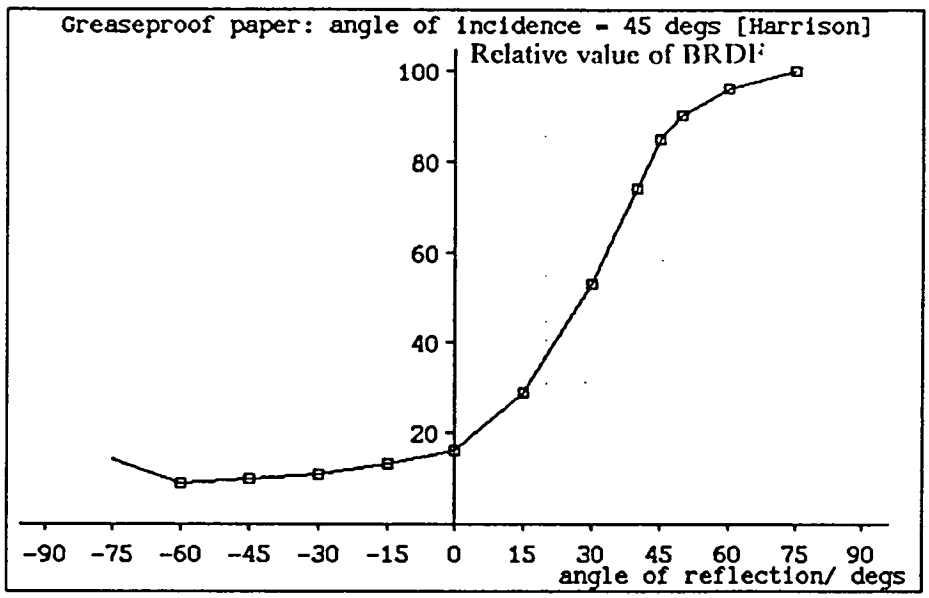


Figure 2–35: *Grease-proof paper:* $\theta_i = 45°$ *[52].*

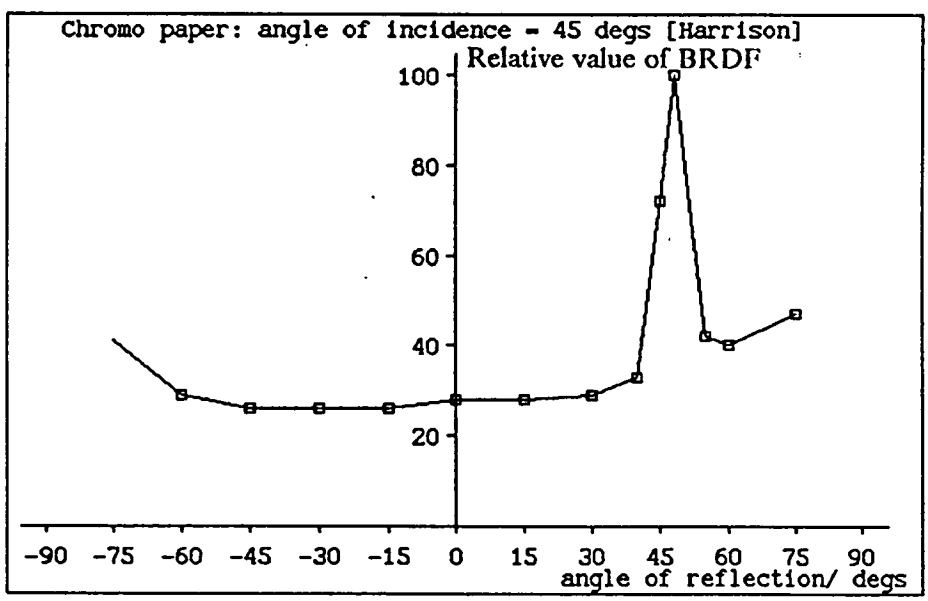


Figure 2–36: *Chromo paper:* $\theta_i = 45°$ *[52].*

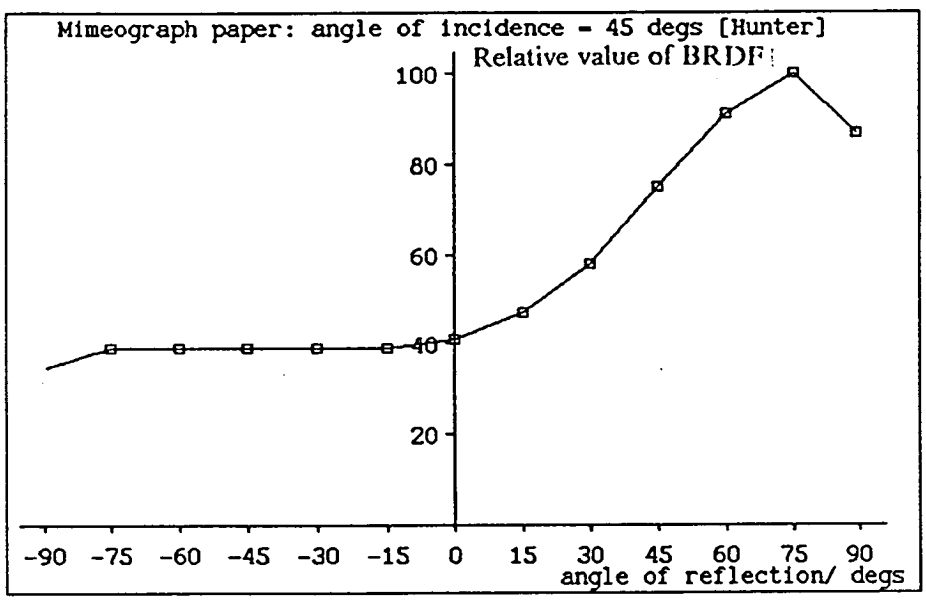


Figure 2–37: *Mimeograph paper:* $\theta_i = 45°$ *[63].*

2.4.8 Constraint on albedo

In chapter 4 the retinex-based test for specularities is introduced: it assumes a dynamic range for the BRDF of a Lambertian surface $r_\rho \leq 10$. For real surfaces encompassed by the model in equation (2.20) this test should still apply as long as effective albedo, $b\rho_d$ varies by less than a factor of 10 within the scene. Below the validity of this assumption is discussed.

Metals

Polished metals form mirror images of the surfaces that surround them. Thus the apparent effective albedo of a point on a metal surface is that of the point in the scene which it is reflecting. This is true even though b and thus effective albedo, $b\rho_d$, is almost zero for most metals. So as long as the point in the scene that is being reflected has an effective albedo within the 10:1 range then the metal also appears to do so. However, if the metal surface is reflecting a light source its effective albedo may appear to exceed this range. In this case the specularity so formed can be detected. Effectively the same argument holds for rough metal surfaces.

Dielectrics

For dielectrics a value of r_ρ can be estimated by considering the range of effective albedo: $b\rho_d$ for real surfaces. The maximum and minimum values of this range define r_ρ. For this purpose measurements of $b\rho_d$ are considered below.

Ideally if a dielectric conforms to the model in equation (2.20) its effective albedo can be directly estimated from its BRDF measured at any non-specular angle. However the angular variation of $f_{r,vol}$ (the sub-surface scattering component) makes this impractical for real dielectrics. In this case an estimate of effective albedo is provided by the ratio of flux reflected into the hemisphere above the surface to that within a beam at normal incidence. This measure is correctly termed normal *directional-hemispherical reflectance* [85] and denoted $\rho(0; 2\pi)$. This measure has two advantages[16]:

---

[16]And an obvious disadvantage: it measures $f_{r,sur}$—the specular component as well

1. it is an average over all angle of reflection for a fixed direction of incidence and so is insensitive to angular variation in $f_{r,vol}$.

2. its value has been measured for many materials.

Coblentz [28] was first to measure $\rho(0;2\pi)$. He invented the Coblentz Hemisphere to do so. The values that he obtained for a large variety of materials are summarised in [55]. However these values are for light of a single wavelength and are not suitable for estimating r_ρ. For this purpose values of $\rho(0;2\pi)$ obtained for 'white' light[17] are necessary. Table 2–2 summarises the ranges of $\rho(0;2\pi)$ measured for a variety of materials by various researchers. [Note: extensive data obtained by Krinov [74] and Gordon [43,44,45,46] are not directly applicable as they are *hemispherical-directional reflectances.*] Taking the maximum and minimum values of $\rho(0;2\pi)$ in the table provides a very large dynamic range: $r_\rho = 98\%/0.3\% \sim 3000$. However if some of the materials such as chemical powders, carbon black and black velvet are discounted upon the basis that they are very rarely encounted then a much lower value prevails: $r_\rho = 90\%/5\% \sim 18$. Using this value the retinex-based test ought to work for a large variety of scenes. The results shown in figures 4–13–4–19 (pages 110–116) demonstrate that it is safe to use a lower value, $r_\rho = 10$ in restricted industrial scenes.

Conclusion: An ordinary spread of materials has (effective) albedo in a 10:1 range—which provides a useful constraint for the Lambertian model.

---

as $f_{r,vol}$. However $f_{r,sur}$ is typically at minimum for normal incidence—see figure 2–4 (page 13)—so its contribution should be small.

[17]'White' light contains wavelengths from all parts of the visible spectrum. The retinex-based test is suitable only for 'white' light—as is provided by the sun, the sky or a tungsten lamp. Variation of r_ρ in general are much greater for narrow wave-bands.

Material	$\rho(o; 2\pi)$	References
Pure powders (e.g. MgCO3)	98–80%	[38]
Sugar, Flour, Talc., Starch	90–80%	[38,55]
White paint, papers, cloth	80–60%	[38,55,94,50]
Coloured pigments:	80–8%	[81]
..Chrome yellow	80%	
..Ultramarine	8%	
Building Materials:	90–9%	[38,55]
..Clays	90–11%	
..Concrete	35–9%	
..Slates	20–10%	
..Roofing Materials	78–11%	
..Bricks	64–11%	
..Pine Wood	40%	
Peach, Pear (ripe or green)	40–20%	[55]
Coloured porcelain-enamel	70–20%	[55]
Black papers	6–5%	[55]
Black velvet	0.4%	[55]
Carbon black in oil	0.3%	[55]

Table 2–2: *Summary of measured values of* $\rho(0; 2\pi)$ *for various non-metallic materials.*

Chapter 3

Detecting specularities: background

3.1 Introduction

This chapter surveys previous attempts to detect specularities. Previous work on shape inference is left until chapter 5. The work discussed here falls into two distinct categories: chromatic approaches and achromatic approaches. Chromatic approaches attempt to exploit colour differences between specularly and diffusely reflected light. They are reviewed in section 3.4. An achromatic approach is designed to succeed in the absence of colour cues. So it is in a sense more fundamental than a chromatic approach. Often complementary evidence for specularities is provided by the two approaches. So a fully rounded vision system might make good use of both.

The only predecessor to the achromatic approach developed in this thesis is Ullman's S operator. It is described and evaluated in section 3.3. Before proposing the S operator Ullman [106] made various attempts to explain the perception of bright non-matt image regions, as are produced by specularities. One of the explanations that he considered was a search for image regions with irradiances that could only correspond to surfaces that were 'brighter than matt white'. More precisely the search was for surface regions that appeared to have albedoes exceeding unity[1], when modelled as entirely matt. It was proposed

[1]For entirely matt surfaces albedo is surface reflectance and has a legal range of zero to unity.

that an existing computational method, known as lightness computation, could provide the estimates of albedo. But because lightness computation could only provide estimates of relative, rather than absolute albedo, Ullman abandoned this approach. Since then, Land [78] has proposed a method that attempts to estimate absolute albedo—see section 3.2.2. However, Land was more interested in predicting perceived surface colour than accurately estimating absolute albedo. So it is unlikely that this method would be applicable to Ullman's approach. One of the tests for specularity developed in this thesis avoids the need to estimate absolute albedo, by considering a constraint on its dynamic range. In this way only estimates of relative albedo are required.

The next section briefly describes lightness computation and explains how it can be used to estimate albedo. It also serves to introduce the concept of the mondrian world—the reflectance model upon which Ullman based the S operator. Discussion of the S operator itself, is included in section 3.3.

3.2 Lightness computation

3.2.1 Introduction

Regardless of the kind of light illuminating it, an object tends to appear the same colour. For instance a green book remains green under natural daylight or under artificial tungsten lighting, even though the spectral composition of the light differs. This phenomenon is known as colour constancy. It poses a computational problem that any vision system which sees in colour must tackle: the problem of recovering spectrally invariant reflectance properties of object surfaces under varying lighting conditions.

An elegant solution to the problem, proposed by Land and McCann [76], is known as lightness computation. It involves applying a computational process, called a retinex, independently to three separate colour channels. For example the red, green and blue colour channels provided by a colour camera. By ap-

plying the retinex, three different fields of lightness values are produced. By suitably combining these fields useful predictions of perceived surface colour can be made [78]. In effect the retinex estimates the relative albedo corresponding to the wave-band of its colour channel.

3.2.2 The retinex

The retinex was developed by considering a simplified model of the real world. This model is known as the mondrian world because of its similarity to some of the paintings of Piet Mondrian. In the mondrian world surfaces are purely matt, so a Lambertian surface model like (A.10) is appropriate. The spatial variation of the surface properties is also restricted. Specifically it is assumed that:

- albedo ρ_d is piecewise constant, varying only in steps—see figure 3–1 (a).
- total illumination, I, vary smoothly across the scene/surface—see figure 3–1 (b).
- surface normal direction with respect to the viewer is constant everywhere. This corresponds to viewing a plane from a large distance. The upshot of this constraint is that there are no shading variations introduced by surface shape.

Under these assumptions, the image irradiance, E, measured by a camera at any point $\mathbf{x}$ in the image is:

$$E(\mathbf{x}) = a_c \rho_d(\mathbf{x}) I(\mathbf{x}). \tag{3.1}$$

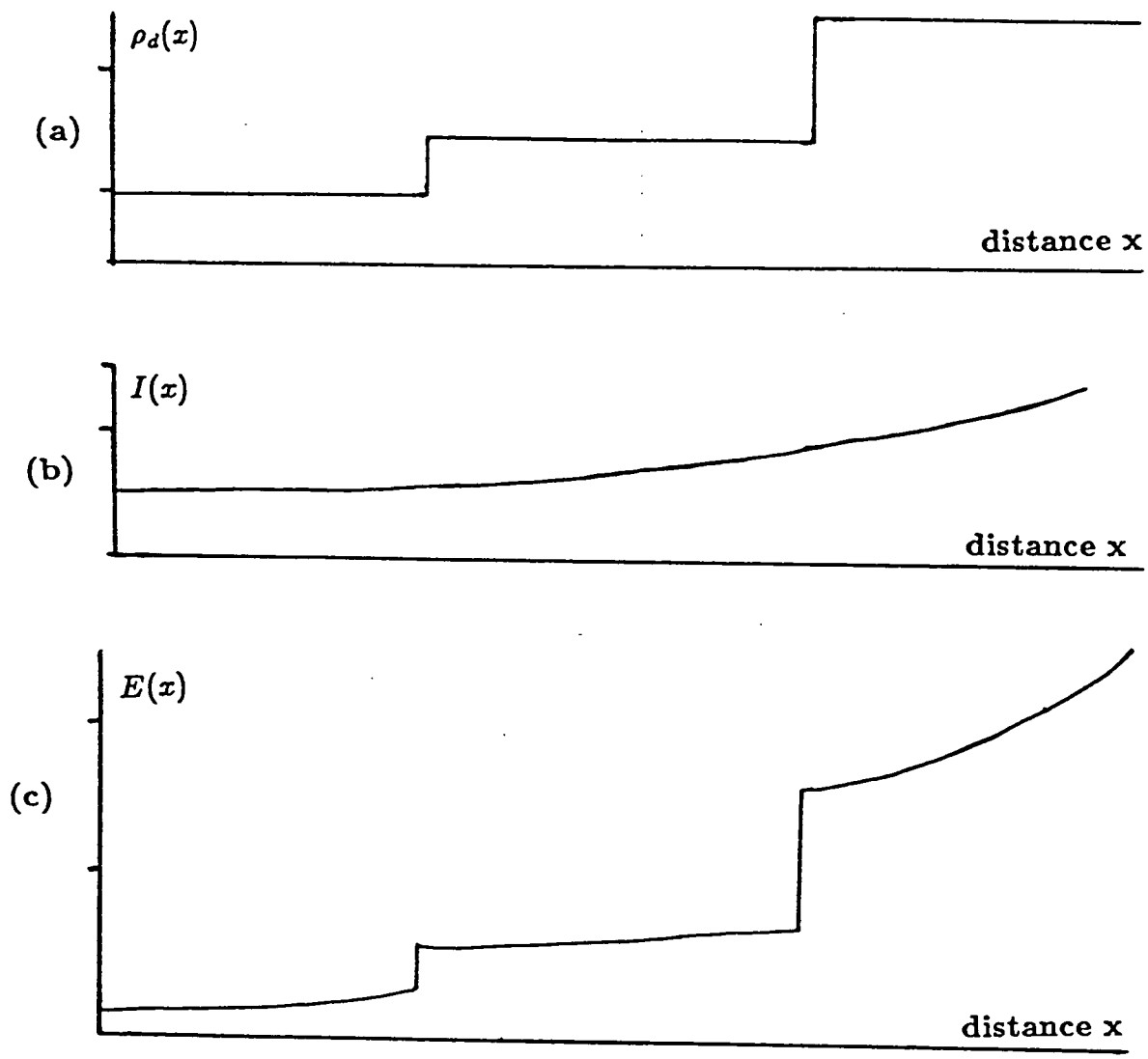


Figure 3–1: *Profiles through a mondrian world. (a) Albedo, ρ_d, is piecewise constant. (b) Total illumination, I, varies smoothly everywhere. (c) Image irradiance, E, is essentially the product of albedo and illumination.*

where ρ_d is piecewise constant albedo and I is the smoothly varying total illumination[2]. a_c is the camera constant—see equation (A.5). Figure 3–1 (c) shows an example of a profile through such an image.

The retinex process consists of two stages:

1. distinguishing between the steps in E due to the steps in ρ_d, and the smooth variations in the illumination function I. This is essentially edge detection—achievable by thresholding the gradient field of E.

2. finding a global solution that reconstructs the original piecewise constant albedo function across the entire image, given only the height and location of the steps detected in 1.

The retinex recovers the albedo function ρ_d up to a multiplicative constant, k. So each lightness value E_l represents relative rather than absolute albedo:

$$E_l(\mathbf{x}) = k\rho_d(\mathbf{x}). \tag{3.2}$$

Land [78] uses a normalisation procedure to estimate the value of k, and thus estimates absolute albedo. This procedure uniformly scales the lightness values, so that their average value over the image corresponds to a mid-grey albedo. The resulting sets of values from all three of the independent colour retinexes are then combined to predict perceived surface colour. Although Land demonstrates that these scaled values are useful for predicting perceived surface colour, it is not claimed that they are accurate estimates of absolute albedo. So they cannot be reliably used for the purpose of detecting specularities—via the retinex-based approach that Ullman considered—see section 3.1. The

---

[2]This expression is the image irradiance equation (2.5) for a Lambertian reflectance map (2.6). Thus I is properly defined as:

$$I(\mathbf{x}) = \frac{1}{\pi}\int_{\omega_i} L_i(\theta_i, \phi_i; \mathbf{x}) \cos\theta_i d\omega_i$$

retinex-based test for specularity developed in this thesis needs no normalisation process, as it uses relative rather than absolute albedo. The estimates of relative albedo are provided directly by a retinex process applied to a single achromatic channel. Blake's retinex algorithm[3] [14] was used for this purpose. Appendix K describes in detail how it was implemented [22].

3.3 Achromatic approaches

3.3.1 Introduction

Motivation for an achromatic approach is provided by the apparent ease with which human subjects recognise specularities in achromatic photographs (i.e. 'black and white' stills). Seemingly, sufficient evidence is available even in the absence of colour or disparity information. Chromatic approaches, that use colour are reviewed in section 3.4, while the disparity information is discussed briefly in chapter 5.

Ullman [106] proposed a computational method to detect light sources in achromatic images. Forbus [37] later demonstrated that the same method could detect certain kinds of specularities. This method is described below, however it turns out to be unsuitable as a general specularity detector.

3.3.2 The work of Ullman

Ullman's work on detecting light sources consisted of three parts:

1. an experimental investigation into human perception of light sources,

---

[3]Hurlbert [64] provides an extensive survey of existing algorithms developed for the retinex. Blake's is unique in its correct treatment of the image boundaries.

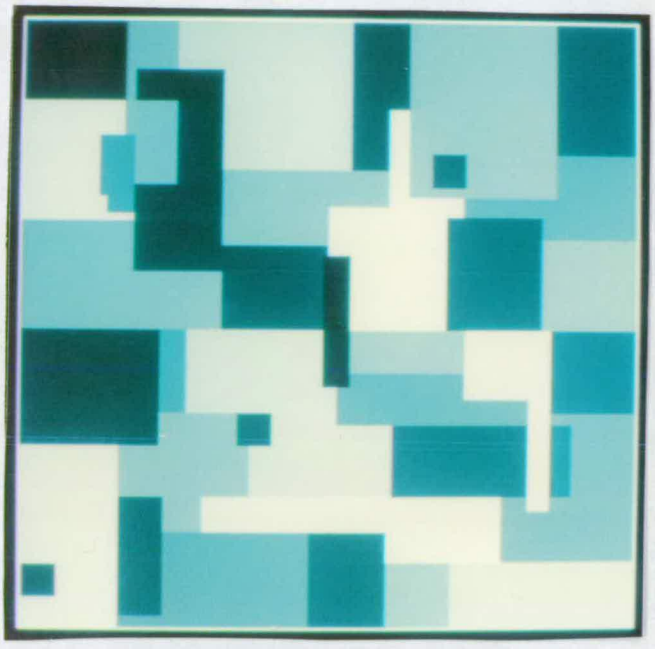


Figure 3–2: *An achromatic mondrian image. It consists of various patches of uniform albedo. Variations of illumination across the image are smooth—this makes them difficult to see.*

2. a discussion of factors that might account for his experimental findings, and

3. the development of a computational method to detect light sources—known as the S operator.

His experiments consisted of disguising a light source as a patch within an achromatic mondrian image—an achromatic version of the type of image discussed in section 3.2. Figure 3–2 depicts such an image. Ullman varied the radiant intensity of the light source patch, and thus investigated under what conditions it was perceived—as a light source. He found that the thresholds of the perception were determined to some extent by the following variables:

- the strength of the illumination that lit the mondrian,

- the strength of the light source patch,

- the albedo of the patches within the mondrian.

His results confirmed that the perception was not simply a matter of recognising the brightest regions within an image. In fact Ullman concluded that no one of a list of six of the most plausible explanations of the perception was individually sufficient.

The explanations that he sought had also to be applicable to machine vision. This meant they involved only the information available to such a system, i.e. the image irradiances provided by an achromatic camera. The same informational restriction is placed upon the achromatic approaches developed in this thesis. The approaches developed in this thesis however, do employ two of the explanatory factors that Ullman dismissed. These factors are high local contrast and lightness computation. Use of these factors is justifiable, because they are not used to provide the complete answer by themselves. Rather they each provide partial evidence—which when combined produces an improved overall result (see chapter 4).

3.3.3 The S operator

The S operator is the computational method designed by Ullman [106] to detect light sources. It has been used to detect specularities by Forbus [37]. A specularity is essentially the virtual image of a light source, so the two detection processes ought to be closely related. However it transpires that the S operator can only distinguish specularities in very limited circumstances. Section 3.3.5 and 3.3.6 detail its limitations.

The S operator was not directly based on Ullman's experimental findings. Rather it was designed by considering the image irradiance equation for the achromatic mondrian images that constituted his experimental stimuli:

$$E(\mathbf{x}) = I(\mathbf{x})\rho_d(\mathbf{x}) + L(\mathbf{x}). \tag{3.3}$$

This equation is a simple extension of that describing the mondrian world (3.1)[4]. The additional term L(**x**) is added to account for contributions to the image irradiance provided by the light source patch—introduced in section 3.3.2. The constraints of piecewise constant albedo, ρ_d, and smoothly varying illumination, I, are as described for the mondrian world in section 3.2.2. In addition L, is constrained to be constant within the light-source patch, and zero elsewhere.

For this modified mondrian world model, discontinuities in E with respect to **x** are due either to step changes in albedo or a light source edge. The S operator was designed specifically to distinguish between these two types of edge. Figure 3–3 shows how this can be done: an irradiance step accompanied by a change in irradiance gradient E_x indicates a step in albedo; however when there is no change in E_x the edge is due to a light source. This is a direct consequence of the modified mondrian world model [106]. It constitutes a simple test that can be implemented on any irradiance profile by:

1. Identifying each step edge.

2. Measuring the value of E_x on either side of the step and comparing them to see if they are approximately equal.

The S operator combines both steps in a single operation to produce a scalar response at each point along the profile. It is defined as:

$$S(x_1, x_2) = E(x_1) - E(x_2)\frac{E_x(x_1)}{E_x(x_2)} - E_x(x_1)d, \quad d = x_2 - x_1. \tag{3.4}$$

S is computed for each pair of points, x_1,x_2, separated by interval d, along the profile[5]. Large positive peaks in S correspond to light source boundaries. indicates a light source edge between x_1 and x_2. An example is shown in figure 3–4.

---

[4]The camera constant a_c is assumed to be unity here. This makes no difference to the arguments that follow.

[5]The last term in (3.4) simply adds linear correction—which allows a small component to the step edge due to the prevailing irradiance gradient.

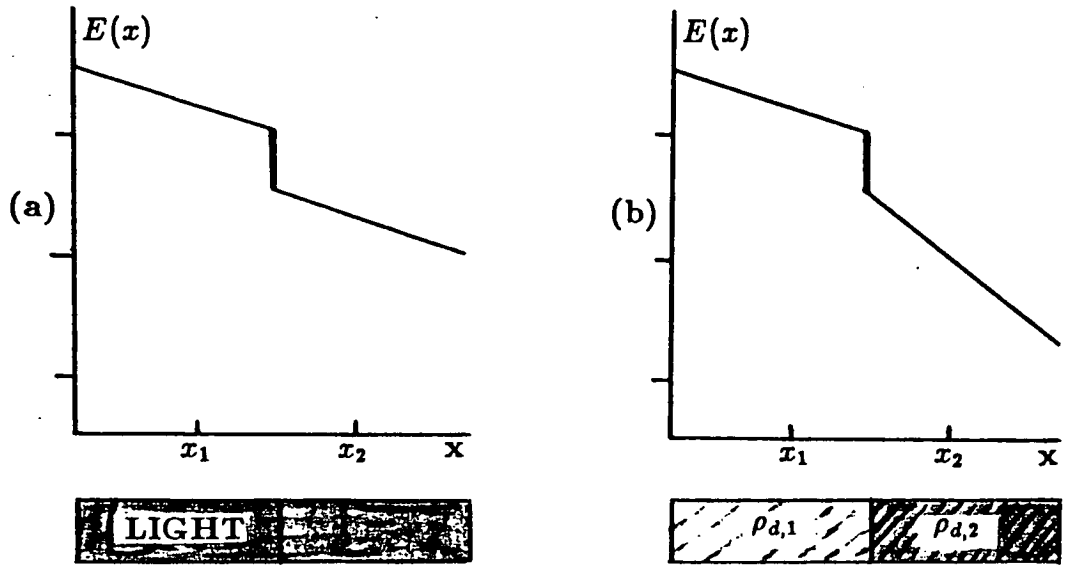


Figure 3–3: *Cross-section through the image irradiance field E. (a) At a discontinuity due only to a light source edge the gradient is unchanged. (b) At a discontinuity due only to a change in albedo the gradient changes.*

The two step test above can be implemented by running the S operator, and marking peaks, except that:

- Ascending steps (i.e. where $E(x_2) > E(x_1)$) are ignored. This means that the operator must be run in both directions (i.e. with d +ve and −ve) in order to detect all light source boundaries along a profile.

- Locations where the ratio $E_x(x_1)/E_x(x_2)$ is not well defined are ignored. This case is rather restrictive as it occurs any where that $E_x \approx 0$—which which need not be uncommon when the illumination gradient is small.

3.3.4 The S operator as specularity detector.

Forbus [37] shows examples of the S operator detecting specularities on some glossy planar surfaces. However no detailed study of its performance is given.

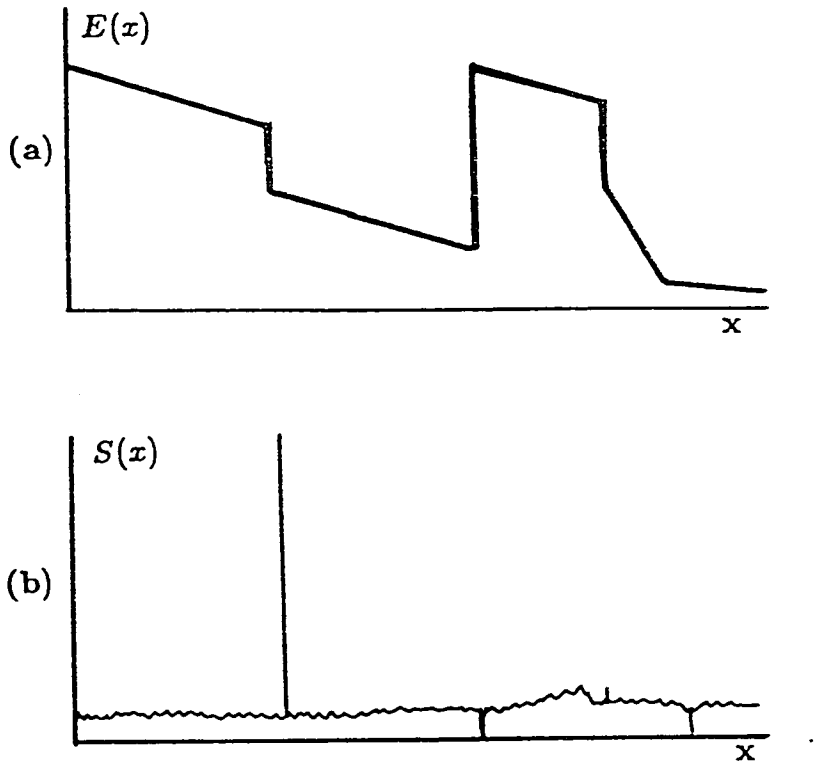


Figure 3–4: *An image irradiance profile (a) over which the S operator has passed. (b) is the operator's response. Note the large positive peak corresponding to a light source edge—a step across which the gradient is unchanged.*

Neither is it claimed that the S operator is suitable for curved surfaces. As a large proportion of the specularities in the real world lie on curved surfaces [73], this places a serious limitation on the applicability of the S operator as a specularity detector. These limitations stem from that fact that the extended mondrian world model, upon which it is based does not model:

- curved surfaces.

- depth variations.

These limitations will be discussed in turn below.

3.3.5 Curved surfaces

Smoothly curved surfaces produce smooth variations in image irradiance—often referred to as shading. The reflectance map discussed in chapter 2 formally specifies the nature of the local shading. For some materials and under some illuminations the shading is mostly due to specular reflection—e.g. as in figure 3–5 (a). In other conditions diffuse reflection dominates—e.g. see figure 3–5 (b). A practical specularity detector must distinguish between specular and diffuse shading. The S operator as it stands is unable to do this. Of course the response obtained from the S operator depends on the value of the sampling interval d employed. Nevertheless a value of d cannot be selected so that the resulting S operator can distinguish between the smooth specular and smooth diffuse shading shown in figure 3–5:

- An S operator with a small sampling interval, d, always produces a small response on any smooth profile, specular or diffuse—$E(x_1) \approx E(x_2)$ and $E_x(x_1) \approx E_x(x_2)$, so that $S(x_1, x_2) \approx -E_x d \rightarrow 0$ as $d \rightarrow 0$.

- So a larger value of d is required to detect a smooth profile (such as that through figure 3–5 (a)) as a specularity. However running an S operator with the same value of d over a region of diffuse shading—like that in figure 3–5 (b)– results in an equally strong response. Simply thresholding the response cannot distinguish the two cases.

The S operator is not able to distinguish between specular and diffuse shading variations. However the cylinder test proposed in chapter 4 often provides a practical way of making this distinction successfully.

(a)

(b)

(c)

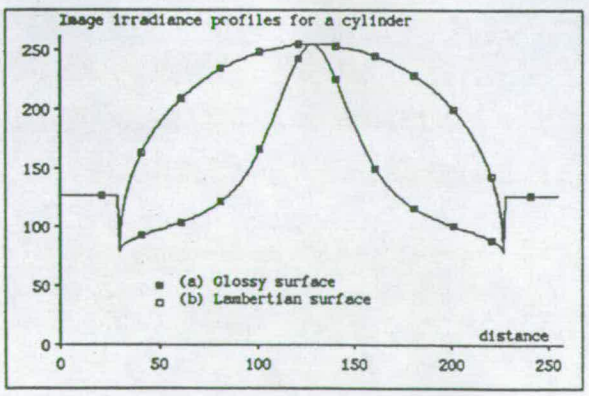


Figure 3–5: *(a) An image of a glossy cylinder. (b) An image of a cylinder in which the shading is due to diffuse reflection. (c) Image irradiance profiles through both (a) and (b).*

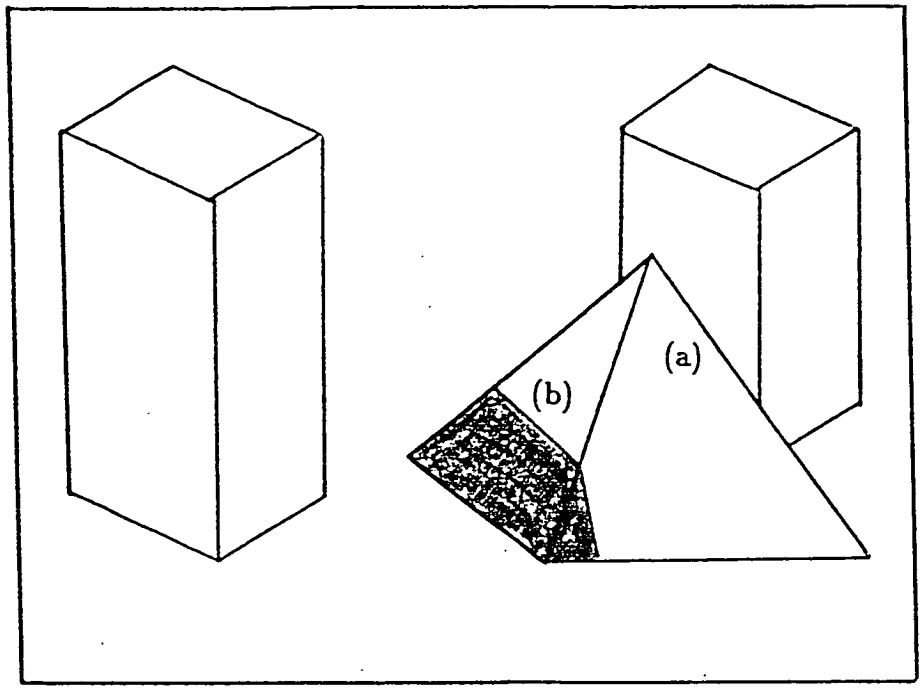


Figure 3–6: *A simple image of a three dimensional scene. (a) The surface of the pyramid occludes the block behind. (b) The other block casts a shadow on to the pyramid. Everywhere but over occluding boundaries and cast shadows the illumination varies slowly.*

3.3.6 Depth variations

The mondrian world for which the S operator is designed is flat. So the depth[6] does not vary across a mondrian image. However depth can and does vary across an image of a three-dimensional scene. Figure 3–6 depicts such a scene. The depth varies smoothly across the smooth surfaces and abruptly at occluding boundaries. In addition objects in three dimensional scenes can also cast shadows upon each other. As a result of both occluding boundaries and cast shadows, discontinuities are introduced into the otherwise smooth illumination function, denoted as I in equation (3.3). These discontinuities, when large enough, produce strong responses as the S operator passes over them. As a

[6]The viewer-to-surface distance

result the S operator will perform badly as a detector either of light sources or of specularities.

These 'rifts' in the illumination field also pose problems for retinex processes that run on real images—see chapter 4.

3.4 Chromatic approaches

3.4.1 Introduction

Recently two different chromatic approaches to detecting specularities have been introduced:

1. Shafer's dichromatic reflection model approach, [72].

2. Gershon's colour shift approach, [39,40].

As mentioned in section 3.1 a chromatic approach can complement an achromatic approach, when colour information is available. A chromatic approach relies on colour differences in specularly and diffusely reflected light. When no colour differences are present—e.g for many metals, or when white light falls on a grey surface—an achromatic method is essential.

The chromatic approaches are variations on the same theme. Below the common aspects of both are discussed before examining the individual approaches.

3.4.2 Aspects common to the chromatic approaches

Both chromatic approaches were developed to be applied to opaque inhomogeneous dielectric materials, such as paints and plastics—see section 2.4. These types of material reflect a proportion of the incident light by sub-surface scattering. Inhomogeneities or pigments in the scattering material selectively absorb light at certain wavelengths. As a result the scattered light reemerges out of the

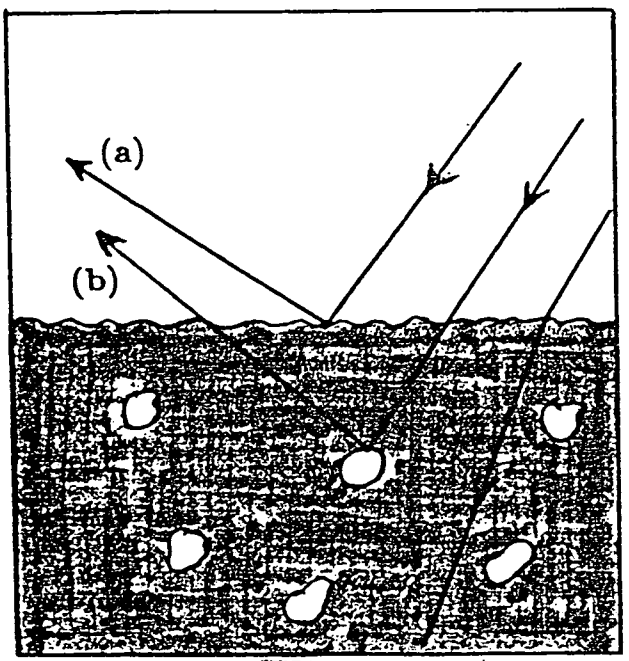


Figure 3–7: *The spectral composition of specularly reflected light (a) is similar to that of the incident light. While the scattered light's composition (b) depends upon the selective absorption of pigment particles within the surface.*

surface with a transformed spectral composition. In sharp contrast, any light that is specularly reflected before it has chance to enter the scattering medium, has substantially the same spectral composition as the incident illumination. Any spectral transformation that does occur is due to quite different physical processes—e.g. the Fresnel relations (see section 2.4). Figure 3–7 depicts the two processes involved. Both of the chromatic approaches exploit the above-mentioned spectral (colour) difference to detect specularities. They do this in four basic stages.

1. A colour camera provides three images of the same view. Each of the three image irradiance patterns is formed by a separate—red, green or blue—wave-band in the visible spectrum.

2. The image is segmented into regions of roughly the same colour. This segmentation must separate specularities from surrounding regions of diffuse shading—if the method is to work.

3. The spectral composition of the light for the points within each region is examined. At each pixel[7] the amount of red, green and blue light constitutes a colour triplet. A colour triplet defines a unique coordinate in a three-dimensional colour space—see figure 3–8. A plot of the colour triplets at each pixel within a given region is made.

4. The colour space plots for each region are analysed in turn, for characteristic features that indicate specularities. If the colour triplets are clustered into a two-line 'dog-leg' feature a specularity is present [72,39].

3.4.3 The dichromatic reflectance model approach

Shafer's approach is based on a simple mathematical model of reflectance called the dichromatic reflection model. Basically this model equates reflected radiance L_r (equivalent to image irradiance via (2.5)) to a sum of a specular and a diffuse component:

$$L_r(\hat{\mathbf{n}}, \lambda) = m_s(\hat{\mathbf{n}})c_s(\lambda) + m_d(\hat{\mathbf{n}})c_d(\lambda), \tag{3.5}$$

where λ denotes a wavelength dependence, and $\hat{\mathbf{n}}$ denotes a dependence on the viewing, lighting and surface geometry—c.f. $R(\hat{\mathbf{n}})$ defined in (2.5). The subscripts s and d respectively denote terms of the specular and diffuse components. Each component is decomposed into independent geometrical and colour terms—m and c respectively.

The discussion in chapter 2 makes clear that models like (3.5) are an oversimplification for real surfaces. They can describe only a small subset of real

[7]Pixel: picture element or digitised image location.

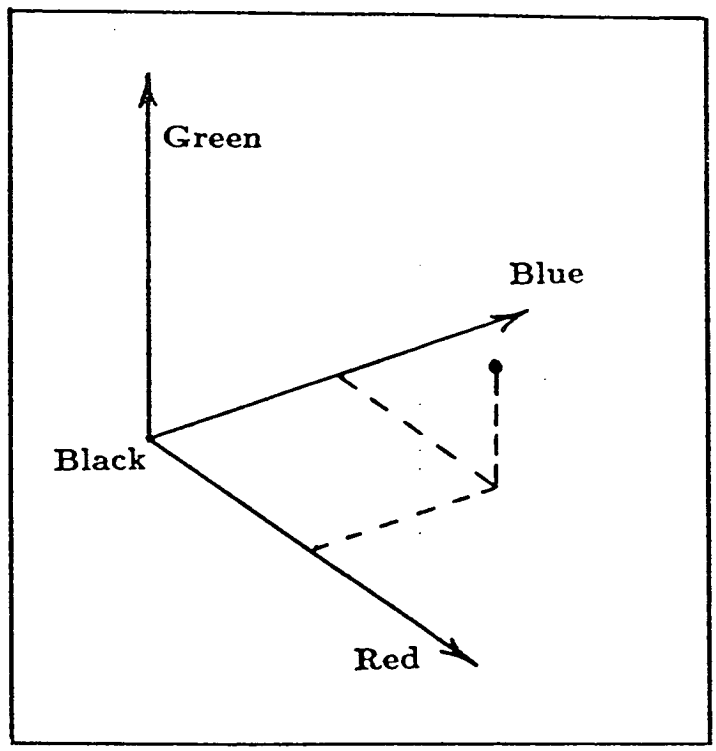


Figure 3–8: *A three-dimensional colour space. A colour triplet defines a unique coordinate within this space.*

world surfaces. However as this approach does not strongly impose this model, it is still able to detect specularities on a variety of opaque dielectric materials.

Finally this approach does not address the non-trivial problem of the image segmentation. A segmentation is assumed to be available.

3.4.4 Gershon's colour shift approach.

Although Gershon's approach [39] is motivated with reference to the Cook and Torrance reflectance model (see section 2.4) it avoids explicit assumptions about surface reflectance. So unlike Shafer's approach, this approach permits any colour shifts caused by specular as well as diffuse reflection.

Gershon's approach differs from Shafer's in two other important respects:

1. it addresses the image segmentation problem. It proposes a practical

if incomplete solution, in terms of a split-and-merge spectrally based scheme [39].

2. it employs colour constant coordinates with which to define the three-dimensional colour space. This allows the scheme to deal simultaneously with the illumination from many light sources—so long as each shares the same spectral power distribution. Details of the transformation into the colour constant space are as yet unpublished. However this transformation must be in many respects similar to the lightness computation described in section 3.2—as both aim to recover a colour constant image description.

Chapter 4

Detecting specularities: new work

4.1 Introduction

A new approach to detecting specularities is proposed in this chapter. A detection process is sought that is to be of immediate use to a practical vision system. Thus processes that require high-level knowledge of the scene such as details of the surface properties and lighting conditions are not considered. A low-level autonomous process that uses only the information provided by the camera is desirable. Here it is assumed that the camera provides an achromatic[1] ('black and white') image. Other approaches that use colour imagery are also being developed—see section 3.4. However these (chromatic) approaches are yet to produce fully autonomous detection processes. If and when such processes are achieved they will be complementary to the approach proposed here—as they tap a different source of information.

Section 4.2 discusses the informational constraints upon a low-level autonomous specularity detector. This leads to the approach based on a Lambertian surface model described in section 4.3. Using this approach three different tests are developed to detect specularities. These tests are based on constraints established by considering the prominent distinguishing features of specularities. The first test assesses whether a region is too bright to be Lambertian. The second if shading gradients are too steep. The third determines if a peak in

[1] Registering light at all visible wavelengths.

shading is too sharp. None relies upon explicit models of specularities—which is a great advantage as was shown in the previous chapters.

Each of the three tests provides evidence for specularities. In section 4.6 a scheme for combining this evidence is proposed. This scheme and the accompanying tests constitute a prototype low-level autonomous process for detecting specularities. The entire process has been implemented by the author as a software system—the details of which are given in appendix J. The results of running this system on a variety of real images are shown in section 4.7.

4.2 A low-level autonomous process to detect specularities

Low-level autonomous processes already play important roles in practical vision systems. Edge detection is a typical example. A description of a scene is produced, using only the information within the image that the camera provides. Because this process needs no high-level knowledge of the scene it is able to run autonomously. The edge map that it produces is two-dimensional and so is not in itself a complete description of a three-dimensional scene. It does however, provide a usefully compact intermediate description. This is a description which other higher level processes in the vision system can conveniently use to create fuller descriptions of the scene. David Marr in his book *Vision* [83] expounded this use of intermediary levels of scene description.

A low-level process that derives a convenient description of the specularities is also required by practical vision systems, for two pressing reasons:

1. Specularities are virtual images (of light sources), not artifacts of the surfaces upon which they lie. They confuse other visual processes such as those concerned with stereo, motion and recognition—because these processes expect all surface features to lie on real surfaces. If they were

provided with a description of the specularities, they could treat specular features as special cases and thus improve their performance.

2. Local surface shape can be inferred from the description of specularities. This can be done by a higher level process like that described in later chapters.

The specular detection process is to be low-level and autonomous so it needs no scene-specific information. Any process that made direct use of the models of specularity described in section 2.4 would require scene-specific information. For example, if the smooth surface model described by equation (2.20) were used to recognise the image irradiance distribution of a specularity, then the following information would be required:

1. the surface orientation at the specularity.

2. the lighting geometry.

3. the refractive index of the surface: η.

4. the proportions of surface and sub-surface reflection: a and b.

For rough surfaces even more information would be required. Without this information, a low-level process cannot make direct use of these models. The next section proposes an alternative approach that does not use explicit models of specularities.

4.3 Constraints on Lambertian surfaces

Intensity profiles on surfaces that reflect light diffusely are subject to certain constraints. A specularity in an image may be recognised because it is not thus constrained. This is the essence of the approach used in this thesis to detect specularities. This approach involves two steps:

1. establishing the constraints.

2. developing tests for violation of the constraints, to indicate the presence of a specularity.

The constraints are established using a model of a diffusely reflecting surface—the Lambertian model. Sections 2.3 and 2.4 show that these constraints usefully apply for a range of real materials. The reflectance map, described in chapter 2, provides the geometrical description of the light reflected by a surface. For a Lambertian surface the reflectance map R_L is particularly simple—see (2.6). Constraints established on R_L provides constraints on the images formed by Lambertian surfaces, via the image irradiance equation (2.5). A previous example of an image irradiance equation for Lambertian surfaces was that describing the mondrian world, (3.1). In section 3.4 it was shown that the mondrian world is an unsuitable model upon which to base a specularity detector—because it does not model curved surfaces. Here a less restrictive approach is made—constraints are sought on the two factors which combine to determine R_L: illumination and effective albedo. For example one constraint restricts the dynamic range that these two factors can induce into R_L. In this case the test for violation of the constraint corresponds to a search for image regions that are 'too bright' to have been formed by Lambertian surfaces. Other tests correspond to searches for other features of specularities that are characteristically non-Lambertian. In all, three different tests are presented here. Table 4–1 summarises the features associated with each. Details of the constraints used

Name of test	Non-Lambertian feature sought
Retinex-based test	Region 'too bright' to be Lambertian
Local contrast test	Shading contrast 'too great' to be Lambertian
Cylinder test	Peak in shading 'too sharp' to be Lambertian

Table 4–1:

by each test are given in the sections that follow. The summary below lists the main assumptions involved in establishing these constraints.

- albedo is piecewise constant—as it was modelled in the mondrian world. This is a reasonable assumption: the reflectance properties of a homogeneous surface are relatively invariant [58]. A notable exception is a 'shaded' drawing, in which the artist mimics chiaroscuro (shape-shading) with extended smooth variations of albedo.

- albedo within the Lambertian model has a dynamic range $r_\rho \leq 10$. The validity of this figure is discussed in section 2.4.7.

- The presence of ambient light ensures that the dynamic range of the Lambertian reflectance map due to directional variations of illumination is no more than 3—see section 2.3.4.

Finally it should be noted, that not every specularly reflecting surface can be detected using the three tests listed in table 4–1. Only specularities with prominently non-Lambertian features can be identified. Other specularities that do not produce such features are missed. Thus there is a fundamental difference between this approach and the chromatic approaches described in section 3.4. Those approaches attempt to solve a much more difficult problem: the complete separation of the diffuse and specular components of image irradiance. So, in addition to specularities having prominent features, they are also trying to detect barely visible low-levels of gloss: e.g. those produced when an overcast sky illuminates lightly varnished timber. In general low levels of gloss are both more difficult and less useful to detect. Only prominent specularities can be

confidently used to infer surface shape. And only prominent specularities will cause significant disruption to other visual processes (stereo, motion and recognition). In fact for edge-based visual processes (e.g. stereo matchers [91,92]), only those specularities that create edge features cause problems.

4.4 The retinex-based test

4.4.1 Introduction

Specularities can be very bright[2] compared with even the lightest matt surfaces. For example specular reflections from a chrome door-handle are often an order of magnitude brighter than the light reflected by the matt white paint covering the rest of the door. However, specularities are not always so prominent—e.g. those on an old bronze drawing pin stuck into the door, may be darker than the surrounding paint. In this section a test is developed for detecting those specularities that are 'too bright' to be matt. Other tests, like those described in later sections, are needed to detect the darker specularities. The test described here, involves *more* than identifying specularities that exceed a certain fixed brightness value—chosen to represent the brightest matt surface. This is because the appropriate value will vary with the level of illumination—which in turn varies widely from image to image and even across a single image. By dealing with relative rather than absolute image brightness the test described below is able to cope with the image-to-image variations of illumination level. Variations within an image itself, are tackled by applying a retinex process, like that described in section 3.2.2. The next section shows how the test evolves from the constraints on the dynamic range of Lambertian reflectance maps.

---

[2]Throughout this introductory section the familiar terms bright and brightness will be used. Strictly speaking the discussion should be in terms of image irradiance.

4.4.2 The constraint

The image irradiance equation describes local variation in image irradiance due to variations in surface orientation (denoted by $\hat{\mathbf{n}}$), via the Lambertian reflectance map function $R_L(\hat{\mathbf{n}})$. In general, the function $R_L(\hat{\mathbf{n}})$ varies with location, $\mathbf{x}$. So, to express the irradiance distribution within an image the equation (A.18) must be extended to include the $\mathbf{x}$ dependence[3]:

$$E(\mathbf{x}) = R_L(\hat{\mathbf{n}}(\mathbf{x}), \mathbf{x}). \tag{4.1}$$

Variations in R_L are of two different types:

1. local variations due to variations of local surface orientation, $\hat{\mathbf{n}}$ or changes in albedo.

2. spatial variations due to variations in the level of illumination, with respect to $\mathbf{x}$.

As established in chapter 2 and summarised in section 4.3, local variations in R_L can be reasonably expected to have a limited dynamic range. In particular, the dynamic range due to albedo variations (assumed to be less than 10) and that due to the local directional distribution of the illumination (assumed to be less than 3), combine to give an overall upper bound of 30. In the absence of the spatial variations in R_L, the same bound would hold for image irradiance, via (4.1). In this circumstance, this Lambertian constraint on dynamic range could be used to detect specularities: in an image with a dynamic range exceeding 30 the bright regions are likely specularities. But, because illumination levels often do vary significantly with respect to $\mathbf{x}$ this constraint cannot be generally applied. However, the retinex process described in section 3.2.2 is specifically

---

[3]The sensitivity of the camera, described in (2.5), as a_c is also a function of $\mathbf{x}$. The attenuation due to this function can be nullified by a once-and-for-all camera calibration. The expression here, assumes that the calibration has been made.

designed to remove such variations. The test proposed here, preprocesses the image irradiance signal through a retinex before applying the dynamic range constraint. Figure 4–1 shows schematically how the retinex is used. The details are given in the next section.

The test can be summarised as two stages:

1. apply the retinex to remove gradual variations in illumination levels. The modified image irradiance signal is denoted E_l.

2. mark candidate specularities where:

$$\frac{E_l(\mathbf{x})}{\min_{\mathbf{x}} E_l} > 30. \tag{4.2}$$

Where the minimum is that of the entire image. Note this test assumes that only bright regions, such as specularities are responsible for breaking the dynamic range constraint—not dark regions. The only phenomenon likely to produce a darkened region is a cast shadow. But ambient illumination[4] falling within these shadows acts to limit how dark they can become. In the unlikely case, that ambient illumination is in places very small the test will be unreliable—e.g. for a view containing an aperture into a very dark cavity.

4.4.3 Retinex preprocessing

This section discusses the retinex process that is applied to an image before the test for specularities (4.2) is made. Section 3.2 describes the retinex process and how it can account for colour constancy. For that task the retinex was employed to eliminate image irradiance variations due to changes in levels of illumination. Here, the retinex is used for the same purpose. Although, originally designed to perform in the flat mondrian world[5], Land [78] shows that the retinex usefully

---

[4]See section 2.3.5.

[5]See section 3.2 for a description of the mondrian world.

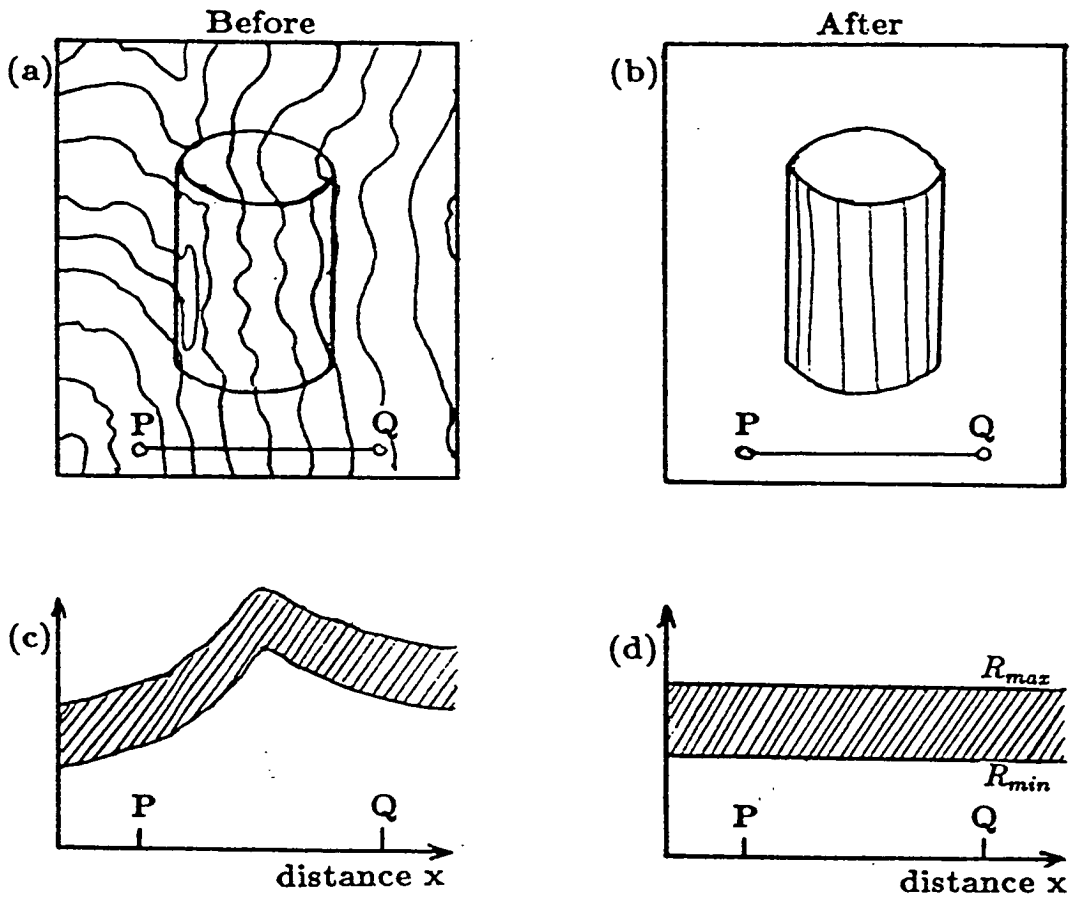


Figure 4–1: *(a) Illumination levels often vary significantly across an image. (b) Retinex preprocessing aims to eliminate these spatial variations. After preprocessing the dynamic range of the image signal, E_l, is that of the local reflectance map $R_L(\hat{\mathbf{n}})$. Consequently, at all points within the image E_l lies between the two extreme values of $R_L(\hat{\mathbf{n}})$: R_{max} and R_{min}—as shown in profile (d). (c) Before preprocessing the image signal E is not so constrained because R_L varies spatially. So, only after preprocessing can the dynamic range constraints be applied uniformly throughout an image.*

eliminates illumination variations present in images of real scenes–which contain curved surfaces. As the specularity test sought here must be able to cope with curved surfaces, it employs a similar retinex process.

Curved surfaces, however, introduce a problem for the retinex that does not exist in the flat mondrian world:

- **In the mondrian world** it is simple to distinguish and remove image irradiance variations due to changes of illumination level. A threshold value can be chosen, below which all the illumination gradients lie.

- **Curved surfaces** introduce additional small gradients into the image irradiance signal. These gradients are *not* caused by changes in illumination levels—but are due to local variations in surface orientation, i.e. shape-shading. So it is no longer possible to completely isolate the changes due to the variations in the illumination level.

Land [78] proposed a practical solution to the problem: he used a gradient threshold large enough to eliminate any gradual variations in illumination level and accepted that some small shape-shading variations will also be lost. In terms of the dynamic range test, (4.2):

- **overestimating** the threshold unnecessarily contracts the dynamic range and so reduces the test's potential to detect specularities—see figure 4–2.

- **underestimating** the threshold will mean not all the illumination variations are eliminated. As a result the dynamic range of the modified signal, E_l, will be left erroneously large. Consequently false detections are likely if the test (4.2) is applied.

On the whole, it is better to overestimate the threshold and accept a few missed specularities, rather than use an underestimate and risk false detections. Those specularities missed due to the contraction of the dynamic range might be detectable using the other tests. The contraction will be severest in images containing extended regions of gentle shading variations. In these regions image

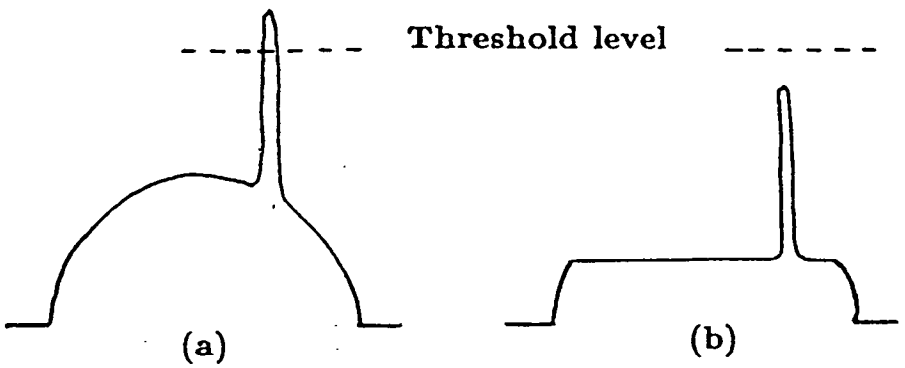


Figure 4–2: *(a) the irradiance profile through a glossy sphere. (b) Overestimating the threshold for the retinex unnecessarily contracts the dynamic range and so a specularity may be missed.*

irradiance gradients are small and thus are likely to fall below threshold. The contraction is less of a problem for images in which most of the shading variation is concentrated into small regions. In these regions image irradiance gradients are large and thus likely to exceed the threshold. In many images the shading variations are concentrated into small regions so it is usually worthwhile applying the test. The extremal boundaries of the handle of the mug shown in figure 4–15 (page 112) provide a typical example of how shading gradients concentrate into small image regions. Such regions are likely to occur anywhere that surface orientation, $\hat{\mathbf{n}}$ varies rapidly with respect to $\mathbf{x}$.

For images of real curved surfaces, Land [77] established a practical choice of the threshold that eliminated any small variations due to illumination while usefully retaining significant shading variations. He chose a threshold on the magnitude of the image irradiance gradient that corresponded to a six percent local contrast between adjacent image locations[6], x_1, x_2: i.e. if

$$E(x_2)/E(x_1) \geq 1.06 \tag{4.3}$$

[6]In Land's experiments image dimensions where roughly 3 or 4 feet and the contrast was measured over a distance of about a third of an inch.

the threshold is exceeded. The retinex process used here, for the purpose of detecting specularities uses the same choice of threshold. However, to ensure that the retinex is robust to noise the image is smoothed before applying the retinex.

As figure 4–3 shows smoothing reduces the magnitude of the signal's gradients—so that the threshold must be reduced as well. Appendix C describes in detail how this is done. Figure 4–15 (page 112) shows that specularities can be detected by applying the retinex followed by the test, (4.2)—more details are given in section 4.7.

The performance of the retinex is impaired in some circumstances. The next section describes these circumstances and the effect on the performance of the test for specularities.

4.4.4 The retinex: associated problems

In most circumstances, the retinex process described in the section above usefully eliminates image irradiance variations corresponding to changes in the level of illumination. However two problems that can arise when using the retinex ought to be considered:

1. **Abrupt changes of illumination level** that occur at cast shadow boundaries or over occluding edges[7] are not eliminated by the retinex. These changes remain in the image irradiance signal after retinex processing—and so may extend its dynamic range causing false detections. In theory, these changes should be identified separately and eliminated. Witkin [110] suggests how they might be identified. In practice, the presence of ambient illumination should ensure that the extension of the dynamic range is insignificant compared with that produced by prominent specularities.

[7] Figure 3–6 (page 69) depicts these two phenomena.

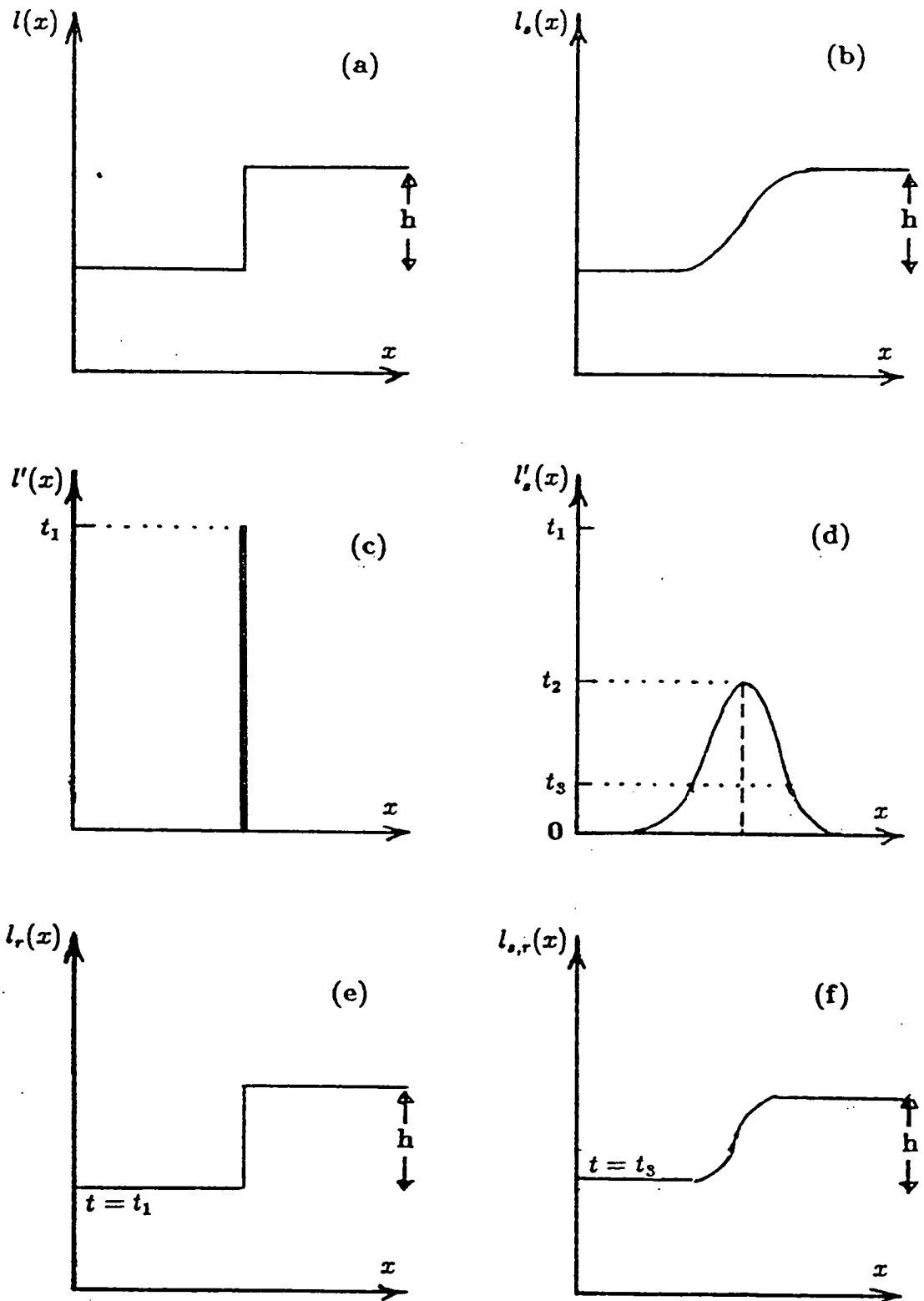


Figure 4–3: *(a) A step, height h, in signal l creates (c) an impulse of magnitude t_1 in the gradient function l'. (e) The retinex reconstructs the step by employing threshold $t = t_1$. (b) The smoothed signal l_s creates a wider pulse of smaller gradients (d). (f) In this case, a lower threshold $t = t_3$ usefully preserves most of the step height while removing small illumination gradients. For more details see appendix C.*

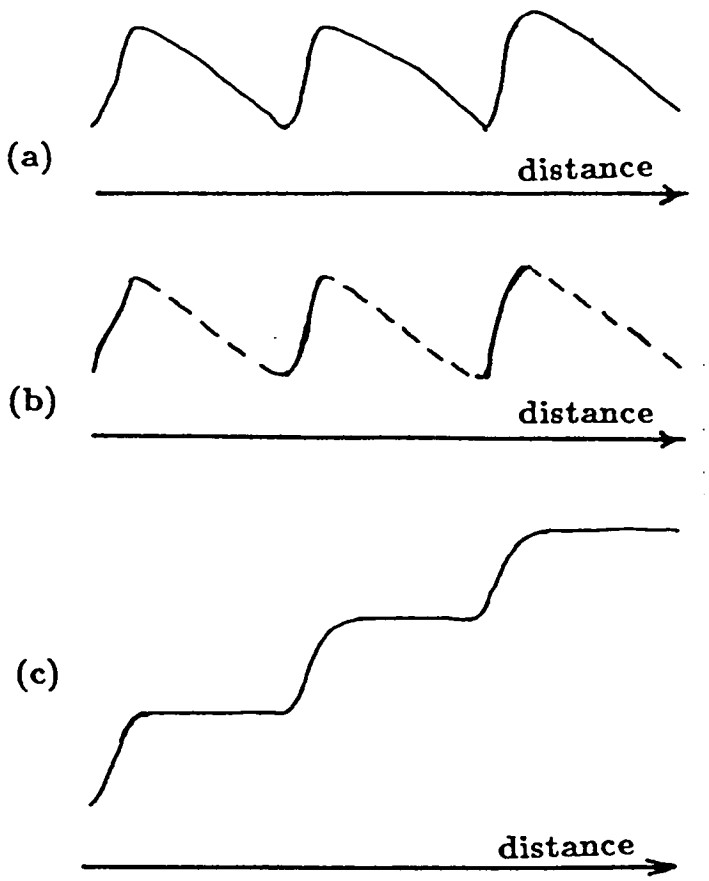


Figure 4–4: *(a) A profile through an image containing a rare combination of steep and gentle gradients. (b) Thresholding preserves only the large gradients (marked as solid lines). (c) The signal after retinex-processing actually has an increased dynamic range.*

2. **In a few somewhat pathological circumstances** combinations of steep and gentle gradients cause the retinex process to actually increase the dynamic range of the image irradiance signal. Figure 4–4 shows an example of how this can arise. In these circumstances the retinex-based test is unreliable as a specularity detector.

4.4.5 Imaging considerations

Before applying the retinex-based test described above, the characteristics of the image irradiance signal provided by the camera should be considered. For example, the original signal must be capable of exceeding the dynamic range limit, before any specularities will be detected. And it should also be proportional to image irradiance. In this section these imaging factors are considered and practical suggestions are made on how the available signal can be used.

The signal provided by a conventional T.V. camera often is not directly related to image irradiance. There are three common reasons for this

1. Some T.V. cameras contain built-in circuitry that automatically controls the gain of the signal. Dark images are made lighter and light images are made darker by an automatic gain control (a.g.c). This is achieved by adding or subtracting a constant value c to the signal received by the camera—over the whole image. The appropriate value of c is selected for an arbitrary image by histograming the irradiance values within it. Although a.g.c. is useful for many applications, it results in a signal with an unknown correspondence to image irradiance. So it is necessary to disable the a.g.c. before the retinex-based test can be applied.

2. Some T.V. cameras have non-linear responses. Sometimes the non-linearity is an inherent property of the imaging device—e.g. a vidicon tube. At other times it is introduced artificially, to make the camera compatible with a display device that it is designed to feed—e.g. some C.C.D. cameras. In both cases the non-linearity can be approximated by a power law. So the signal g can be expressed as

$$g = E^{\gamma}, \tag{4.4}$$

where E is image irradiance and γ is the index of the power law. So before applying the retinex-based test to the signal, it ought to be processed to eliminate the non-linearity (i.e. $g \rightarrow g^{1/\gamma}$).

3. It is conventional photographic practice to use an aperture stop that saturates the signal at specularities. Although this technique often results in images containing a more interesting range of tones, it truncates the dynamic range of the signal. This truncation reduces the potential of the retinex based test to detect specularities. In order to avoid this problem, a smaller aperture should be used, so that specularities do not saturate, even if the rest of the image does appear rather dark.

Although the signal provided by a conventional T.V. camera is represented with 8 bit precision, its dynamic range is typically no better than 6 bits, (64). The retinex-based test only requires a signal capable of exceeding a dynamic range of 30. So it is just feasible to apply the test and detect specularities in the image provided by conventional T.V. cameras. The result shown in section 4.7 demonstrate that this is often a practical proposition. In practice a dynamic range greater than 6 bits is desirable because:

- noise often corrupts the lower bits in the image.
- specularities will often be saturated if only 6 bits are used.
- the figure 30 applies to the signal after retinex processing.

One way of obtaining a signal with an extended dynamic range would be to combine the signal provided by two images of the same view taken at different aperture stops. The upper end of the dynamic range would be supplied when the aperture is narrow, the lower end when it is wide. A practical vision system could acquire the two images in rapid succession using a motorised stop adjuster. Another possibility would be to take control of the a.g.c. and adjust the signal electronically.

Results for the retinex-based test are shown later (section 4.7, pages 110–116).

4.5 Local tests

4.5.1 Introduction

Both the local contrast test and the cylinder test, mentioned in Table 4–1 are local tests for specularities. Below the notion of a local test is introduced and the differences between a local and a global test (like the retinex based test, above) are described. Both local tests are based on a similar principle: they look for shading variations that are inconsistent with the Lambertian shading model (see section 4.3). Section 4.5.3 describes where and in what direction specularities produce the greatest inconsistencies. It transpires that the tests are best applied to one-dimensional profiles extracted from the image, along straight lines. The straight lines that intersect local shading maxima are used. Details of how the local contrast test and the cylinder test are applied along these profiles are left until sections 4.5.4 and 4.5.5 respectively.

4.5.2 Local and global tests

The retinex-based test described in section 4.4 is a global test because

- the retinex process itself involves a global minimisation—see appendix K for details.

- the dynamic range test (4.2) requires the quantity $\min_{\mathbf{x}} E_l$—to be computed this involves the signal over the entire image.

As such, it combines information from all parts of the image when deciding which regions to mark as specular. The tests described below are local tests: when deciding whether a region is specular they use only the shading information in the local neighbourhood of the region. Local tests are limited in what they are able to achieve. For example they cannot be expected

to confidently detect any specularity that loses its specular appearance when most of the surrounding image is masked out. The dimmer, crescent-shaped specularity in figure 4–13 (page 110) is of this type: seen out of context of the rest of the image it might well be Lambertian shading. Consequently it is best to treat a local test as providing somewhat incomplete and inconclusive evidence for specularities. But, by combining the evidence provided by both local tests and the global retinex-based test, a more complete, more certain description of the specularities within an image can be obtained. Section 4.6 proposes a scheme to do this.

4.5.3 Local tests on shading variations

A glance at figure 3–5 (page 68) shows that the shading variation over specular and Lambertian surfaces can differ significantly—even if the surfaces are the same shape. Profiles through these shading variations—like those shown in figure 3–5 (c)—can be used to distinguish the specular and Lambertian cases. The two local tests exploit the information provided by such profiles, in order to provide evidence for specularities. The local contrast test assesses whether the contrast over each edge in the profile is consistent with that of a Lambertian surface—while the cylinder test assesses whether a profile's peak is 'too sharp' to be Lambertian. It does this by measuring the second derivative at the peak.

For both tests the best evidence of specularity is available from profiles along the direction in which the image irradiance varies most rapidly. Figure 4–5 depicts some typical specularities and indicates the profile-lines which provide the best evidence. It shows the two most common forms of specularities:

1. an elongated blob containing a spine of local maxima[8]. For this form, the

[8]One might imagine that a distributed, fairly uniform light source like the sun would produce a specularity containing maxima extended in two dimensions—rather than a spine. However this is only the case if the surface is planar and viewed along its normal

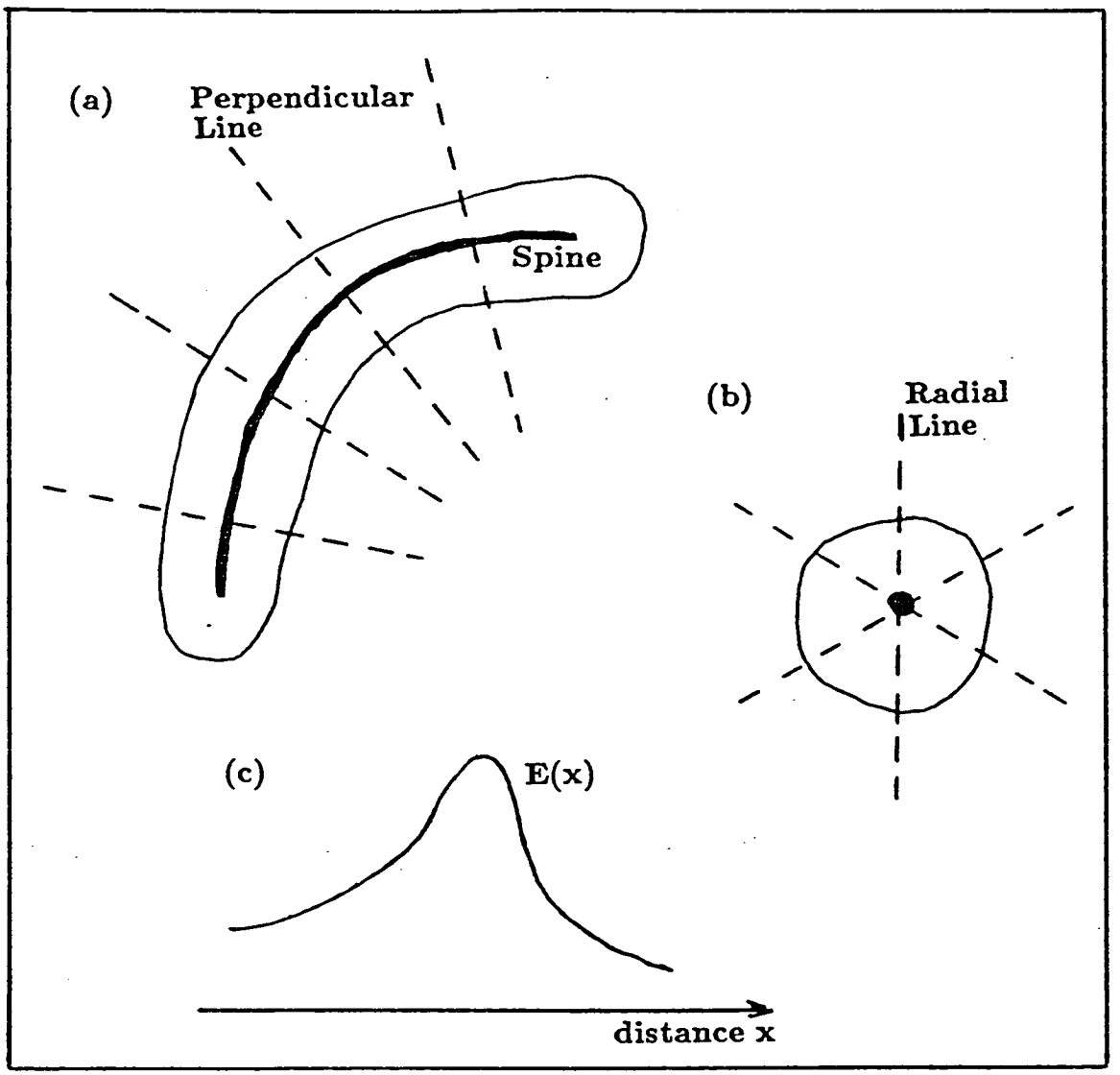


Figure 4–5: *Two forms of specularity: (a) An elongated blob containing a spine of local maxima. In this case the best evidence of specularity is obtained along the perpendiculars to the spine (shown as dashed lines). (b) A compact blob containing a single local maximum. In this case the best evidence is obtained along the radial lines shown. In both cases the profile is characterised as a local maximum flanked on each side by steep edges—as in (c).*

tests should be applied along profiles that intersect the spine perpendicularly—as in figure 4–5 (a).

2. a compact blob containing a single local maximum. This form can be produced by highly curved and specular surfaces. In this case, a set of evenly spaced radial profiles are appropriate—as in figure 4–5 (b).

For both forms of specularity the profile will be characterised—like that shown in figure 4–5 (c)—as having a local maximum flanked on both sides by steep edges. These edges correspond to the sudden drop off in the specularly reflected light occurring at the border of the virtual image of the light source. Often these edge features are sufficiently prominent to be marked by an edge detector. So they may consequently disrupt other visual processes that use edges—as mentioned in section 4.3. The specularities that create such prominent edge features are the most important to detect—so that the disruption that they cause can be avoided. In fact, the method used in this thesis, to extract profiles along which the local tests are applied, actually uses the output of an edge detector.

Profiles are extracted by following certain paths between edge locations in the image—as outlined in figures 4–6 and 4–7. Full details of the method used are given in Appendix D. The section below describes the local contrast test which measures the contrast over these edges in order to provide evidence for specularities.

4.5.4 The local contrast test

The local contrasts across the edges of specularities can be much greater than those produced by Lambertian shading variation. This fact is exploited by the local contrast test to provide evidence for specularities. Again the test uses the model of Lambertian surface reflection, equation (4.1). For this equation

---

direction. Otherwise the effects of foreshortening and surface curvature combine to create spine maxima.

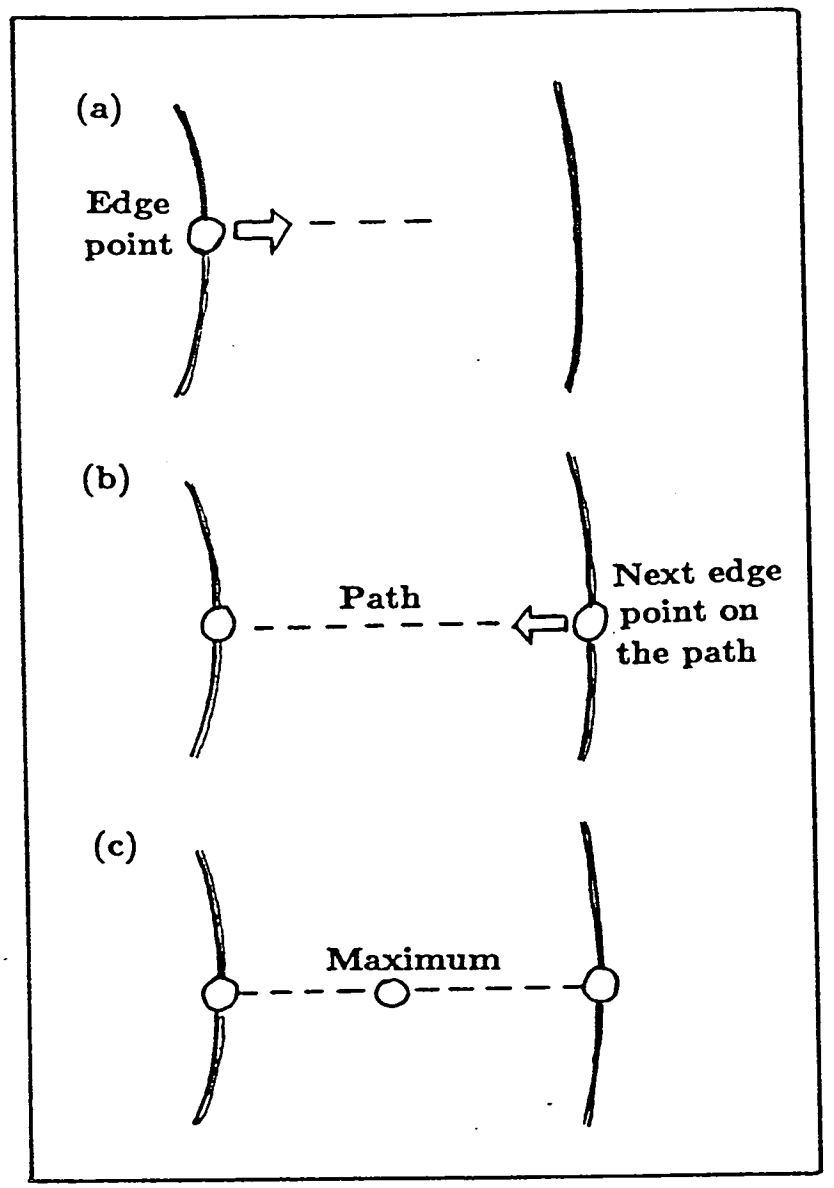


Figure 4–6: *Extracting profiles from an image: (a) A path is initiated from an edge location in the direction of greatest ascent. (b) If the next edge point encounted is in the opposite direction then the profile along the path is extracted. The profile's specification consists of the path's direction and the location of the maximum signal along the path (c).*

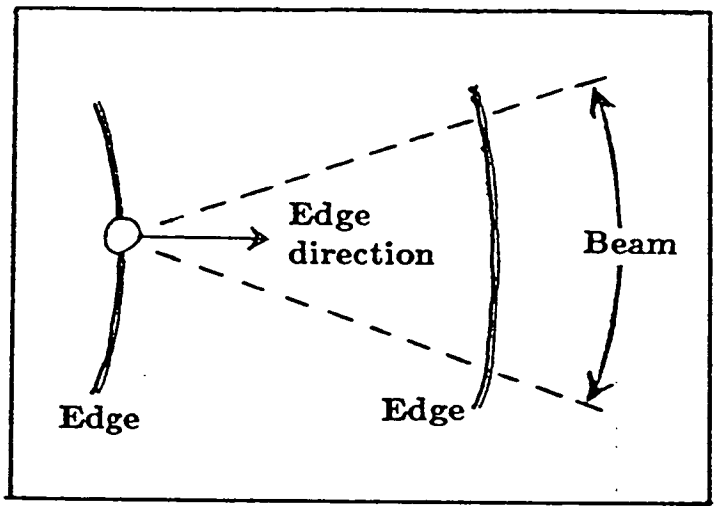


Figure 4–7: *To make the profile extraction method noise resistant a beam of paths rather than a single path is initiated from each edge point. The beam is centred about the direction of the original path.*

only the local variations—due to surface orientation and albedo—produce significant local contrasts, so only they are considered by the test. Again the local variations in R_L can expected to have a dynamic range with an upper bound of 30 (10 for albedo and 3 for illumination). These bounds are used to help distinguish the different types of edges that can occur in image profiles. Figure 4–8 shows a profile through a shading maximum. This might be due to a specularity—like the profiles discussed above in section 4.5.3—or it may be due to Lambertian shading. The contrasts across the edges of the profile are marked. Edges over which the local contrast $c > 30$ are inconsistent with the Lambertian shading model—they are candidate specularities. Edges over which $c > 3$ are only consistent with the model if albedo differs on either side. Across a homogeneous surface albedo is constant[9] so only values of $c \leq 3$ are consistent with the model. In fact this upper bound of 3 is more useful than the higher bound of 30. Very rarely does local contrast exceed 30 and in such cases the

[9]See section 4.3.

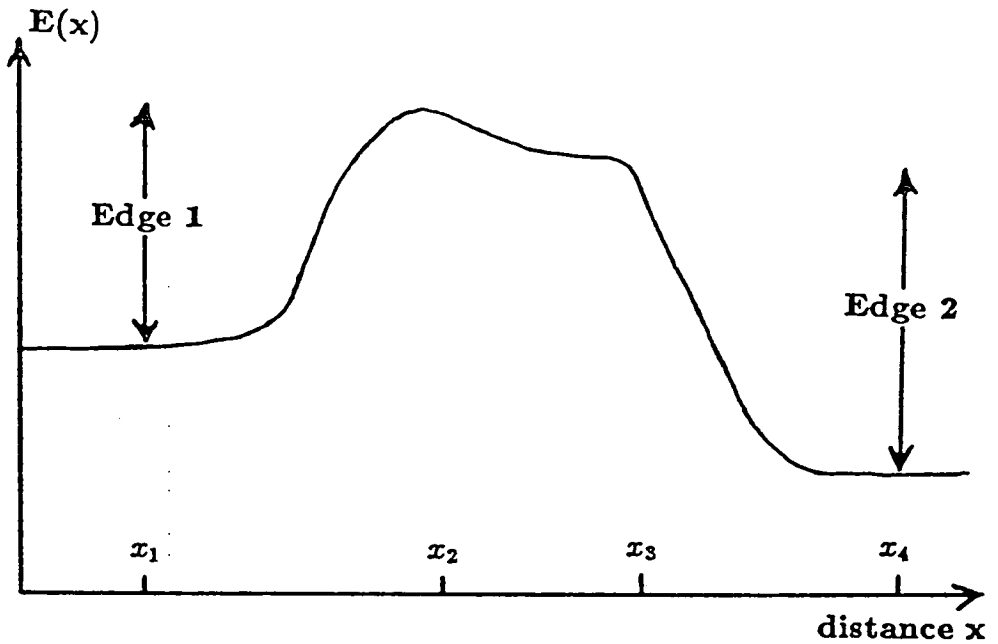


Figure 4–8: *Contrasts over edges in a profile. Over edge 1 contrast is* $c_1 = \frac{E(x_2)-E(x_1)}{E(x_1)}$, *and over edge 2 it is* $c_2 = \frac{E(x_3)-E(x_4)}{E(x_4)}$.

retinex-based method is likely to be triggered[10]. On the other hand values of $c > 3$ are quite common: so the local contrast test is based on this upper bound.

So the test provides circumstantial evidence for specularities across an edge in a profile for which $c > 3$. The evidence is only circumstantial because other phenomena, such as cast shadow, occluding and material boundaries can also create local contrasts exceeding 3. However this circumstantial evidence is used to reinforce the evidence provided by other tests—as described in section 4.6.

Finally, it should be noted: that as for the retinex-based test, the image signal has been assumed to be directly proportional to image irradiance. In order to ensure that this is so, consideration of the imaging process should be made—as described in section 4.4.5.

---

[10]Indeed none of the real images obtained for this thesis contained local contrasts exceeding 30.

4.5.5 The cylinder test

In general a profile produced by a specularity is more likely to have a sharp peak than that produced by Lambertian shading. An example of this is shown in figure 3–5 (page 68). The cylinder test attempts to distinguish the profiles produced by smooth specular shading from Lambertian shading. If the magnitude of the second derivative at the shading maximum exceeds an upper bound then smooth Lambertian shading is unlikely—there is evidence for a specularity. The upper bound itself is computed from the gross properties of the profile as described later. However there are two limitations to this approach:

1. Not all specularities create profiles that have sharp peaks. In particular, specularities produced by slowly curving surfaces and extended light sources can produce rather flattened peaks. So the cylinder test will not detect these specularities. They might be detected using other tests that exploit different specular features.

2. Other phenomena can also create profiles with sharp peaks. Indeed even purely Lambertian surfaces can do so. For example a thin white line on a dark background will produce a very sharp peak. But this can be easily discounted because of the narrowness of the profile. Consequently the test is used only where the profile is wide enough for its shape to be effectively resolved by the image array.

Despite these limitations the cylinder test—as formulated below—provides a useful contribution to the evidence collated by the scheme for detecting specularities.

Like the other two tests the cylinder test is based on the Lambertian surface model. However in this case, additional assumptions are added to the model. Specifically it is assumed that:

1. a profile produced using the model has uniform albedo between its two flanking edges. So that all variations between the edges are due to changes

of surface orientation. This is not an unreasonable assumption: albedo is fairly constant over the surface of a homogeneous material[11]—at a material boundary the abrupt change in albedo ought to create a detectable edge in the profile.

2. the illumination is modelled using the collimated-source-plus-ambient-light model, (2.8). This model provides a worst case: the sharpest Lambertian profiles are produced by highly directional illuminations. Consequently an upper bound on the sharpness of the peak established using this illumination model will apply for other more distributed illuminations (like those described in section 2.3.7).

3. the shape of the surface is assumed to be locally cylindrical. On the face of it, this is a rather restrictive assumption: in many circumstances other surface shapes occur. To formulate a reasonable constraint on the sharpness of a profile produced by an unconstrained shape is very difficult. The formulation below, that uses the cylindrical assumption is at least tractable. Scope exists for further work in this area. Some justification for the cylindrical assumption is provided by the facts:

    - The test is only applied to profiles through spine maxima—as shown in figure 4–5 (a). The presence of a spine, is itself a strong indication of approximate local cylindrical shape[12].
    - The dominant curvature of the surface is therefore orthogonal to the spine. Since the test requires only a bound on, rather than an accurate estimate of shading gradient, it is robust to modest curvature along the spine.

---

[11] See section 4.3.

[12] see section 5.1.3.

For the Lambertian surface model embodying all three of the above assumptions, the following simple constraint on the sharpness of a peak in Lambertian shading can be derived. The magnitude of the second derivative of the irradiance signal at the peak in a profile has an upper bound:

$$U = \frac{E(x_P)}{r_1 r_2}, \tag{4.5}$$

where E is the image irradiance—and as shown in figure 4-9—x_P is the peak location and r_1 and r_2 are the distances in the image from x_P to the two flanking edges. The derivation of this constraint in appendix E shows that it applies for all:

- viewer-sources geometries.
- cylinder radii.
- levels of ambient illumination.
- size of patch—regardless of whether it is a whole cylinder. or just a little piece.

Profiles with a second derivative D at x_P, such that

$$D > U \tag{4.6}$$

are inconsistent with the model and are thus likely to be specularities. An estimate of D can be obtained directly from the profile using the second difference:

$$D = (1/h^2)(G_1 + G_2), \quad \text{where} \quad G_1 = E(x_P + h) - E(x_P), \quad \text{and} \quad G_2 = E(x_P - h) - E(x_P). \tag{4.7}$$

h is the sampling interval of the profile. For profiles along oblique directions within the image, values of D can be obtained by interpolation[13] .

---

[13]Interpolation involves averaging which can significantly reduce the peak value of D for a narrow profile. To avoid this problem profiles were obtained only along 12 special directions, at well spaced orientations—those associated with the 3,4,5 triangle.

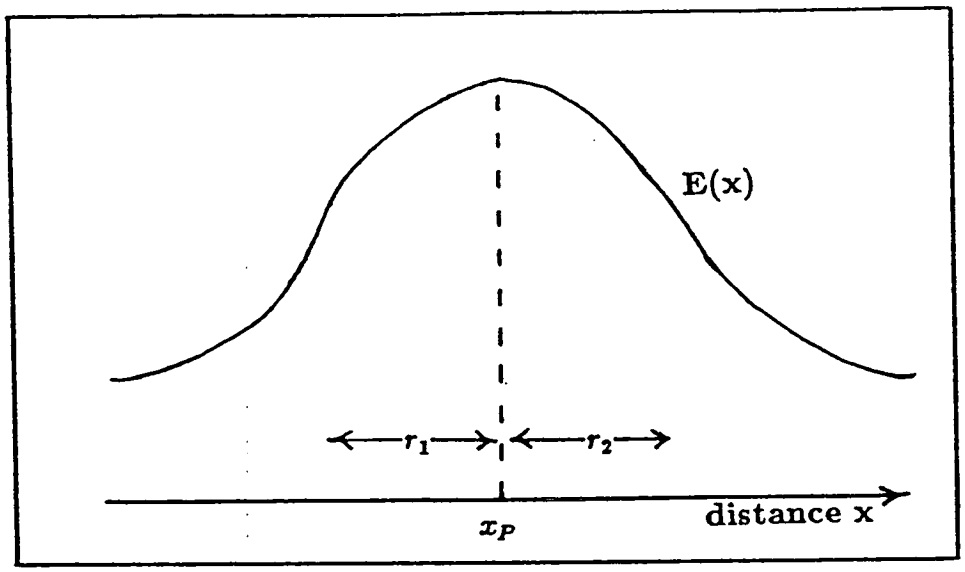


Figure 4–9: *The cylinder test is applied to an image irradiance profile $E(x)$ through a spine. x_P is the peak location and r_1 and r_2 are the distances in the image from x_P to the two flanking edges. These quantities are used to evaluate the upper bound on the second derivative at the peak, (4.5).*

For flattened peaks, the value of D estimated using (4.7) becomes unreliable. The reason for this is that T.V. cameras quantise their signal into a limited number of grey-levels (usually 256), so when the differences G_1 and G_2 are of the order of one grey level, large errors are introduced into the value of D. In this case, it is safest not to apply the test and accept that some wide-profile specularities will be lost, rather than risk marking Lambertian profiles as likely specularities. For the results obtained in this thesis the test is suppressed for any profile in which G_1 or G_2 is less than 6 grey levels. In general, the method of suppression should depend on the noise-levels in the imaging process.

For very narrow profiles the estimates of the image distances r_1 and r_2 are prone to relatively large errors. So the test is also suppressed for profiles in which r_1 or r_2 are less than a minimum width r_0. For the results obtained in this thesis a value of ~ 3 pixels was used for r_0. (In general the value of r_0 should depend on the scale of the specularities sought.)

4.6 Combining the evidence

4.6.1 Introduction

A variety of tests for gathering evidence for specularities can be developed by establishing constraints on Lambertian surface reflection, using the approach proposed in section 4.3. A specularity detector of the type described in section 4.2 must combine the evidence provided by each test to create descriptions of the likely specularities in an image. By itself, no single test can provide a complete description because it only exploits one particular characteristically specular feature. This feature need not be present in all specularities. For example the retinex-based test detects only very bright specularities. In addition, some tests are more reliable than others because some characteristic features provide better evidence than others. Thus in order to create useful descriptions of the likely specularities the detector must take into account the quality of the evidence provided by each test. In some cases the evidence provided by one test reinforces that provided by another. Ideally the specularity detector should:

1. Gather the evidence from a large number of different tests.

2. Simultaneously feed the evidence provided by each test into an inference scheme, tailored to combine the evidence in the most efficacious manner.

4.6.2 A scheme to combine the evidence

Described below is the prototype scheme used to combine the evidence for specularities gathered by the retinex-based, local contrast and cylinder tests. The end product of the scheme is a map of the specular regions within an image. Section 4.6.3 discusses how this specularity map can be adapted for use by other visual processes. Two classes are used to describe points in the map where there is evidence for specularities. These classes are called *good* evidence

and *some* evidence. Where the scheme indicates *good* evidence it is usually safe to apply the shape inference techniques described in chapter 6. Where only *some* evidence is indicated it is unwise to rely on the image region for stereo, motion or recognition processes. This classification scheme is shown in the table below and was arrived at by observing the performance of the tests when applied to a variety of different images.

condition		inference
Retinex-based test passed	→	*good* evidence of specularity
Cylinder test passed	→	*good* evidence of specularity
Local contrast test passed at a very narrow profile	→	*some* evidence of specularity
Any other results	→	no evidence of specularity

Both the retinex-based test and the cylinder test were found to provide what is called *good* evidence—while the evidence provided by the local contrast test proved more circumstantial. So it is only used at very narrow profiles where the cylinder test is inapplicable due to insufficient image resolution (see section 4.5.5): even there it provides only what is called *some* evidence.

Evidence is collected in this way at local maxima in the image. For the local contrast and cylinder tests these are the peaks in the profiles. For the retinex-based test they are the local maxima of the regions that pass the dynamic range test. To obtain a specularity map, the evidence obtained at the local maxima should be associated with the blobs that surround them. In the ideal situation, depicted in figure 4–10, each blob is simply defined by a closed contour in the edge map. So the evidence can be propagated from a local maximum to all points within the contour. If the contour contains maxima with both *good* and *some* evidence the classification *good* predominates.

In reality however, the edge contours are usually broken for a variety of reasons: image noise and the confusion of the edge detector by surface features in the neighbourhood of specularities. A practical method is required to determine the extent of the blob surrounding each local maximum where there is evidence

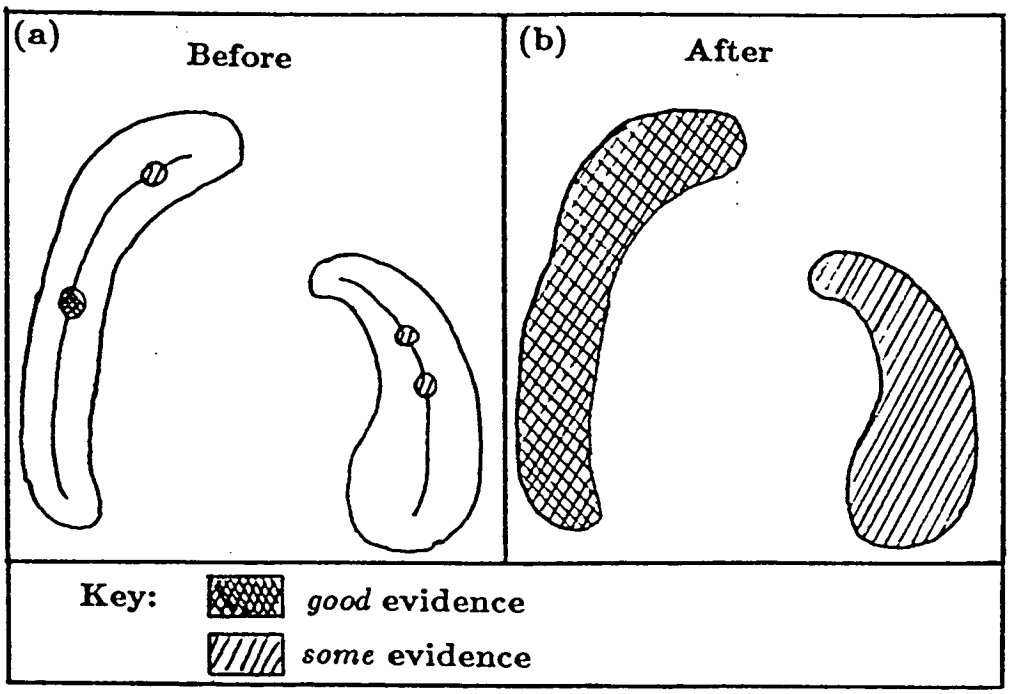


Figure 4–10: *(a) Evidence is combined at the local maxima within each blob. Ideally the extent of each blob is defined by a closed contour. (b) So the evidence can be simply propagated throughout the blob. If the contour contains both* good *and* some *evidence, then the classification* good *predominates.*

of specularity. Once the extent is determined the evidence can be propagated as discussed above. Where an edge contour is missing the image irradiance distribution may be used to determine a blob's extent. To obtain the results in this thesis a very simple method was used. This method involved both the image irradiance distribution and any remnants of the edge contour that were available. At each local maximum where there was evidence for specularity, the following process is initiated:

- The image location is marked as being within the blob.
- The process is then recursively applied to any adjacent image location which is not a part of the contour and at which image irradiance exceeds $2/3^{rds}$ of the peak value.

Although this purely empirical method often produces reasonable results—see section 4.7—there is scope for developing a method on a more principled basis. For example the broken contours might be connected up in some way. This might be achieved by extending the existing contour fragments in a manner that minimises a combination of the contour's length and the tangent gradients around it. Another approach might be to fit a quadratic patch to the image irradiance distribution in the neighbourhood of a maximum. In this case the contour would be determined by the perimeter of the best fitting patch.

4.6.3 Specularity descriptions

The scheme described above provides a map of likely specularities within an image. At each point in the map there is *good*, *some* or no evidence of specularity. This description is in a form that can be used directly by some visual processes—for others a more compact description is appropriate:

- For an example of its direct use, consider a binocular stereo process—as used in chapter 6. It establishes correspondences between edge maps extracted from a right and left view. Once the correspondences are found

the depth of the lines represented by the edge maps can be determined. An edge of a specularity creates a false depth estimate and should be excluded from the edge maps. It can be, by masking out any edge in the neighbourhood of a point marked in the specularity maps. Figure 4–20 shows this being done for a real stereo-pair.

- A compact description is required by the shape from specularity scheme described in chapter 6 for two reasons:

  1. So that monocular inferences can be made.
  2. So that the specular blobs in a stereo pair can be matched—only then can the stereo inferences be made.

Details of both monocular and stereo shape inferences are given in chapter 6. The format used to describe the specular blobs is as below—appendix F gives details of how it is obtained from the specularity map. In addition to its evidence class (*good* or *some*) each blob in the map is described using the following parameters (some of which are shown in figure 4–11):

- the image coordinates of its centroid, (x_c, y_c).
- its approximate dimensions: height h and width w.
- if it is elongated, the direction of its major axis, θ.
- a boolean variable e, that is set if the blob's shape is approximately elliptical—see figure 4–12.

Note: for compactly shaped blobs $h \approx w$, and the information represented by θ is not needed.

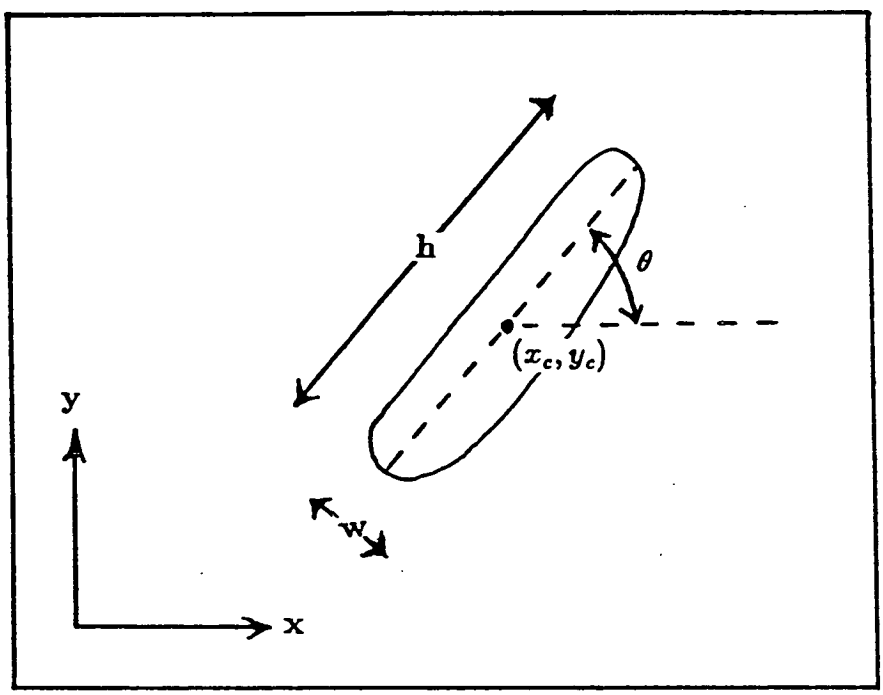


Figure 4–11: *Parameters used to describe a blob: (x_c, y_c)—the location of its centroid, h and w its dimensions and θ the direction of its major axis.*

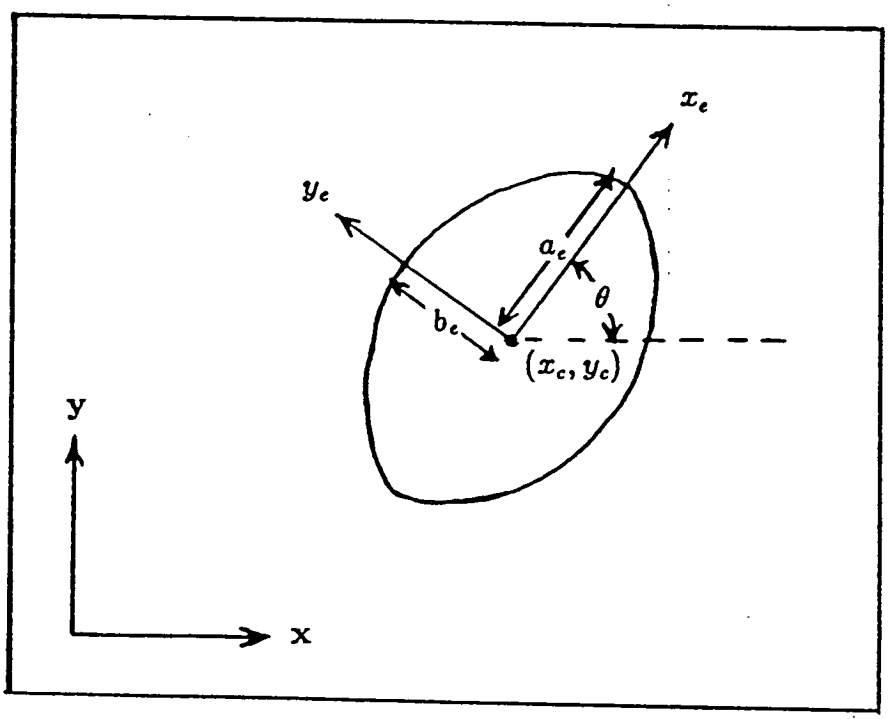


Figure 4–12: *Fitting an ellipse: an ellipse with semi-major and semi-minor axes of lengths a_e and b_e respectively is fitted so that its origin coincides with the centroid of the blob (x_c, y_c). (x, y) are the global image coordinates and (x_e, y_e) are the ellipse's local coordinates. The ellipse's major axis is oriented at angle θ to the global x-axis.*

4.7 Results

4.7.1 Introduction

Figures 4–13 – 4–19 show how the prototype specularity detector, described above, performs on a set of real images. The sections below detail both how the images were obtained and how the various stages of the scheme were applied. The results are discussed in section 4.7.8.

4.7.2 Image acquisition

Each image was represented as a 256 square matrix of elements each themselves represented as an 8 bit grey-level. Care was taken to ensure that the image signal was a good approximation to relative image irradiance—so that the tests could be applied directly. This was achieved by using a camera with a linear response, no a.g.c. and by avoiding saturation of the specularities[14]. Real images were acquired in two ways:

1. A Link-Electronics 109 vidicon camera was used for images in figures 4–13 – 4–19. These images suffer a slight distortion: 'vertical banding' near the left borders. However, this does not unduly hamper the specularity detection process.

2. A stereo rig containing two Panasonic WV-CD50 CCD cameras was used for the stereo pair of images in figure 4–20. All other real stereo pairs shown in this thesis were obtained using the same rig.

In both cases the dynamic range of the signal was about 6 bits.

---

[14]Note: the photographs of these image shown in the figures have been enhanced for display purposes—the tests were applied to the unenhanced signal.

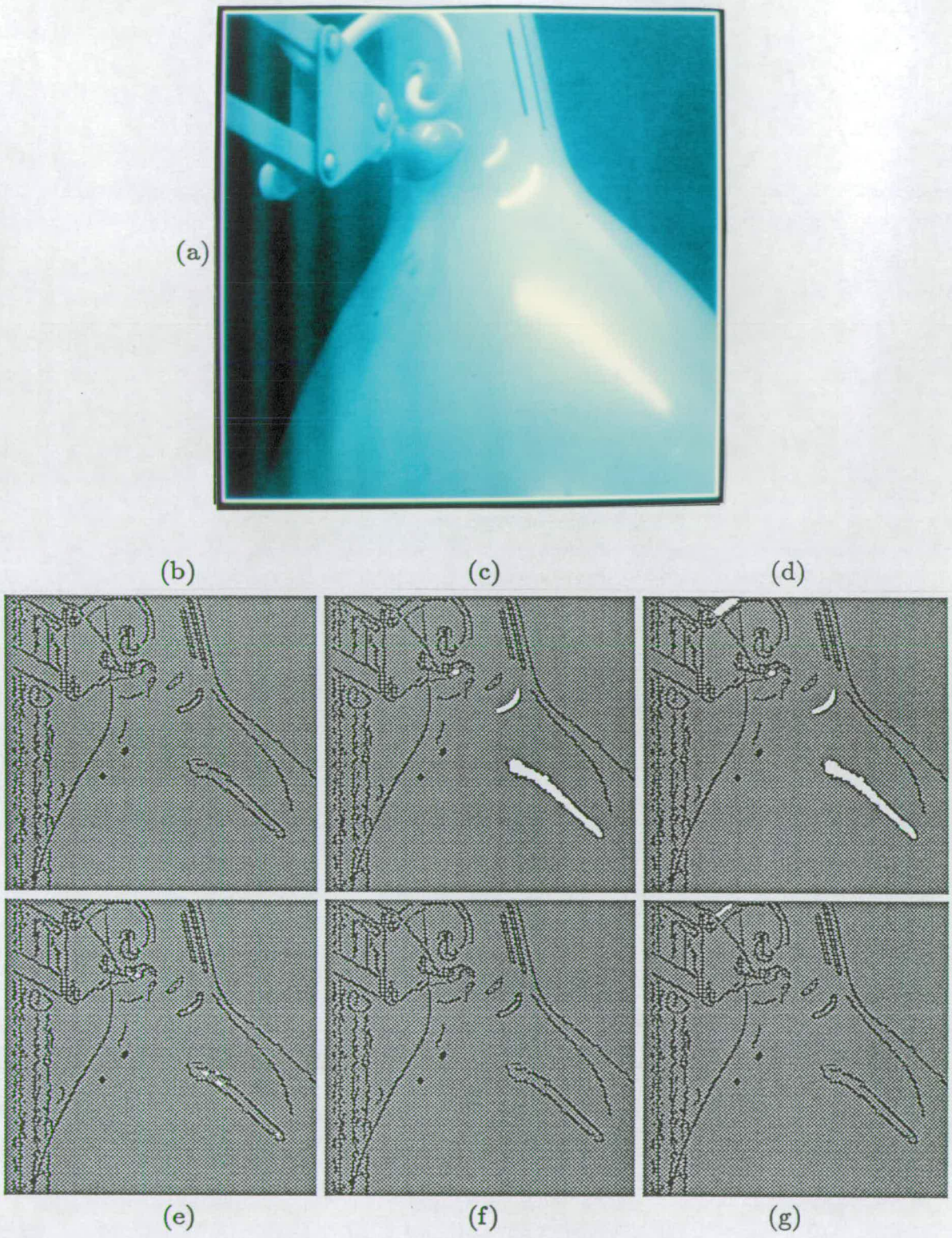


Figure 4–13: *(a) An image of* **an anglepoise lamp.** *(b) The edge map—edges in black. Regions (in white) where the prototype scheme finds evidence for specularities: (c)* good *and (d)* good *plus* some *evidence (after propagation). Local maxima (in white) where tests are passed: (e) cylinder, (f) retinex-based and (g) local contrast test.*

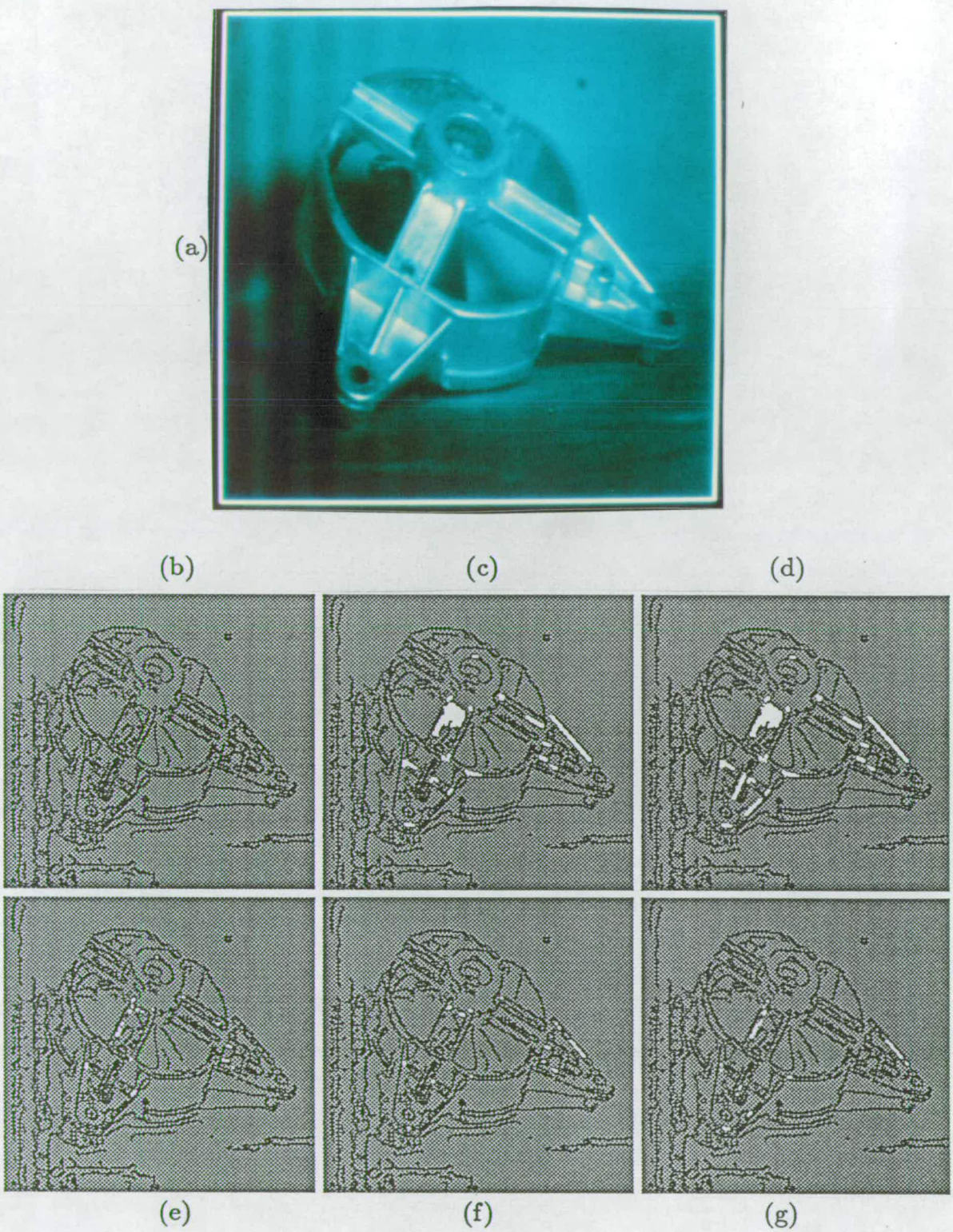


Figure 4–14: *(a) An image of* **an industrial casting.** *(b) The edge map—edges in black. Regions (in white) where the prototype scheme finds evidence for specularities: (c)* good *and (d)* good *plus* some *evidence (after propagation). Local maxima (in white) where tests are passed: (e) cylinder, (f) retinex-based and (g) local contrast test.*

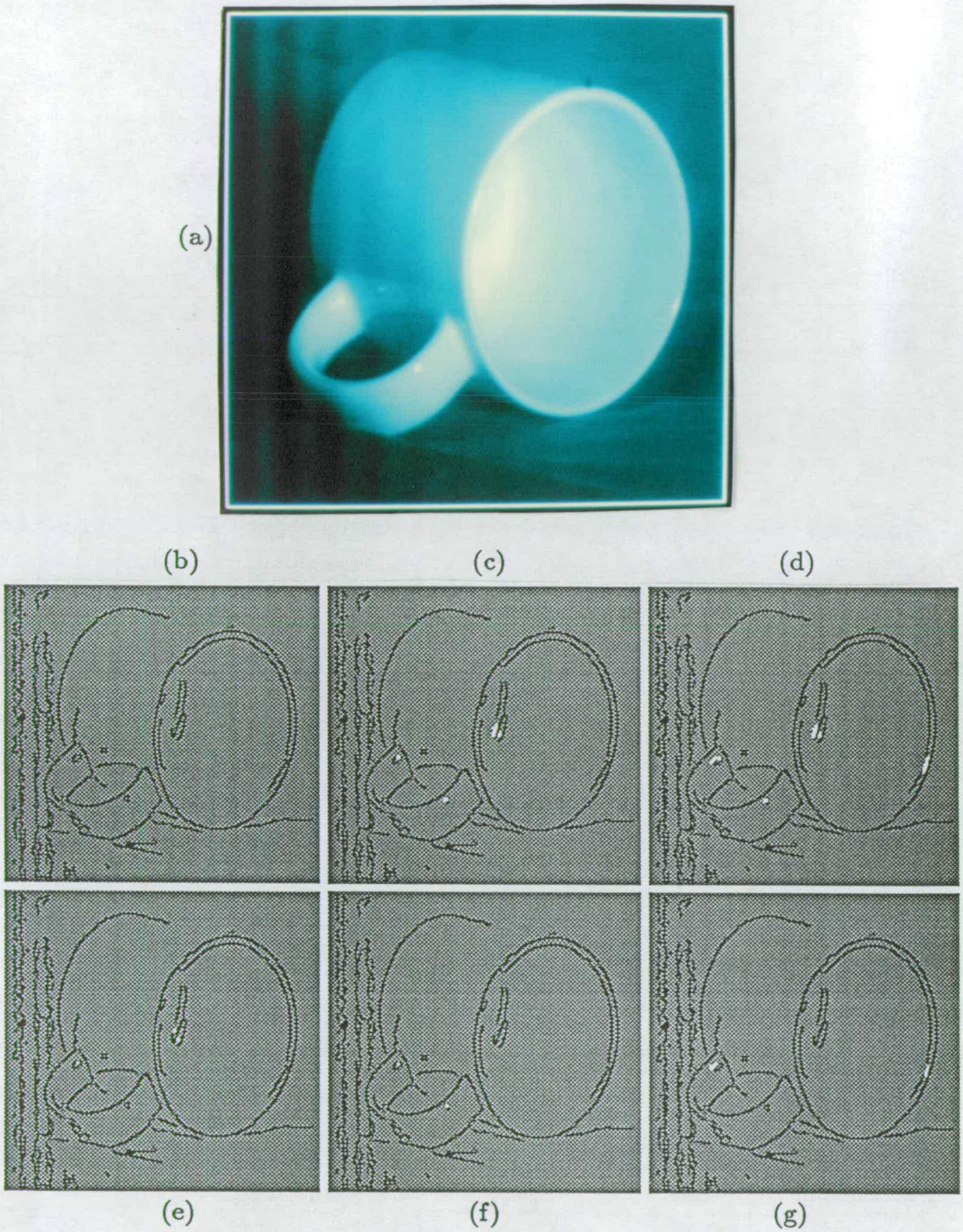


Figure 4–15: *(a) An image of* a mug. *(b) The edge map—edges in black. Regions (in white) where the prototype scheme finds evidence for specularities: (c)* good *and (d)* good *plus* some *evidence (after propagation). Local maxima (in white) where tests are passed: (e) cylinder, (f) retinex-based and (g) local contrast test.*

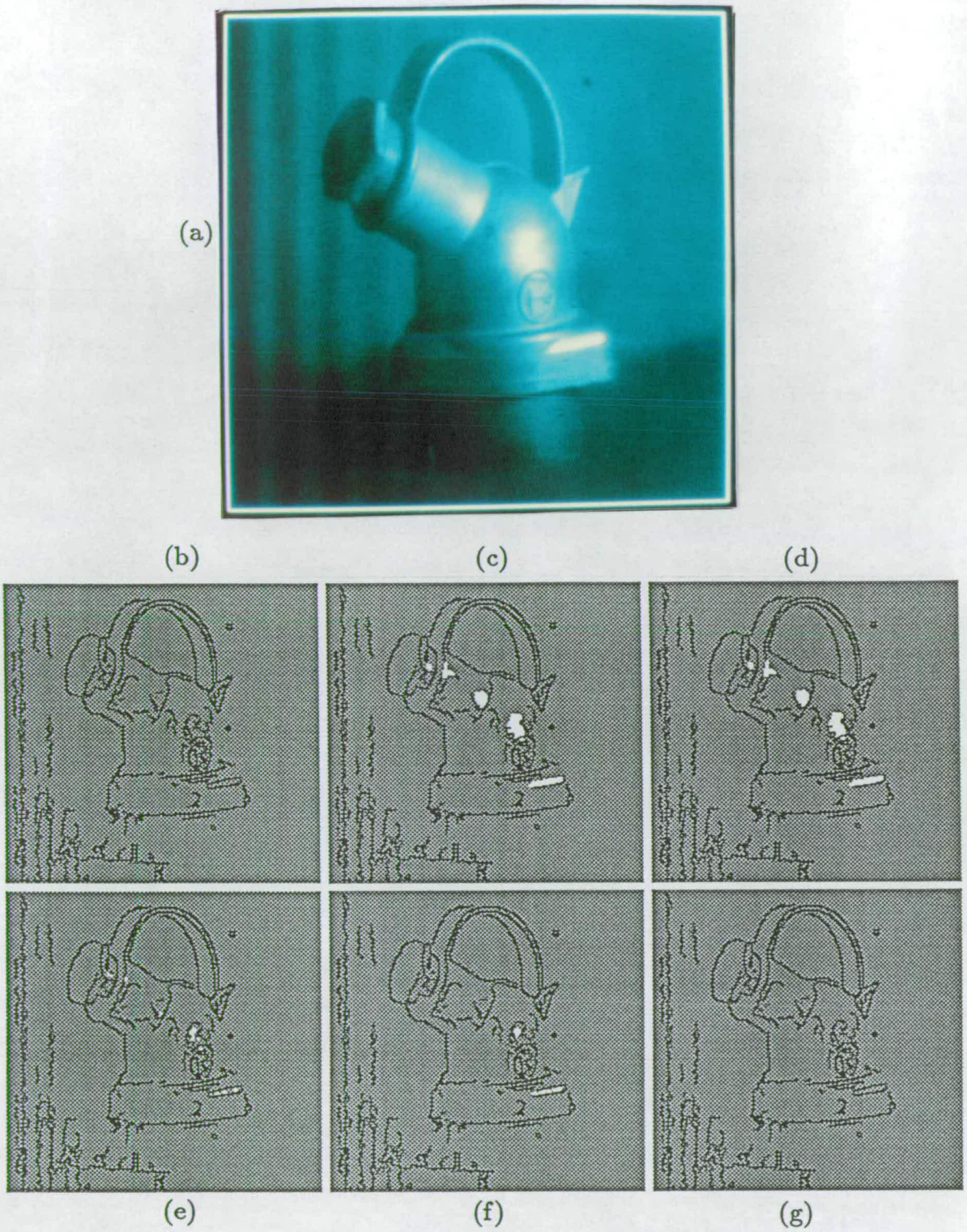


Figure 4–16: *(a) An image of* **an industrial plug.** *(b) The edge map—edges in black. Regions (in white) where the prototype scheme finds evidence for specularities: (c)* good *and (d)* good *plus* some *evidence (after propagation). Local maxima (in white) where tests are passed: (e) cylinder, (f) retinex-based and (g) local contrast test.*

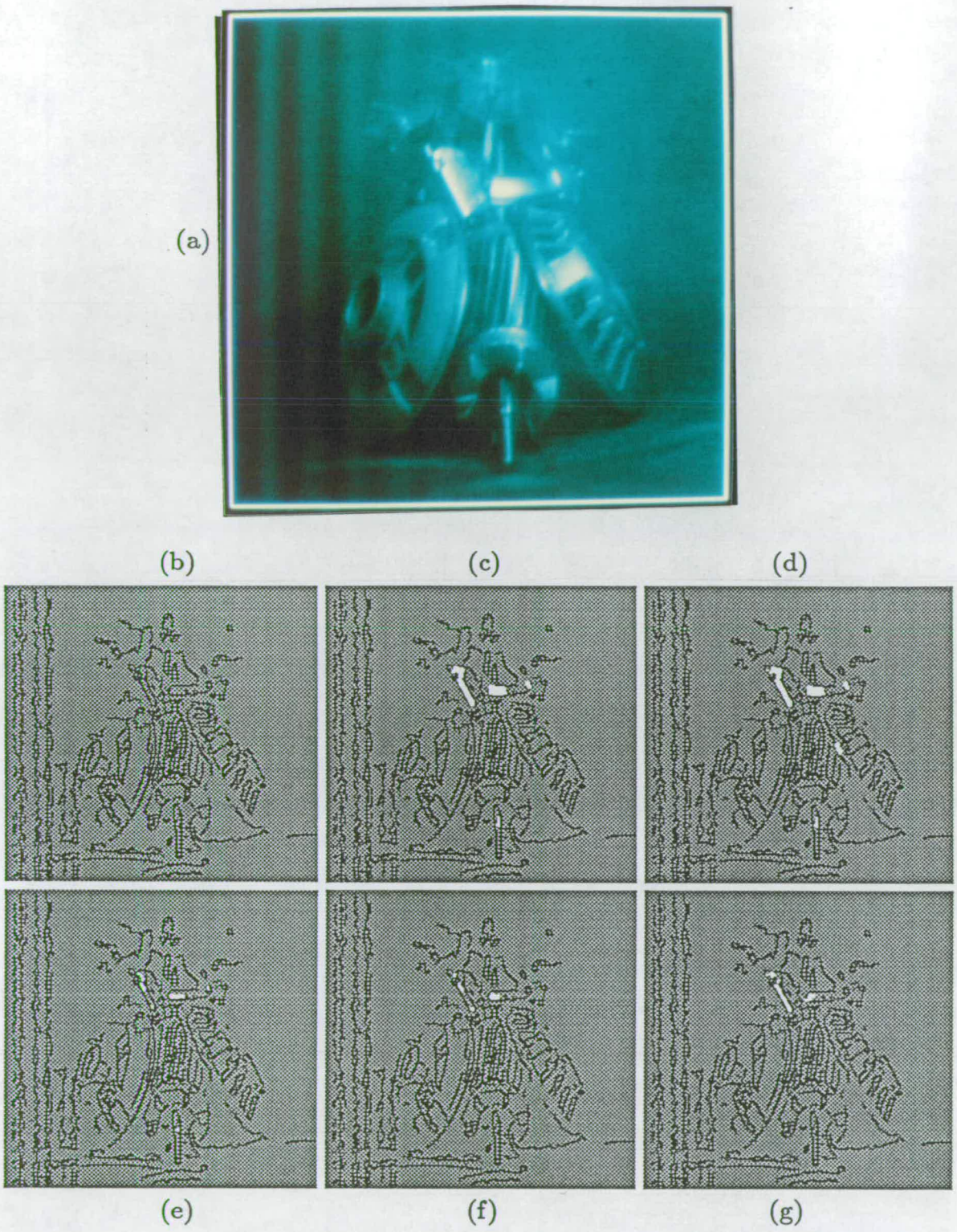


Figure 4–17: *(a) An image of* **a jumble of parts.** *(b) The edge map—edges in black. Regions (in white) where the prototype scheme finds evidence for specularities: (c)* good *and (d)* good *plus* some *evidence (after propagation). Local maxima (in white) where tests are passed: (e) cylinder, (f) retinex-based and (g) local contrast test.*

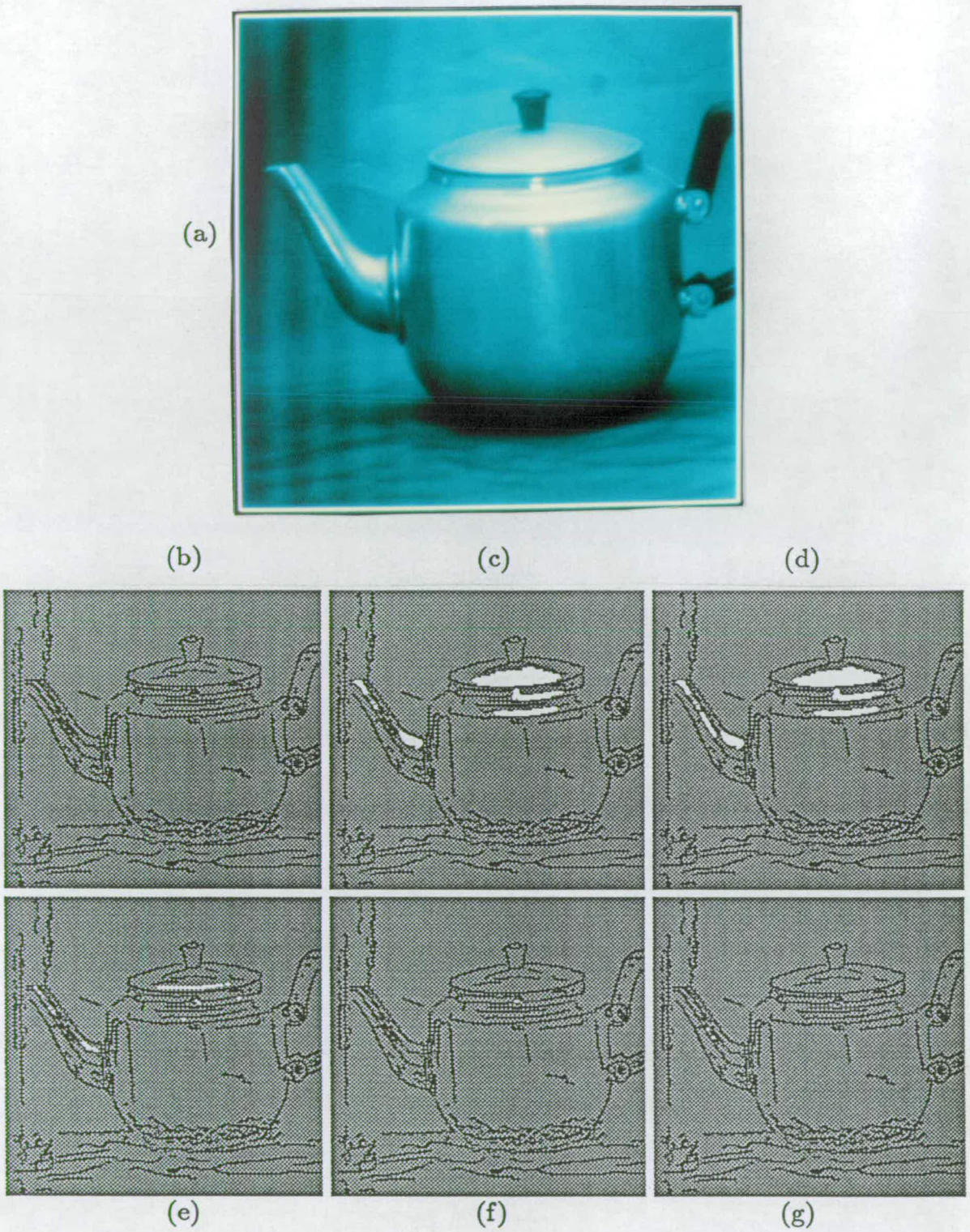


Figure 4–18: *(a) An image of* **a teapot.** *(b) The edge map—edges in black. Regions (in white) where the prototype scheme finds evidence for specularities: (c)* good *and (d)* good *plus* some *evidence (after propagation). Local maxima (in white) where tests are passed: (e) cylinder, (f) retinex-based and (g) local contrast test.*

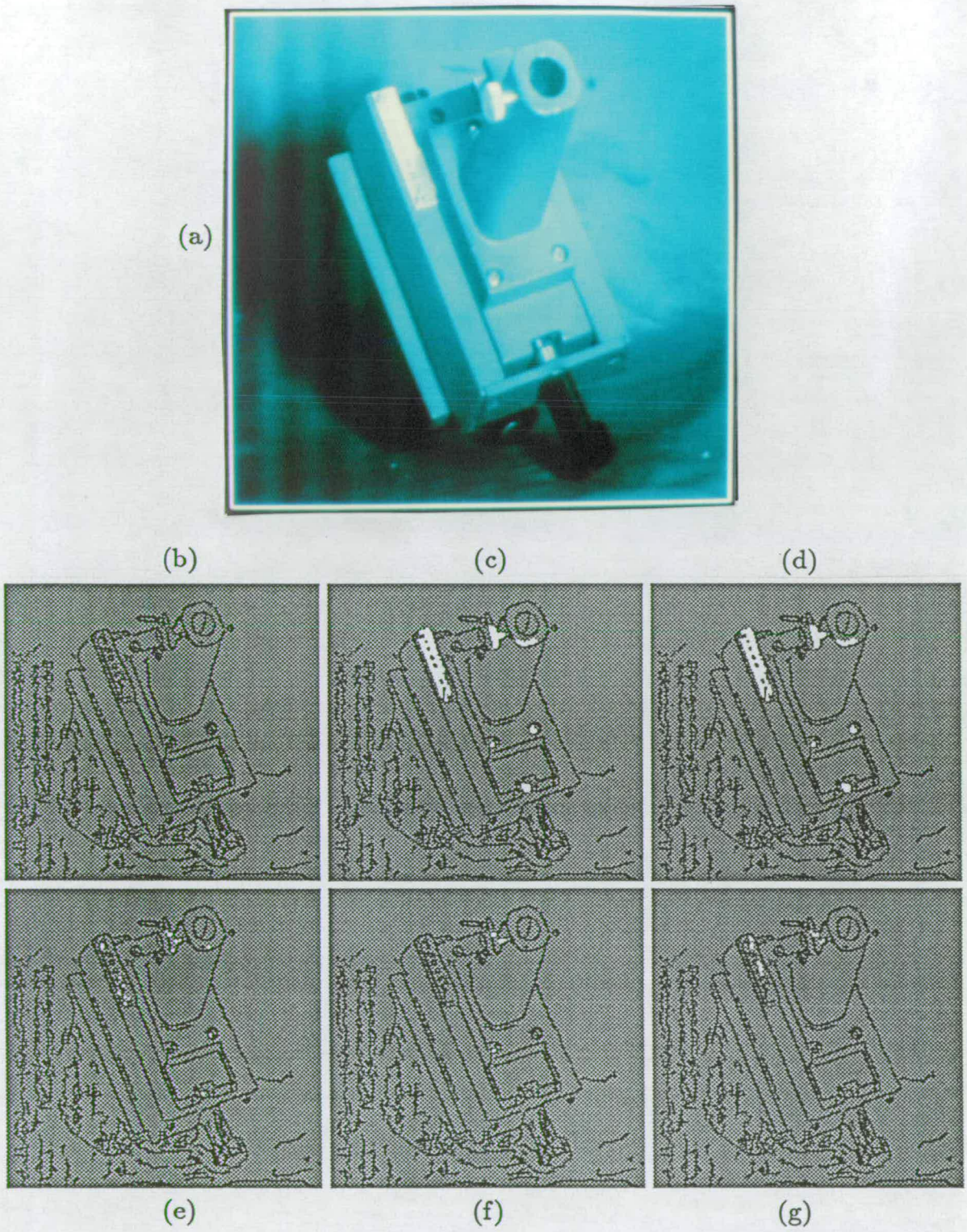


Figure 4–19: *(a) An image of* **a clamp stand.** *(b) The edge map—edges in black. Regions (in white) where the prototype scheme finds evidence for specularities: (c)* good *and (d)* good *plus* some *evidence (after propagation). Local maxima (in white) where tests are passed: (e) cylinder, (f) retinex-based and (g) local contrast test.*

4.7.3 Profile extraction

The profiles along which it was appropriate to apply local tests were extracted from each image using the method detailed in appendix D. This method must be supplied with a description of the locations and directions of the edges in each image. A local implementation of the Canny operator was used for this purpose. Details of the Canny operator are given in [24]: here only the parameters employed are listed. The smoothing width was $\sigma = 1$ and hysteresis thresholds corresponded to 20% (high) and 2% (low) contrasts between adjacent pixels. The edges obtained using these parameters correspond roughly to those visible in the images. In each of figures 4–13 – 4–20, (b) shows the edge map used.

4.7.4 Retinex preprocessing

Before applying the retinex-based test, each image was preprocessed using a retinex - as described in section 4.4.3. Details of the scheme used are given in the paper in appendix K. The threshold operation used by the retinex is detailed in appendix C.

4.7.5 The tests

The following three tests were applied to each image:

1. The cylinder test—as in equation (4.6).

2. The retinex-based test—as in equation (4.2).

3. The local contrast test—as in section 4.5.4.

The local maxima at which each of these tests fired are shown respectively in (e), (f) and (g) of figures 4–13 – 4–19.

4.7.6 Combining evidence

The scheme described in section 4.6.2 was used to combine and propagate the evidence gathered by each test. In figures 4–13 – 4–19 (c) shows regions where this scheme indicates *good* evidence for specularities and (d) where either *good* or *some* evidence is indicated.

4.7.7 An application: masking out specular edges

Figure 4–20 shows an example of the direct use of the specularity maps obtained using the scheme. For the reasons described in earlier it is useful to mask out the specular edges from the edge map. This was achieved by simply masking out any edge located within 4 pixels of any detected specularity. Figure 4–20 shows a stereo pair (a), its edge maps (b) it specularity maps (c) and the edge maps after masking (d). Figure 4–21 shows the improvement in depths (near the large specularity) obtained by binocular stereo after excising known specular features. The apparent fragmentation of the smooth surface is eliminated.

4.7.8 Evaluation of the results

The images shown in figures 4–13 – 4–19 are fairly representative of those viewed by an industrial vision system. The results obtained for theses images provide an indication of how the detection scheme is likely to perform in an industrial scenario. Ideally the scheme ought to be implemented as part of a real time vision system[15] and tested for a wider variety of environments. However it is still useful to evaluate how the scheme and its constituent tests perform for the limited image-set to which it has been applied. Three criteria are used when making the evaluation:

1. Are *genuine* specularities detected?

---

[15]like that proposed in [18].

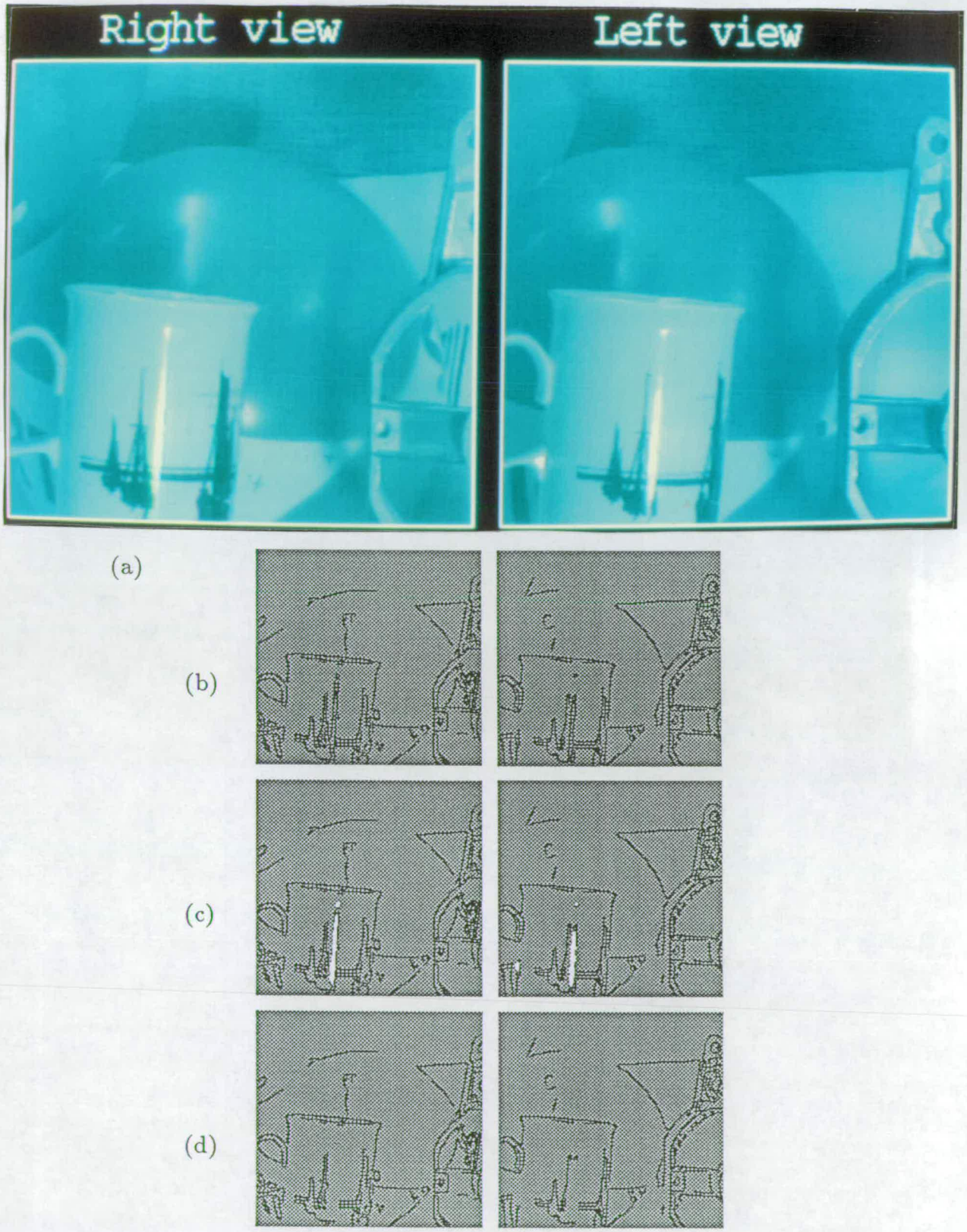


Figure 4–20: *(a) A stereo pair of images (aligned for cross-eyed viewing). (b) The edge maps—edges in black. (c) Detected specularities—in white. (d) The edge maps with specular edges removed.*

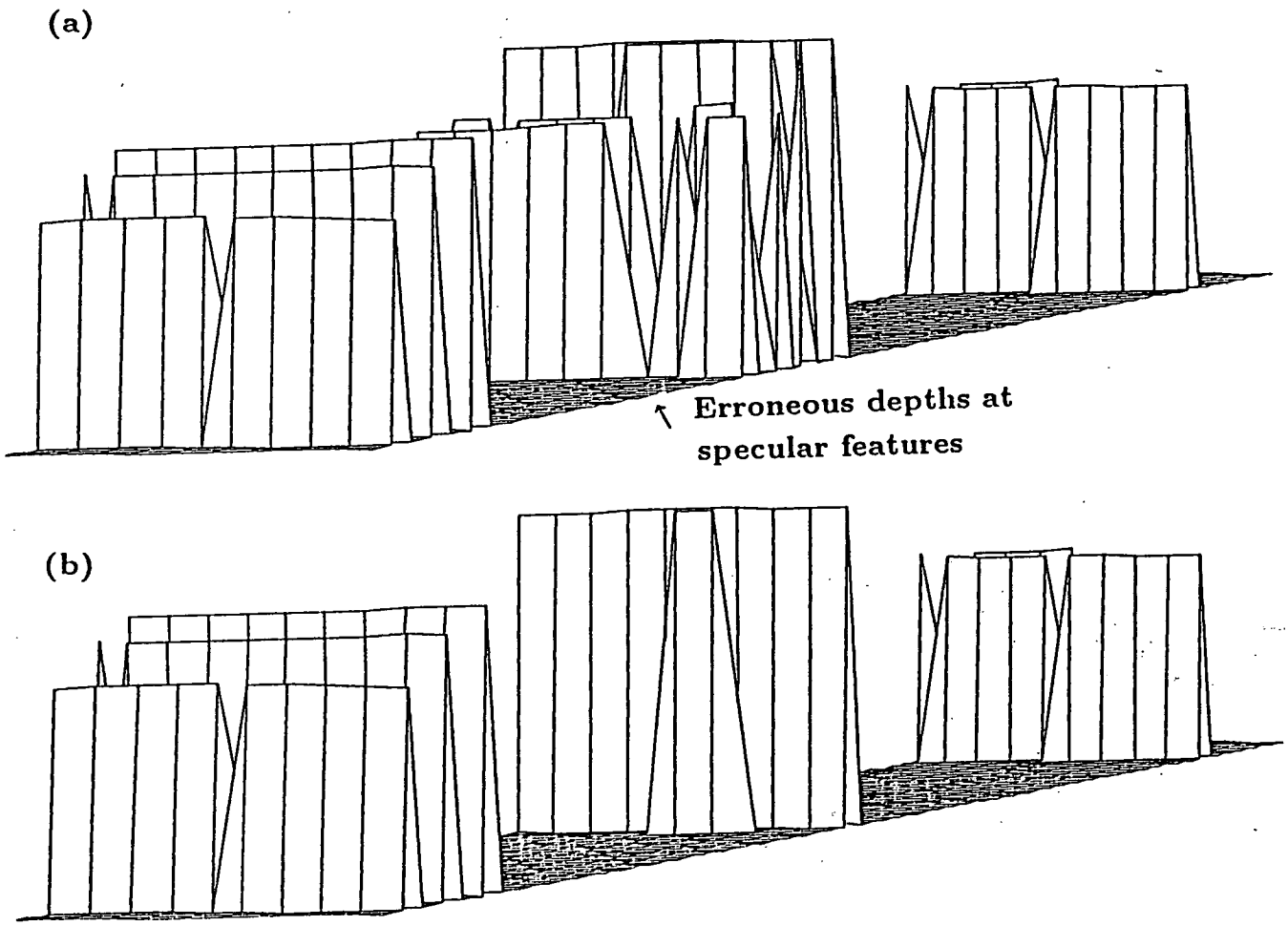


Figure 4–21: *Surface plots of part of the sparse depth map from the previous figure, (a) before and (b) after excising the specular features.*

2. Are *genuine* specularities missed?

3. Are *false* specularities detected?

For prominent specularities and dark matt regions there is often little doubt what constitutes a *genuine* specularity. The distinction is more difficult to make for other image regions. Two factors confound a naked-eye comparison of the image to the results obtained:

1. The human interpretation of images is subjective.

2. The interpretation of specularities depends critically upon the method used to display the images. In particular, when photographs are used—as in this thesis—the interpretation may differ from that obtained using a correctly calibrated T.V. monitor[16].

Ideally an objective method of comparison should be used—one possibility could be to use the polarisation of light to distinguish specular regions in a test control. Here, as no such method is available, the comparison has been made as objective as possible.

The figures 4–13 – 4–19 show that all three tests can and often do detect *genuine* specularities. No single test detects all the *genuine* specularities in any of the images—this is to be expected, as the specular feature exploited by any individual test need not be present in every specularity. Neither the retinex-based test nor the cylinder test create any *false specularities*. However the local contrast test does: notably on the handle of the mug in figure 4–15 and perhaps also on the plastic cord in figure 4–13. This reinforces the interpretation—already justified—of the local contrast test as providing only partial evidence for specularities and being susceptible to false alarms.

---

[16]See [25] for a discussion of monitor calibration.

On most occasions the evidence is satisfactorily propagated from the local maxima into the surrounding blobs—using the method described in section 4.6.2. Occasionally the evidence is propagated too far: e.g. at the top of the object in figure 4–19[17]. This occurs where the edge contour surrounding the specularity is incomplete. Some remedies for this problem were suggested in section 4.6.2.

Barring this problem, the scheme successfully combines the evidence provided by the tests and creates a specularity map. However the maps obtained for each image are all incomplete: at least one *genuine* specularity is missed in each case. Thus there is room for improvement. The specularities that are missed are often both too dim for the retinex-based test and too narrow for the cylinder test—e.g. those on the left spoke of the casting in figure 4–14. New tests need to be developed in order to detect these specularities. In some special environments the upper bounds imposed by some of the existing tests may be found to be too conservative. For example, if the dynamic range of albedo is less than 10 then the bound used by the retinex-based test may be reduced to detect more *genuine* specularities without introducing any *false* ones.

[17]The largest blobs in figures 4–14 and 4–18 may also be regarded as overspilling their bounds—although this is a more subjective judgement.

Chapter 5

Inferring surface shape: background

5.1 Introduction

The task of inferring the shape of a surface from visual cues is complex. It often requires the combination of disparate kinds of information. For example: binocular stereo, silhouettes, texture and shading each supply some—but seldom all—of the necessary information. For a glossy surface, specularities are often a powerful cue. The scheme described in chapter 6 infers surface shape using specularities. In this chapter existing approaches to the inference of shape are reviewed—concentrating in particular on the role of specularities. To begin with basic representations for surface shape are introduced.

5.1.1 Surface depth

A surface may simply be represented by the depth of each point visible upon it. Traditionally a depth map $z(x, y)$ has been used for this purpose. Depth maps have direct application in robotics tasks such as path planning and collision avoidance. However further processing is necessary to obtain more concise descriptions if they are to be used for more complex tasks such as object recognition. If the depth map is sparse (as is typical of those obtained from binocular stereo), further processing may also be necessary—in the form of interpolation or surface reconstruction. Specularities appear to lie above or below but generally not on the reconstructed surface. In chapter 6, a formula for surface depth at a specularity is obtained, relative to the depth of a nearby point. For

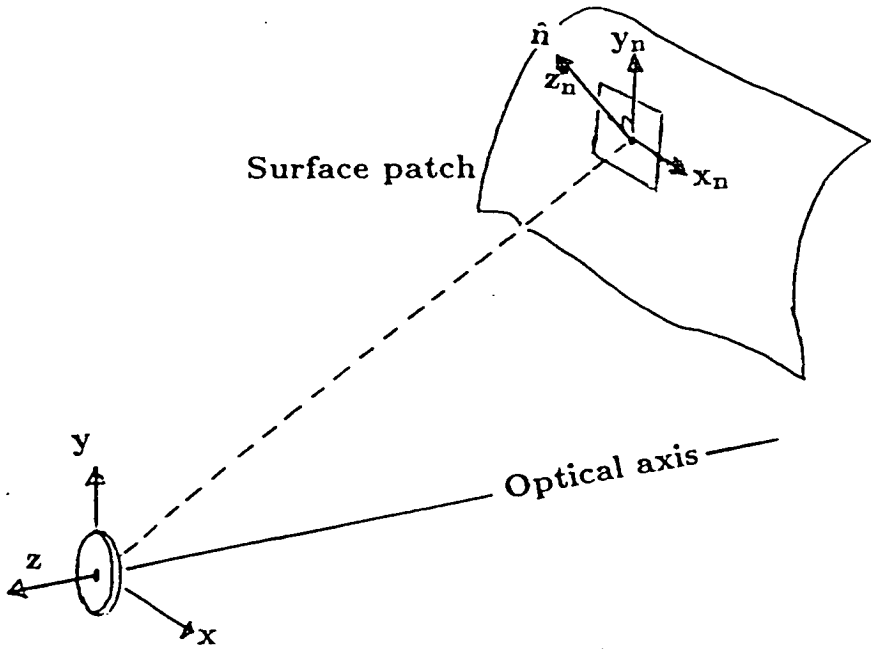


Figure 5–1: *Surface depth may be defined using both the camera coordinate system* (x, y, z) *or the local surface system* (x_n, y_n, z_n). *The* z_n*-axis lies along the local surface normal* $\hat{\mathbf{n}}$.

this purpose depth estimates are provided by binocular stereo—see section 5.2. Figure 5–1 depicts the two different 3-D right-handed cartesian coordinate system used to define depth:

1. The (x, y, z) camera system for which the z-axis along the optical axis of the camera. The x and y axes are horizontal and vertical respectively. In this system the visible surface is a function $z(x, y)$.

2. The (x_n, y_n, z_n) local system for which the z_n-axis is along the local surface normal $\hat{\mathbf{n}}$. In this system the visible surface is a function $z_n(x_n, y_n)$.

Both are used below.

5.1.2 Surface orientation

Local surface orientation is defined using $\hat{\mathbf{n}}$: the normal direction to the local tangent plane—shown in figure 5–1. Surface orientation is only an applicable representation for smooth[1] surfaces with well defined tangent planes. Two different specifications of $\hat{\mathbf{n}}$ are used below:

- Angles of slant σ and tilt τ which are defined:

$$\begin{aligned} \cos\sigma &= \hat{\mathbf{n}} \cdot \hat{\mathbf{V}}; \quad 0 \leq \sigma \leq \pi/2, \\ \tan\tau &= n_y/n_x; \quad 0 \leq \tau \leq 2\pi, \end{aligned} \tag{5.1}$$

  where n_x and n_y are the x and y components, in the camera frame[2], of $\hat{\mathbf{n}}$ respectively and $\hat{\mathbf{V}}$ is the viewing direction.

- Gradients p and q which are defined as:

$$(p,q) = \left(\frac{\partial z_n}{\partial x_n}, \frac{\partial z_n}{\partial y_n}\right) \text{ and } \hat{\mathbf{n}} = (n_x, n_y, n_z) = \frac{(p,q,-1)}{\sqrt{1+p^2+q^2}}. \tag{5.2}$$

Both specifications have been applied to a variety of problems: e.g. (σ,τ) for texture analysis [109] and (p,q) for shading analysis—discussed later. Surface orientation data by itself can be used to identify known objects in cluttered scenes—e.g. bin picking [59]. Used in conjunction with other information it can produce surface descriptions suitable for more difficult model matching tasks. For known source position, surface orientation can of course be obtained directly from an observed specularity.

---

[1]Here *smooth* means smooth at a macroscopic scale, so it applies to the microscopically rough surfaces referred to in section 2.4.4.

[2]Slant is traditionally defined for parallel projection where $\hat{\mathbf{V}}$ is aligned with the camera's optical axis (e.g. [109]).

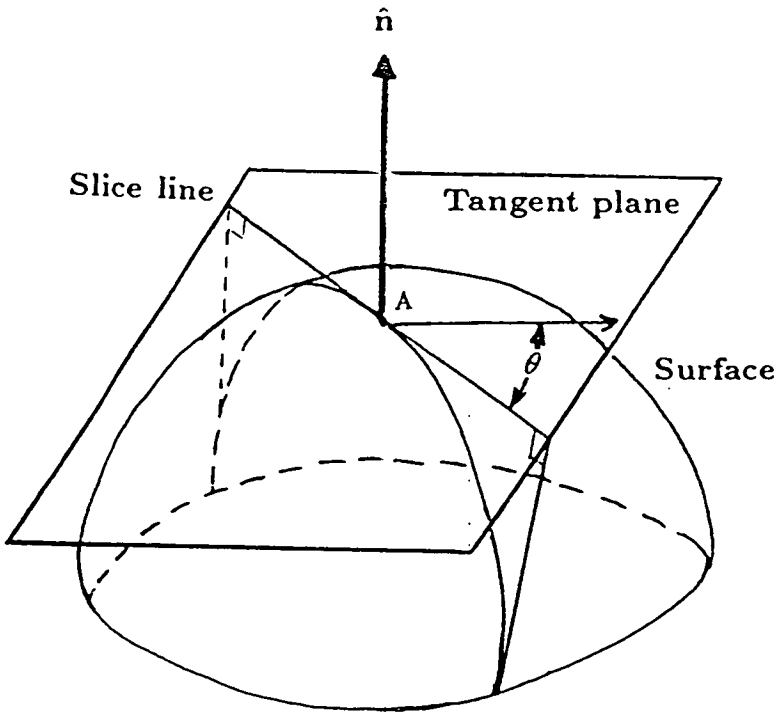


Figure 5–2: *Geometry used to define local curvature at point A on a smooth surface.*

5.1.3 Surface curvature

Here curvature is introduced with reference to a simple shape: a sphere[3]. Every point on the surface of a sphere lies at the same distance (radius r) from the centre of curvature (the centre of the sphere). The *radii of curvature* at any point on the sphere are simply r and the *magnitude of curvature* is $k = 1/r$. The *sign of the curvature* depends on whether the sphere appears convex or concave. Curvature is positive for a convex surface and negative for a concave one[4]. In the limit as $r \to \infty$, $k \to 0$ and the surface is locally planar. The local curvature of an arbitrarily shaped, smooth surface at any point A upon it, is defined by the rate at which it slopes away from the local tangent plane. To make this

[3]For a more formal introduction in terms of differential geometry see [32,79].

[4]Some texts define curvature with the opposite sign used here.

more precise, consider a slice through the surface at A—as shown in figure 5–2. This slice is a curve in space, along which the surface intercepts a plane—a plane that passes through A and contains $\hat{\mathbf{n}}$, the local surface normal. The orientation of the slice is defined by the angle θ—i.e. that between the tangent vector of the slice curve at A and the axis. Thus θ quantifies the rotation of the slice about $\hat{\mathbf{n}}$.

If the surface is locally spherical at A, it slopes away from the tangent plane at a constant rate k, in all directions (all θ). So that $k(\theta)$ is constant. In general, surfaces are not locally spherical and $k(\theta)$ is not constant. Nevertheless, $k(\theta)$ is constrained for locally smooth surfaces. It varies monotonically between a maximum and minimum value: k_1 and k_2 respectively. These values lie along orthogonal directions, i.e. if $k_1 = k(\theta_0)$ then $k_2 = k(\theta_0 \pm \pi/2)$. In fact local curvature upon any smooth surface is defined by the three parameters (k_1, k_2, θ_0).

The hessian matrix H provides a concise description of curvature which is directly related to (k_1, k_2, θ_0). It is of particular use in the analysis of shape from specularities described in section 5.6 and chapter 6. Expressed in terms of (k_1, k_2, θ_0);

$$\mathsf{H} = \mathsf{R}(\theta_0) \begin{pmatrix} k_1 & \cdot \\ \cdot & k_2 \end{pmatrix} \mathsf{R}(-\theta_0), \tag{5.3}$$

where $\mathsf{R}(\psi)$ is the matrix describing a rotation through ψ about $\hat{\mathbf{n}}$:

$$\mathsf{R}(\psi) = \begin{pmatrix} \cos\psi & \sin\psi \\ -\sin\psi & \cos\psi \end{pmatrix}. \tag{5.4}$$

H is also the matrix the 2nd partial derivatives of surface function $z_n(x_n, y_n)$:

$$\mathsf{H} = \begin{pmatrix} \frac{\partial^2 z_n}{\partial x_n^2} & \frac{\partial^2 z_n}{\partial y_n \partial x_n} \\ \frac{\partial^2 z_n}{\partial x_n \partial y_n} & \frac{\partial^2 z_n}{\partial y_n^2} \end{pmatrix}. \tag{5.5}$$

Although, in this form H has four elements there remain only three unknowns of curvature because, by definition, $\frac{\partial^2 z_n}{\partial x_n \partial y_n} = \frac{\partial^2 z_n}{\partial y_n \partial x_n}$. The *principal curvatures* k_1 and k_2 are intrinsic surface properties independent of any particular viewing geometry. As such they are particularly useful for model matching. Table 5–1

Condition	Classification	Example surface
sign(k_1)=sign(k_2)	elliptic	ellipsoid
sign(k_1)≠sign(k_2)	hyperbolic	saddle
$k_1 = 0$ or $k_2 = 0$	parabolic	cylinder

Table 5–1:

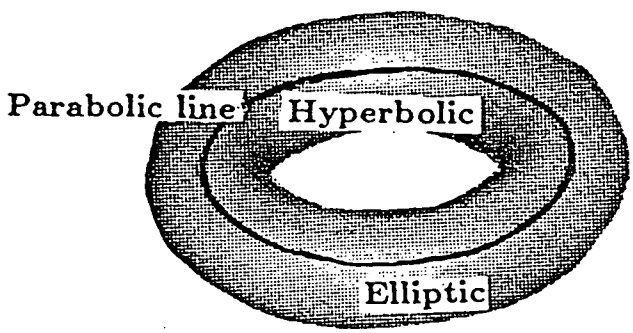


Figure 5–3: *A torus consists of one elliptic and one hyperbolic region separated by parabolic lines.*

shows a classification of local surface shape in terms of the signs of k_1 and k_2. Smoothly curving surfaces are tessellated into regions containing either only elliptic or hyperbolic points (*elliptic* or *hyperbolic regions*). These regions are separated by closed curves of parabolic points (*parabolic lines*). Parabolic lines and the classification into elliptic and hyperbolic regions provide an attractive representation of the surface structure of an object. For example a torus can be represented by a graph consisting of one elliptic and one hyperbolic region separated by parabolic lines—see figure 5–3. Such a representation provides a useful framework with which to fit surfaces to model surfaces. The curvature at all points on the model need not be specified—rather a range of permissible curvatures within each region might be used. So that when visual cues provide some curvature information, even if only at a single point within the region, the fit of surface to model can be evaluated. For example: the curvature information

provided by a single specularity can be enough to determine whether an elliptic region is convex or concave.

Koenderink and van Doorn [73] provide an analysis of the surface representation described above. They relate it to shading features that are invariant with change of viewing and lighting geometry. In particular they note that:

- Specularities[5] are a particular type of singularity of the shading field which tend to cling to surface points of high curvature.

- As viewing or lighting changes specularities are created or annihilated in pairs on parabolic lines. At the moment of their creation or destruction they move transversely to such lines.

These observations constitute qualitative constraints on the formation of specularities. Such constraints might be exploited by specularity detection schemes. Indeed the fact that specularities cling to 'spines' of high curvature is already used by the detector described in chapter 4.

5.2 Obtaining surface depth

5.2.1 Introduction

Existing methods to estimate surface depth fall into two categories:

- passive methods that estimate the depth from conventional imagery.

- active methods which bombard the scene with a signal and then measure its position (e.g laser [97]) of time of flight (e.g. sonar).

---

[5]Referred to in [73] as brilliant points.

Here, only passive methods are considered: active methods can provide dense depth maps for specific robotics tasks [36,49]. Most passive methods require two or more images of a scene in order to operate—e.g. binocular stereo. However monocular depth cues do exist and have been investigated, e.g: focus-blur [89], texture gradients and perspective [41]. Binocular stereo estimates depth using two views of a scene (a stereo-pair)—with similarities to human vision. Other methods that use three or more images have also been developed [5,88]. The scheme in chapter 6 employs an existing binocular stereo algorithm to estimate surface depth at specularities. It is reviewed below. Note however, the depth is estimated for surface features nearby each specularity, because the method is unsuitable for specular features.

5.2.2 Binocular stereo

Seen from two different vantage points the same scene forms two similar images. Similarities arise because the two images are the projections of roughly the same set of surface points. Differences are due to separation of the vantage points. Binocular stereo uses the differences to estimate surface depth. It consists of two steps:

1. Identifying pairs of points in the left and right images that correspond to the same physical locations. This is the *correspondence problem* and the matched points are called *conjugate pairs*.

2. Using *triangulation* to estimate the depth of the surface point that they represent. This requires the camera geometry to be known.

The correspondence problem: The first step is the hardest. Two cameras may be positioned in many different ways in order to produce two images suitable for binocular stereo processing. However, as long as their relative position and orientation are known the images can be *rectified* [59] so that a standard simple camera geometry is applicable. This geometry is shown in figure 5–4: the optical axes of the left and right cameras are parallel. In this case conjugate

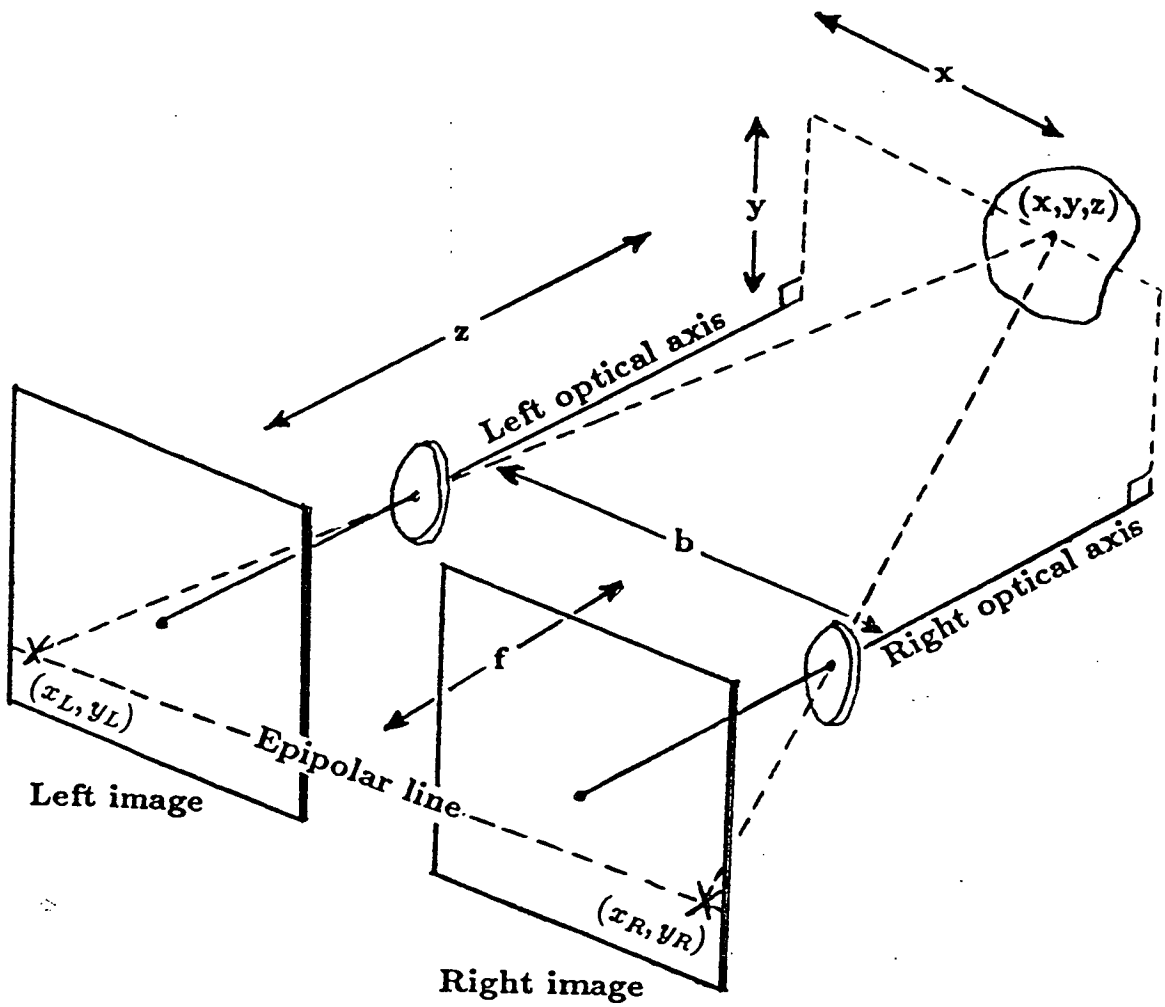


Figure 5–4: *Standard rectified camera geometry for binocular stereo. Optical axes of left and right cameras are parallel, with a base-line separation b. Both cameras have focal length f. Surface point (x, y, z) projects onto left and right images at (x_L, y_L) and (x_R, y_R) respectively.*

pairs have identical vertical coordinates in their respective images—i.e. they lie on the same raster. This is the *epipolar constraint*[6]. It simplifies the correspondence problem by reducing the search for each match. It has become 1-D rather than 2-D. Algorithms have been developed to provide global solutions to this problem [8,10]. In generally only prominent surface features can be identified and matched. The PMF algorithm [91,92] used in this thesis matches edge segments. The density of the depth map obtained is limited by the density of prominent edges. For a smooth untextured surface patch, depth estimates are only obtained at the edges that bound it. To obtain surface descriptions within such a patch another approach is necessary—see sections 5.3 and 5.4.

Triangulation: The geometry in figure 5–4 shows how depth is estimated from a matched conjugate pair. A point (x, y, z) on a surface[7] projects onto the left image at (x_L, y_L) and onto the right image at (x_R, y_R).

Simple trigonometry provides:

- The epipolar constraint:

$$\frac{y_L}{f} = \frac{y}{z} \text{ and } \frac{y_R}{f} = \frac{y}{z} \text{ therefore } \quad y_L - y_R = 0, \tag{5.6}$$

  where f is the focal length of both cameras.

- The depth, z: estimated using $x_L - x_R$, the *horizontal disparity*:

$$\frac{x_L}{f} = \frac{x}{z} \text{ and } \frac{x_R}{f} = \frac{x-b}{z} \text{ therefore } \quad z = \frac{bf}{x_L - x_R}, \tag{5.7}$$

---

[6]For parallel cameras.

[7]Using the (x, y, z) coordinates originally defined for figure 5–1 (page 124), choosing the focal point of the rectified left camera as the origin. The horizontal and vertical coordinates of the rectified left image are x_L, y_L respectively. The optical axis of the left camera passes through their origin. The horizontal and vertical coordinates of the rectified right image are x_R, y_R respectively. The optical axis of the right camera passes through their origin. Note: below subscripts R and L are used for rectified images only.

where b is the base-line separating the focal points of the two cameras.

As figure 4–20 (page 119) demonstrates, specularities produce prominent edge features but they should not be used to estimate depth because:

1. They are virtual images reflected by the surface. Thus they produce horizontal disparities corresponding to depths lying above or below the actual surface.

2. They can create *vertical disparities*, where $y_L - y_R \neq 0$. An example is shown in figure 7–23 (page 199). So the epipolar constraint does not hold for specularities.

Consequences of these facts are:

- A separate matching process is necessary for specular features. The shape inference scheme of chapter 6 employs a simple process that matches specularities using the blob-like descriptions introduced in section 4.6.3.

- Features that violate the epipolar constraint are very likely to be specular. This fact remains untapped by the current specular detection scheme (chapter 4)—i.e. the scheme is merely tolerant of vertical disparity, but in principle vertical disparity could be used as an *indicator* for specularity.

5.3 Obtaining surface orientation

5.3.1 Introduction

A variety of techniques exist for estimating local surface orientation. If a dense depth map $z(x, y)$ is available then estimates can be obtained simply by numerically differentiating z with respect to x and y—see equation 5.2. In many circumstances only sparse depth maps or monocular imagery are available and other techniques are necessary. These include shape from contour [3,21,69],

shape from texture [2,65,70,71,109], shape from shading and photometric stereo. The latter two are considered in more detail below, as specularities play an important role in them both. The surface orientation at a specularity is more easily obtained when both lighting and viewing geometries are known—it is simply the local mirror-like orientation—see later. Buchanan [23] proposés a method to determine surface orientation when the light source geometry is unknown. It requires constraints both on the ratio of the viewing to light source distance and on the surface roughness. The same criticism that applied to the work of Healey, reviewed later, applies here, as it is assumed both that the "specular" component of the light is isolated and that the degree of roughness of each surface is known.

5.3.2 Shape from shading

Horn [57,67,60,59] and others [101,82] have developed methods that exploit the shading variations over smooth surface patches to recover surface orientation. The methods are based on the image irradiance equation (equation (2.5), page 17):

$$E(x,y) = R(p,q), \tag{5.8}$$

rewritten[8] here using the (p,q) notation to emphasise that two unknowns are involved. This equation provides only a single constraint—another is required to fully determine surface orientation. Several different constraints have been proposed embodying the fact that neighbouring points on a smooth surface cannot assume arbitrary orientations—somehow the surface must fit smoothly together. The constraints usually result in a large set of 1st order partial differential equations which when solved yield values of (p,q) across the surface. Parallel algorithms have been developed to provide efficient solutions to these equations.

---

[8]Strictly, this formulation assumes orthographic projection: i.e. a distant viewer.

The major drawback of shape from shading is that it can only be applied when the reflectance map R is known. Chapter 2 shows that this is not practical in all but very controlled circumstances. Most results have been obtained by assuming Lambertian reflectance maps. Babu et al' [6] determine surface orientation for some specular surfaces assuming particular models of R. However the models contain many unknowns and their method works only for planar surfaces.

5.3.3 Photometric stereo

If the same view is imaged under two different lighting conditions then two image irradiance equations are obtained for each image point (x, y):

$$E_A(x, y) = R_A(p, q) \text{ and } E_B(x, y) = R_B(p, q), \tag{5.9}$$

where E_A and E_B are the irradiances in the two images and R_A and R_B are the reflectance maps corresponding to the two lighting conditions. In this case the two equations provide sufficient constraints to recover both p and q at each point (x, y). This method of obtaining (p, q) is called photometric stereo. Woodham [111] demonstrates that the method can be made to work in real-time using controlled lighting conditions. In order to simplify the computation three, rather than two, reflectance maps are used. Each reflectance map corresponds to a Lambertian surface illuminated by a collimated source in a particular direction. Ikeuchi [66] demonstrates that the method works with specular surfaces if distributed light sources are used. Coleman and Jain [30] show that when four reflectance maps are employed specularities can be identified by the photometric system while simultaneously estimating (p, q). More recently Sanderson et al' [96] show that solder inspection tasks, involving the estimation of that local surface height and orientation from specularities, can be achieved by more elaborate control of the lighting. Photometric stereo requires careful regulation of the illumination and knowledge of the surface properties, thus it is only suitable for precisely controlled environments.

5.4 Integrating shading with stereo

Feature-based binocular stereo cannot estimate the depth across an untextured surface patch. Shape from shading can only do so if the reflectance map R is known. Various approaches have been made to overcome these two problems by integrating the shading and stereo information. The obvious approach is to apply binocular stereo to match irradiance within the right and left images of the patch. This is ineffective for Lambertian surfaces and—as Grimson [48] shows—it is unstable when there are specularities. Another approach: area-correlation [7], first creates a skeleton of edge features enclosing patches—using traditional stereo methods—before matching irradiance gradients within each patch. Blake [16] demonstrates that this approach is only good for short range viewing. Unacceptable errors accumulate at large viewing distances.

Until fairly recently monocular shape from shading schemes could only be applied to patches enclosed by extremal contours[9]. Ikeuchi [67] modifies the scheme so that any 3-D contour (obtained from binocular stereo) can be used. Nevertheless, knowledge of R is still required.

Grimson [47] and Terzopoulos [102] developed a scheme to reconstruct depth over smooth surface patches. The scheme involves fitting a thin plate of minimum energy to the depth—estimated at features by binocular stereo. Shading information is used only to ensure that the surface is smooth where it appeared to be so. No explicit knowledge of R is needed. Blake and Zisserman [19] show that Grimson's scheme provides solutions even where the combined stereo and shading information is ambiguous. Clearly it is dangerous to use such solutions. Blake and Zisserman argue that explicit depth reconstruction is often inappropriate and suggest [16] a qualitative approach—as in figure 5–3, page 128—when information is ambiguous.

---

[9]Edges along which the surface slant is 90 degrees

5.5 Obtaining curvature

5.5.1 Introduction

Woodham [112] used the Hessian to describe the curvature of Lambertian generalised cylinders. Work on the direct inference of surface curvature is scarce. That which has been done exploits specularities. Thrift and Lee [103] establish curvatures of some simple surface shapes such as spheres, cylinders and generalised cylinders, using specularities. However their method requires the class of shape to be known beforehand and it does not extend to determine curvatures of arbitrary shaped surfaces. For instance locally elliptical regions are not dealt with. Neither are stereoscopic cues considered. The work reviewed in the rest of this chapter attempts to solve more general problems.

5.5.2 The work of Healey and Binford

Healey and Binford [53] estimate surface curvature at specularities using monocular shading analysis. They employ an image irradiance equation (for the specular component only):

$$E_s(x,y) = E_o \exp[-\alpha^2(x,y)/m^2]. \tag{5.10}$$

Here, image irradiance E_s is defined with respect to the nearest specularity. E_o is the irradiance at the peak of the specularity, m specifies surface roughness and the half-angle α parameterises local surface orientation at an arbitrary image point (x,y). The expression is based on the micro-facet model of a rough surface—equation (2.18), page 34. The formulation requires seven main assumptions:

1. The specular component of image irradiance E_s has been extracted at all points in the image, beforehand. The chromatic specularity detection methods (see section 3.4) aim to achieve this.

2. The illumination is collimated along a single known direction—like that of a distant point source.

3. The surface is rough enough to employ the micro-facet model.

4. An ad-hoc micro-facet model $D(\alpha) = \exp(-\alpha^2/m^2)$ adequately represents the surfaces.

5. The roughness parameter m is known for each surface material in view.

6. Off-specular glints can be identified and avoided. This permits the factors G and F in (2.18) to be set to unity.

7. The surface depth at the specularity is known. The task of estimating this depth is not broached. The same information is needed by the scheme described in chapter 6—where a solution is proposed.

When all the above assumptions are valid—e.g. for a roughened metal under special lighting conditions—then equation (5.10) can be inverted to estimate $|\alpha|$ in the vicinity of each specularity:

$$|\alpha| = m\sqrt{\ln E_s(x,y)/E_o}. \tag{5.11}$$

This value is computed along radial lines passing through each specular peak. The magnitude of curvature is estimated along each radial line using:

$$|k(\theta)| = \partial|\alpha|/\partial s, \tag{5.12}$$

where s is the arc length along the radial line and θ parameterises its orientation as in figure 5-2 (page 126). The greatest and least values of $|k(\theta)|$ provide estimates of the magnitude of the principal curvatures at each specular peak. This method has been shown to work for metal spheres and cylinders of various radii.

Healey's work provides a different approach to that proposed in this thesis. His inference of surface curvature is based on precise measurements of image

irradiance at each specularity rather their positions and approximate shape. In principle, the two approaches are complementary and one of them might be used to confirm the other's results. However, in practice, the approach described in this thesis can be applied under much more general conditions. This is because:

- It is *not* restricted to collimated illumination.
- It is *not* restricted to surface materials of a known roughness parameter.
- It does *not* require the specular component of image irradiance to be measured. Specularities need only be marked.

5.6 The work of Blake

5.6.1 Introduction

Blake [13,15] introduces a theory for inferring constraints on surface curvature using specular imagery. The scheme described in chapter 6 builds on his original theory. His linear formulation aids an investigation of some important problems: e.g. focusing effects. However, it also introduces an unnecessary assumption: that of a short stereo base-line, which is discarded in the simpler formulation in chapter 6. Below the basic theory is reviewed. Details of the linear formulation are restricted to major results. The basic theory investigates a number of different questions:

- How do the stereo disparities of a specularity constrain the curvature of a surface?
- Under what conditions do the focusing effects disrupt the constraints?
- How are constraints obtained from a stereo-pair, assuming a distant point source?

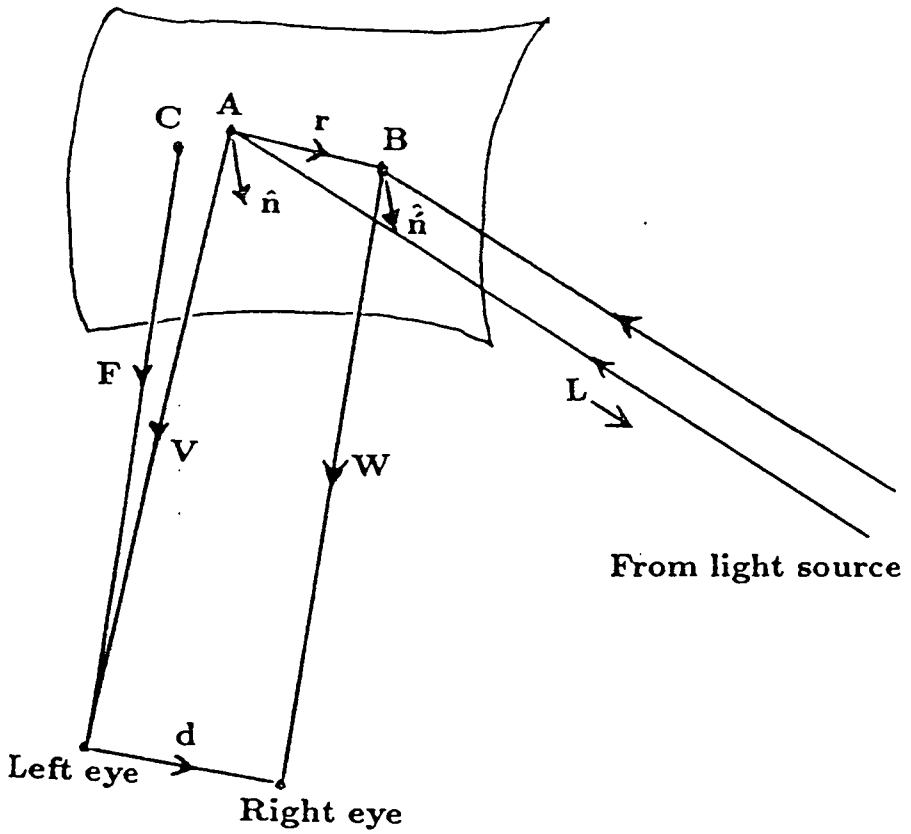


Figure 5–5: *A smooth patch of surface is illuminated by a distant point source along vector* **L**. *Light striking point A is specularly reflected along vector* **V** *into the left eye. Similarly, light incident at B traverses* **W** *into the right eye. Surface normals at A and B are* $\hat{\mathbf{n}}$ *and* $\hat{\mathbf{n}}$ *respectively. Vector* **r** *separates A and B. The stereo base-line lies along vector* **d**. *A surface marking lies nearby A at C.*

- How are constraints obtained from a stereo-pair, assuming a point source at a known finite distance?

- How are constraints obtained from a single image, assuming a distributed light source of known silhouette-shape?

- When can local curvature be fully inferred?

5.6.2 The viewing geometry

Figure 5–5 shows the stereo viewing geometry used by Blake—it is essentially the same as the one used in chapter 6. A smooth surface patch is illuminated along a single direction specified by the light source vector **L**. At point A the surface specularly reflects incident light along vector **V** and into the left eye. At B the reflection is along **W** and into the right eye. Surface normals at A and B are $\hat{\mathbf{n}}$ and $\hat{\mathbf{n}}$ respectively. Vector **r** describes the displacement from A to B. The stereo base-line lies along vector **d**.

Before curvature can be constrained the surface orientation at A must be estimated. This is done using the equation of mirror alignment:

$$\hat{\mathbf{n}} = \hat{\mathbf{h}} = \frac{\hat{\mathbf{V}} + \hat{\mathbf{L}}}{|\hat{\mathbf{V}} + \hat{\mathbf{L}}|}. \tag{5.13}$$

In general this calculation needs A's position to be estimated. The estimate is provided by binocular stereo. The estimate must be obtained from a nearby marking. In figure 5–5 C represents such a point. As long as C lies approximately in the tangent plane of A the estimate is good. An iterative technique for refining the estimate is described in chapter 6.

The relative orientation of the left and right eyes remain unspecified in figure 5–5. "Left eye" and "right eye" labels refer to positions of respective optical centres. For instance they need not be configured with parallel optical axes—as they are in figure 5–4, page 131. Care must of course be taken to express the depth estimates in the original (un*rectified*) left eye coordinate frame when estimating A's position.

5.6.3 Equation for a smooth surface

The surface passing through point A is locally described by a Taylor series written in terms of the 2-D vector in the tangent plane at A, **x**:

$$\mathbf{f}(\mathbf{x}) = (1/2)\mathbf{x} \cdot (\mathbf{H}\mathbf{x}) + O(|\mathbf{x}|^3), \tag{5.14}$$

where H is again the hessian matrix and $\mathbf{x} = (R_1, R_2)^T$ where[10] $\mathbf{r} = (R_1, R_2, R_3)^T$. The smoothness of the surface is assumed to be such that the higher order terms $O(|\mathbf{x}|^3)$ can be ignored. In this case:

$$(\delta n_1, \delta n_2)^T \approx -\mathsf{H}\mathbf{x} \text{ where } \boldsymbol{\delta n} = (\delta n_1, \delta n_2, \delta n_3)^T = \hat{\mathbf{n}} - \mathbf{n}. \tag{5.15}$$

5.6.4 Linear equation

A linear equation is derived that relates curvature (H) to stereo disparities (via $\mathbf{x}$). The derivation [13] combines equation (5.15) with the law of mirror-like reflection (5.13), expressed at both A and B and the vector cycle $\mathbf{V}+\mathbf{d}-\mathbf{W}-\mathbf{r}=0$. The latter is apparent from figure 5–5. The stereo base-line is assumed to be short, ($|\mathbf{d}| \ll |\mathbf{V}| \cos\sigma$) and the surface assumed not to focus light to a point or line near the viewer. The result for a point source at a finite distance is:

$$2V(\mathsf{MH} - \kappa_{VL}\mathsf{I})\mathbf{x} = \mathbf{w}, \tag{5.16}$$

where

$$\mathsf{M} = \begin{pmatrix} \sec\sigma & 0 \\ 0 & \cos\sigma \end{pmatrix}, \tag{5.17}$$

$$\kappa_{VL} = (1/2)(1/V + 1/L), \tag{5.18}$$

$$\mathbf{w} = (-d_1 + d_3 \tan\sigma, -d_2)^T, \tag{5.19}$$

$V = |\mathbf{V}|$, $L = |\mathbf{L}|$ and σ is surface slant at A—defined by equation (5.1), page 125.

Equation (5.16) can be inverted under circumstances described below. The inverted equation:

$$\mathsf{H}\mathbf{x} = \mathsf{M}^{-1}((1/2V)\mathbf{w} + \kappa_{VL}\mathbf{x}) \approx -(\delta n_1, \delta n_2)^T \tag{5.20}$$

---

[10]Here vectors are expressed in the local coordinate frame using three orthogonal components subscripted by 1, 2 and 3. The 3rd component lies along the surface normal, while the 1st lies along the projection of view vector V onto the tangent plane.

provides two constraints on curvature. Curvature is specified by H which has three unknowns—see equation (5.3). Thus another constraint is required to infer curvature in full. Blake's suggestions for augmenting equation (5.20) to obtain more information are discussed later. Equation (5.16) inverts only if $\det(\mathsf{M}) \neq 0$. This is guaranteed by a short stereo base-line, where $|\mathbf{d}| \leq V/2$.

When focusing effects occur no useful curvature information is provided by equation (5.20). Blake shows that these effects occur rarely: only on non-convex surfaces where $\det(\mathsf{MH}-\kappa_{VL}\mathsf{I}) = 0$. In these cases a specularity is either focussed outside of the bounds of the image or at more than one point. In the latter case, a line specularity occurs when the rank of $(\mathsf{MH}-\kappa_{VL}\mathsf{I})$ is 1 and an extended blob when the rank is 0. This analysis applies only for point light sources. Blobs and lines are more usually consequences of distributed light sources.

In practical circumstances when the illumination is provided by a distributed source the same equation can be usefully applied at the central peak of the specularity

5.6.5 Measuring x

To obtain constraints on H using (5.20) an estimate of x is required. Blake indicates that it can be estimated from measurements of the "angular position" of the specularity and a nearby reference point in the left and right images: $\mathbf{x} = V\mathsf{P}\boldsymbol{\delta}$ where

$$\mathsf{P} = \begin{pmatrix} \sec\sigma & 0 \\ 0 & 1 \end{pmatrix} \tag{5.21}$$

and the vector $\boldsymbol{\delta}$ compares the "angular positions". To be precise:

$$\boldsymbol{\delta} = (\widehat{\mathbf{A}_\mathbf{l}} - \widehat{\mathbf{C}_\mathbf{l}}) - (\widehat{\mathbf{A}_\mathbf{r}} - \widehat{\mathbf{C}_\mathbf{r}}) \tag{5.22}$$

where $\widehat{\mathbf{A}_\mathbf{l}}$, $\widehat{\mathbf{A}_\mathbf{r}}$, $\widehat{\mathbf{C}_\mathbf{l}}$ and $\widehat{\mathbf{C}_\mathbf{r}}$ are the "angular positions", respectively, of the specularity in the left image, the specularity in the right image, a nearby reference point in the left image and the reference point in the right image. Each of these "angular positions" can be measured using the same method—that which

is described for the quantity $\hat{\mathbf{V}}$ in chapter 6. There, the simpler formulation of the stereo analysis demonstrates this way of measuring $\mathbf{x}$ to be unnecessarily complicated[11]. However an analogue of vector $\boldsymbol{\delta}$ still proves useful in the monocular analysis.

5.6.6 Monocular analysis

Another approach is adopted for distributed sources that exploits the shape of the specularities. The analysis is again based on the constraint equation (5.15), $\mathbb{H}\mathbf{x} \approx -(\delta n_1, \delta n_2)^T$. The only difference between this monocular approach and the stereoscopic approach is the way in which the measurements of $\mathbf{x}$ and $(\delta n_1, \delta n_2)^T$ are obtained. In the stereoscopic case two views of the point on a light source are required while in the monocular case a single view of two points on a distributed light source (figure 5–6) is used. Two different analyses are possible:

- An analysis based on two points on the source separated by a known baseline (similar to the stereoscopic analysis).
- An analysis based on the distortion of the shape contour of the light source in the reflected specularity.

Both analyses use the same monocular linear constraint equation:

$$2L(\mathbb{M}\mathbb{H} - \kappa_{VL}\mathbb{I})\mathbf{x} = \mathbf{w_m}. \tag{5.23}$$

This is identical to the stereo linear constraint equation (5.16), except that the viewer and light positions are interchanged. The switch from a baseline between two eyes to one on the source alters the definition of the $\mathbf{w}$ term to become $\mathbf{w_m}$:

$$\mathbf{w_m} = L\mathbb{P}\boldsymbol{\alpha}, \tag{5.24}$$

---

[11]The reference point is really only needed for estimating depth.

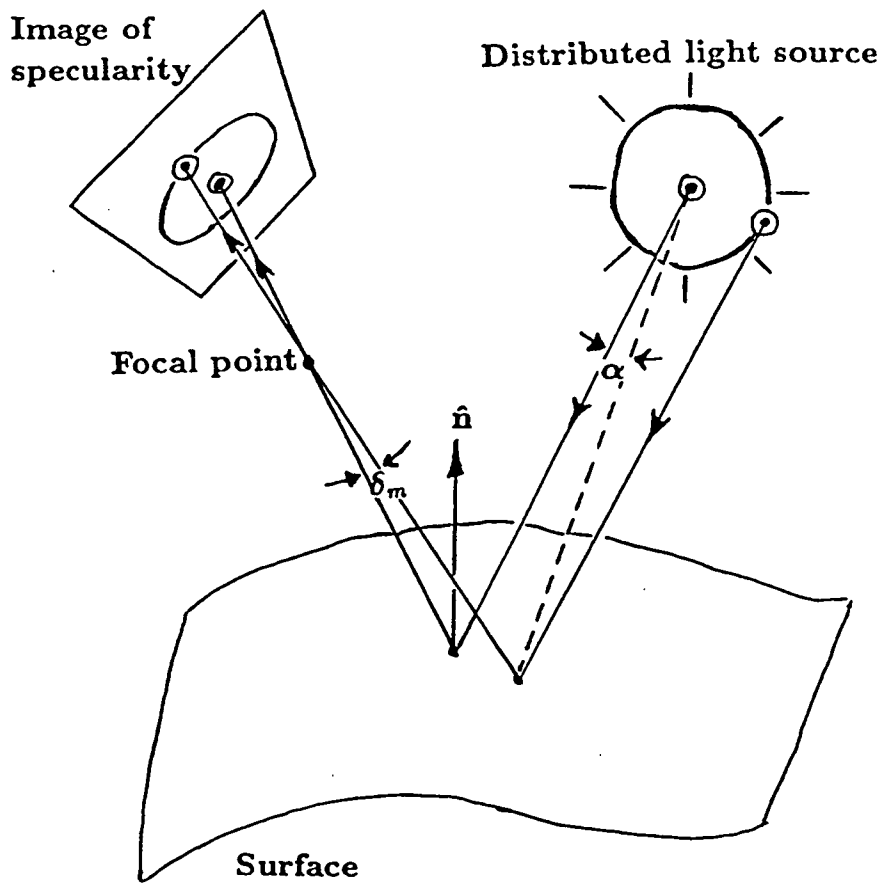


Figure 5–6: *Monocular analysis: A distributed light source of known shape is reflected by a curved surface as a specular image region. Surface curvature information may be inferred by measuring the shape in the image. A point on the the source, transforms to a point on of the specularity. The angular separation of two such points is specified using α and δ_m of the source and specularity respectively.*

where α is the angular separation of one point on the source from another. The expression for x also changes to:

$$\mathbf{x} = V\mathsf{P}\delta_m, \tag{5.25}$$

where δ_m is the angular separation of the two reflected points on the specularity. P is the projection matrix (5.21). Before going into more details the assumptions underlie this monocular approach will be clarified:

1. The formulation of equation (5.23) requires the light source's dimension s to be small compared to its distance from the surface L, more precisely: $L\cos\sigma \gg s$.

2. Curvature is assumed to be constant over the region of surface covered by the specularity. This might not hold for a widely distributed source or a broken surface patch.

3. The contour of the specularity is assumed to be the mirror-reflection of the contour of the light source. Two processes can undermine this assumption:

   - Focusing effects - that cause blob-like specularities where a point is otherwise expected.
   - Rough surfaces—with wide spread functions that extend the contour beyond mirror-like orientations. The effect of a surface with a wide spread function reflecting a point source is equivalent to that of a smooth surface reflecting a distributed source. However the monocular analysis is unable to disambiguate these two effects.

4. The shape of the light source, as seen from the surface is assumed to be known. If it is to be approximated by the shape as seen from the viewer then the light source must be relatively distant—as required in 1 above.

5. The monocular analysis assumes that the viewing distance V is known. However, if the analysis is to be truly monocular then binocular stereo

will not be available to estimate this depth. But as monocular analysis is chiefly intended to supplement the stereoscopic an estimate is usually available.

Two point monocular analysis

Blake omits a discussion of the two point monocular analysis mentioned above. It is in fact much less applicable than the contour shape based analysis because of its requirement for detailed knowledge of the light source:

- Two points must be unambiguously identifiable on the source. This is only feasible for sources of a certain shape—e.g the two vertices on the crescent of the moon might be used. Additionally the reflections of each point must also be unambiguously identifiable on the specularity.

- The baseline vector **s** between the two points on the source must be known. For the stereo analysis it was reasonable to assume that the interocular baseline was known—it is after all an integral and often fixed part of the viewing apparatus. However it is unreasonable to expect knowledge of **s** for any particular source involved, except under carefully controlled lighting conditions.

When detailed knowledge of the source is available equation (5.23) can be applied to provide two constraints on curvature—just as in the stereoscopic case. To avoid the short baseline assumption required by the linear formulation a new formulation, given in appendix H, can be used.

Source contour shape transformation

Blake does discuss the contour shape based analysis mentioned above. He describes the linear transformation $\mathbb{T}^{-1}$ of the shape of the light source into the shape of the specularity by:

$$\mathbb{T}\delta_m = \alpha, \tag{5.26}$$

where an arbitrary point on the contour of the light source is specified by α—its angular position with respect to a reference point on the source (say the

centroid). The point specified by $\boldsymbol{\alpha}$ is reflected at a point on the specularity's contour, specified by $\boldsymbol{\delta}_m$—its angular position with respect to the image of the reference point. Figure 5–6 illustrates the case for a circular contour. From equations (5.23), (5.24) and (5.25) the expression for T is obtained:

$$\mathsf{T} = 2V\mathsf{P}^{-1}\mathsf{M}\mathsf{H}\mathsf{P} - 2V\kappa_{VL}\mathsf{I}. \tag{5.27}$$

The shape of the specularity is simply the transformed shape of the the light source. The particular transformation depends on both the viewing and lighting geometry (via V, P, M and κ_{VL}) and the surface curvature (via H). In general the transformation may be a combination of a scaling and a rotation.

Here are some special cases:

- When the surface is **planar**, $\mathsf{H}=0$ so that $\mathsf{T}=-2V\kappa_{VL}\mathsf{I}$. The transformation is simply a shape preserving, isotropic scaling. The scaling factor is $-(2V\kappa_{VL})^{-1}$. For a very distant source for which $V \ll L$, the scaling factor is simply -1, (a mirror inversion).

- When the surface is **curved** and it is orientated at zero slant ($\sigma=0$) then $\mathsf{T}=2V(\mathsf{H}-\kappa_{VL}\mathsf{I})$. Which is again a simple scaling transformation. The scaling factors aligned along the surface's principal curvature directions are $-(2V(k_1-\kappa_{VL}))^{-1}$ and $-(2V(k_2-\kappa_{VL}))^{-1}$. For example, a cylindrical surface for which $(k_1,k_2)=(k_c,0)$ and $k_c \gg \kappa_{VL}$ squashes the shape by a large amount in the direction perpendicular to the cylinder's axis—corresponding to a small scaling factor $-(2Vk_c)^{-1}$. In the axial direction the scaling factor is as for a plane $-(2V\kappa_{VL})^{-1}$. A characteristically elongated specularity is guaranteed, as long as the light source is not elongated itself. In this case, an elongated specularity is *monocular evidence* that the surface has one large and one small principal curvature.

- When the surface is **curved** but $\sigma \neq 0$ then the transformation generally involves a rotation as well as a scaling. The nature of such a transformation is further investigated in chapter 6, where the circumstances are

discussed in which the axis of elongation of a specularity can be interpreted as the projection of one of the principal curvature directions. This direction might then be used in combination with stereoscopic constraints to infer full curvature, as described in the next section.

- When the shape of the **light source** is **circular**—as is often the case, e.g. the sun or a spot lamp—then the shape of the specularity is an ellipse. By measuring the length and direction of the ellipse's axes, $\mathbb{T}$ and hence $\mathbb{H}$—full curvature—can be found. Exactly how this can be achieved is described in chapter 6. Because of the symmetry of the circular source it transpires that there is a fourfold ambiguity of the inferred curvatures.

5.6.7 Inferring curvature

Equation (5.20) provides two constraints on the three unknowns of local surface curvature. In the absence of other information these constraints ensure that the principal curvatures (k_1, k_2) lie in a restricted space—which can be represented graphically, see chapter 6. Often however, additional information is available and curvature can be obtained in full. There are four situations:

1. **A locally spherical surface:** There is only one unknown $k_s = k_1 = k_2$. From (5.3), $\mathbb{H} = k_s\mathbb{I}$ so that equation (5.20) supplies two estimates of $k_s = -\delta n_1/R_1$ and $k_s = -\delta n_2/R_2$. (As before $\boldsymbol{\delta n} = (\delta n_1, \delta n_2, \delta n_3)^T$ and $\mathbf{x} = (R_1, R_2)^T$.) A least-squares combination of these provides a more robust measure:

$$k_s = -(R_1\delta n_1 + R_2\delta n_2)/(R_1^2 + R_2^2). \tag{5.28}$$

2. **A locally cylindrical surface:** There are two unknowns—one principal curvature is known to be zero, the other has an unknown value k_c. The remaining unknown is the axial direction of the cylinder θ_c. In this case equation (5.3) simplifies so that equation (5.20) gives

$$\tan\theta_c = \delta n_2/\delta n_1, \quad k_c = \delta n_1/(R_1\sin^2\theta_c + R_2\sin\theta_c\cos\theta_c). \tag{5.29}$$

3. **Known orientation of principal axes:** This case corresponds to knowing θ_0 in equation (5.3). Chapter 6 discusses how θ_0 can be estimated from the shape of some specularities. The remaining unknowns k_1 and k_2 are obtained by rotation of the local coordinate frame, i.e.

$$\acute{\mathsf{H}}\acute{\mathbf{x}} = -(\acute{\delta}n_1, \acute{\delta}n_2)^T. \tag{5.30}$$

where

$$\acute{\mathsf{H}} = \mathsf{R}(-\theta_0)\mathsf{H}\mathsf{R}(\theta_0) = \begin{pmatrix} k_1 & . \\ . & k_2 \end{pmatrix}, \tag{5.31}$$

$$\acute{\mathbf{x}} = (\acute{R}_1, \acute{R}_2)^T, = \mathsf{R}(-\theta_0)\mathbf{x}, \tag{5.32}$$

$$(\acute{\delta}n_1, \acute{\delta}n_2)^T = \mathsf{R}(-\theta_0)[(\delta n_1, \delta n_2)^T]. \tag{5.33}$$

Now

$$(k_1, k_2) = -(\acute{\delta}n_1/\acute{R}_1, \acute{\delta}n_2/\acute{R}_2). \tag{5.34}$$

4. **Three cameras:** Two stereo pairs obtained from a set of three cameras provide four constraints on local curvature. So it can be recovered in full. The process works best for roughly orthogonal stereo base-lines—see [13] for the mathematical details. This approach is pursued no further in this thesis.

Chapter 6

Inferring surface shape: new work

6.1 Introduction

The new work presented below refines and extends the work of Blake that was reviewed above. It exploits specularities to infer local curvature information on smooth surfaces. To do so, it requires descriptions of specularities within images. The specularity detector presented in chapter 4 is able to provide these descriptions. This chapter provides the theory—which when implemented as a computational scheme (in chapter 7) produces useful curvature information. The theory divides into four sections:

1. **Stereo analysis:** a simple formulation of the constraints obtainable from a stereo-pair, a graphical method of representing the constraints and two new ways to interpret them.

2. **Stereo measurement:** how to obtain the positional and angular measurements required as input to the stereo analysis. This section includes a treatment of the errors and ends by summarising the various inferences to be drawn from the stereo analysis.

3. **Monocular analysis:** Analysis for a distributed light source with a circular contour and a discussion of the conditions necessary for inferring a principal curvature direction from a specularity's axis of elongation.

4. **Combined inference:** how the stereo and monocular analyses can combine to infer fuller curvature information.

6.2 Stereo analysis

6.2.1 Simple formulation

Here the Taylor series expansion, in the form of equation (5.15), is again used to represent a smooth surface. In place of Blake's linear formulation a simple formulation is presented using vector algebra. As a consequence, no assumption of a short stereo base-line is required. Both formulations assume the same information to be initially available, i.e:

- $\mathbf{V}$—the view vector, with respect to the left eye. The vectors $\mathbf{V}$ as well as $\mathbf{W}$, $\mathbf{d}$ and $\mathbf{S}$, defined below are all shown in figure 6–1. This is estimated from the position of a marking near to the specularity. Section 6.3.3 describes how the estimate is obtained and refined.

- $\hat{\mathbf{W}}$—the direction (angular position) of the the view vector, with respect to the right eye. Section 6.3.2 describes how specularities are identified, matched and their angular positions estimated.

- $\mathbf{d}$—the stereo base-line vector. This describes the displacement between the two eyes. It is reasonable to assume that $\mathbf{d}$ is known.

- $\mathbf{S}$—the position of the particular light source that creates the specularity. It is less reasonable to assume that $\mathbf{S}$ is known. There may be may different light sources to choose from. A high level process that infers and reasons about the positions of the possible light sources might be used to supply S. Scope exists for further work in this area.

Again a specular ray is incident at a point A when viewed from the left eye and at point B when viewed from the right.

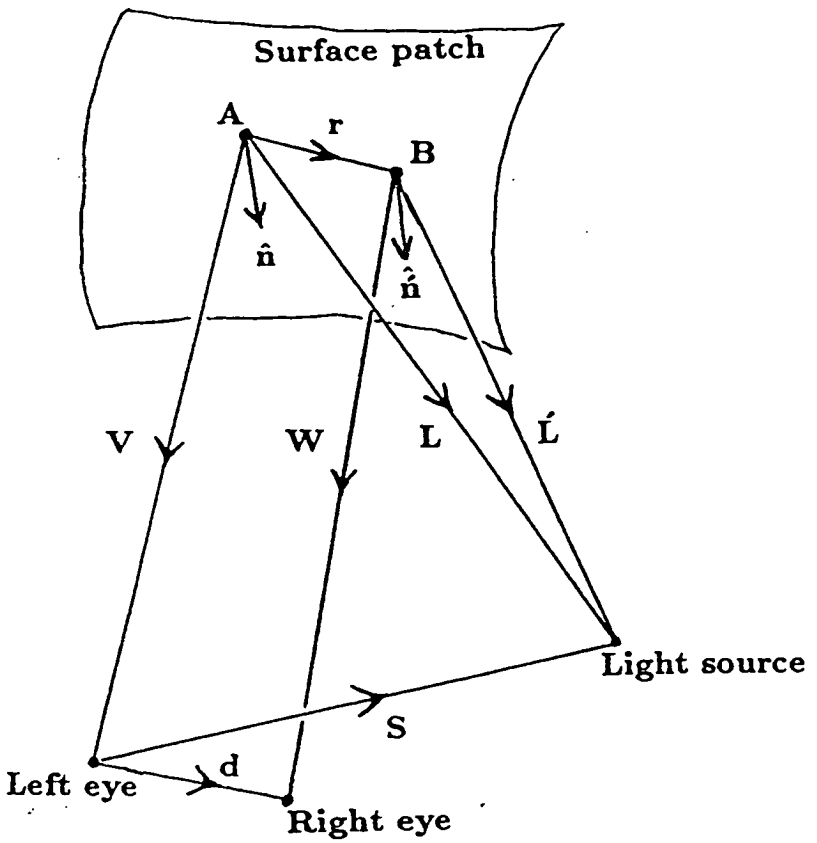


Figure 6–1: *Geometry for the simple formulation of the stereo analysis (c.f. figure 5–5, page 140). A specularity appears to lie at point A when viewed from the left eye and at point B when viewed from the right. The vectors marked are used in the accompanying analysis.*

The formulation is as follows. The surface normal at A is given by the law of mirror-like reflection:

$$\hat{\mathbf{n}} = (\hat{\mathbf{V}} + \hat{\mathbf{L}})/|\hat{\mathbf{V}} + \hat{\mathbf{L}}|. \tag{6.1}$$

where $\mathbf{L} = \mathbf{V} + \mathbf{S}$. The angular position of the specularity in the right image is transformed into the left-eye coordinate frame by a simple transformation matrix $\mathsf{T}_{r\to l}$ which depends only on the known camera geometry. The particular form of the $\mathsf{T}_{r\to l}$ appropriate for the camera geometry used in this thesis is described in appendix I.1. The magnitude W of $\mathbf{W}$ is obtained from the following vector equality:

$$\begin{aligned} \mathbf{W} &= \mathbf{V} + \mathbf{d} - \mathbf{r}, \\ \text{i.e. } W\hat{\mathbf{W}} &= \mathbf{V} + \mathbf{d} - \mathbf{r} \end{aligned} \tag{6.2}$$

—apparent from figure 6–1 . Now taking the scalar product of both sides with $\hat{\mathbf{n}}$, observing that $\mathbf{r} \cdot \hat{\mathbf{n}} \approx 0$ and rearranging, the magnitude is obtained:

$$W \approx \frac{(\mathbf{V} + \mathbf{d}) \cdot \hat{\mathbf{n}}}{\hat{\mathbf{W}} \cdot \hat{\mathbf{n}}}. \tag{6.3}$$

Now vector $\mathbf{W} = W\hat{\mathbf{W}}$ is used to derive $\mathbf{r}$,

$$\mathbf{r} \approx \mathbf{V} + \mathbf{d} - \mathbf{W} \tag{6.4}$$

and $\hat{\hat{\mathbf{n}}}$—the surface normal at B (again using the law of mirror-like reflection):

$$\hat{\hat{\mathbf{n}}} = (\hat{\mathbf{W}} + \hat{\acute{\mathbf{L}}})/|\hat{\mathbf{W}} + \hat{\acute{\mathbf{L}}}|, \tag{6.5}$$

where $\acute{\mathbf{L}} \approx \mathbf{W} - \mathbf{d} + \mathbf{S}$. This is apparent from figure 6–1. Finally, the formulation of equation (5.15), page 142 is:

$$\begin{aligned} \mathsf{H}\mathbf{x} = - \quad & (\delta n_1, \delta n_2)^T, \\ \text{where} \quad (\delta n_1, \delta n_2)^T &= \left((\hat{\hat{\mathbf{n}}} - \hat{\mathbf{n}}) \cdot \hat{\mathbf{i}},\ (\hat{\hat{\mathbf{n}}} - \hat{\mathbf{n}}) \cdot \hat{\mathbf{j}}\right)^T \\ \text{and} \quad \mathbf{x} &= \left(\mathbf{r} \cdot \hat{\mathbf{i}},\ \mathbf{r} \cdot \hat{\mathbf{j}}\right)^T, \end{aligned} \tag{6.6}$$

where $\left(\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}\right)$ are the component unit vectors along the axes of the local coordinate frame at A. They are defined in terms of vectors readily available in the left-eye frame:

$$\begin{aligned} \hat{\mathbf{k}} &= \hat{\mathbf{n}}, \\ \hat{\mathbf{i}} &= (\hat{\mathbf{V}} - \hat{\mathbf{L}})/|\hat{\mathbf{V}} - \hat{\mathbf{L}}|, \\ \hat{\mathbf{j}} &= \hat{\mathbf{k}} \times \hat{\mathbf{i}}. \end{aligned} \tag{6.7}$$

The methods of inferring curvature listed in section 5.6.7 also apply for this new formulation. Below, in sections 6.2.3 and 6.2.4 two new methods are added to the list.

Note however, whenever a component of $\mathbf{x}$ is zero (i.e. $R_1 = 0$ or $R_2 = 0$) then one of the two curvature constraints embodied by (6.6) is lost. Information is not available along the direction of the missing component. For example: when there is no vertical disparity then curvature is often undetermined in the vertical direction. The sphere shown in figure 7–27 (page 212) shows such a case.

6.2.2 Constraint graphs

The constraints (6.6) can be represented graphically—below it is shown how. This representation is used in chapter 7 in the absence of additional constraints.

Two of the three elements of the hessian matrix can be obtained by rotating[1] the local coordinate frame (in 2-D) through an angle α such that:

$$\mathsf{R}(\alpha)\mathbf{x} = (|\mathbf{x}|, 0)^T. \tag{6.8}$$

In this new frame the hessian becomes:

$$\ddot{\mathsf{H}} = \mathsf{R}(\alpha)\mathsf{H}\mathsf{R}(-\alpha) \tag{6.9}$$

or expanding in terms of (k_1, k_2, θ_0) via equation (5.3), (page 127):

$$\ddot{\mathsf{H}} = \mathsf{R}(\varphi) \begin{pmatrix} k_1 & 0 \\ 0 & k_2 \end{pmatrix} \mathsf{R}(-\varphi), \tag{6.10}$$

where $\varphi = \theta_0 + \alpha$. In the new frame the constraint equation (6.6) becomes:

$$\begin{aligned} \ddot{\mathsf{H}}(|\mathbf{x}|, 0)^T &\approx - \ (\delta\tilde{n}_1, \delta\tilde{n}_2)^T, \\ \text{where} \quad (\delta\tilde{n}_1, \delta\tilde{n}_2)^T &= \mathsf{R}(\alpha)(\delta n_1, \delta n_2)^T. \end{aligned} \tag{6.11}$$

---

[1] R is the rotation matrix—see equation (5.4) on page 127.

Denoting the hessian as follows:

$$\ddot{\mathsf{H}} = \begin{pmatrix} H_{xx} & H_{xy} \\ H_{yx} & H_{yy} \end{pmatrix}, \tag{6.12}$$

equation (6.11) supplies the two elements:

$$\begin{aligned} H_{xx} &= \delta\tilde{n}_1/|\mathrm{x}|, \\ H_{xy} &= \delta\tilde{n}_2/|\mathrm{x}|, \end{aligned} \tag{6.13}$$

which are related to (k_1, k_2, θ_0) by equation (6.10), so that:

$$\begin{aligned} H_{xx} &= k_1 \cos^2\varphi + k_2 \sin^2\varphi, \\ H_{xy} &= (k_2 - k_1)\cos\varphi \sin\varphi. \end{aligned} \tag{6.14}$$

Now the values of H_{xx} and H_{xy}, obtained from equation (6.13), restrict the permissible values of the principal curvatures k_1, k_2 to lie upon a right-rectangular hyperbola in (k_1, k_2) space. This follows from the observation that:

$$\begin{aligned} -{H_{xy}}^2 &= -\ (k_2 - k_1)^2 \cos^2\varphi \sin^2\varphi, \\ &= (k_1 \sin^2\varphi - k_2 \sin^2\varphi)(k_2 \cos^2\varphi - k_1 \cos^2\varphi), \\ &= (k_1 - k_1 \cos^2\varphi - k_2 \sin^2\varphi)(k_2 - k_1 \cos^2\varphi - k_2 \sin^2\varphi), \\ &= (k_1 - H_{xx})(k_2 - H_{xx}) \end{aligned} \tag{6.15}$$

is the equation of a hyperbola. As figure 6–2 (a) shows the hyperbola has its origin at (H_{xx}, H_{xx}). Note: by definition $k_1 \geq k_2$ so the permissible values of k_1 and k_2 are restricted to one curve of the hyperbola. In chapter 7 the constraints are represented using the directly measurable quantities $r_1 = 1/k_1$ and $r_2 = 1/k_2$ rather than k_1 and k_2. It is easily shown (see appendix G) that r_1 and r_2 are similarly restricted to a right-rectangular hyperbola in (r_1, r_2) space:

$$-B^2 = (r_1 - A)(r_2 - A), \tag{6.16}$$

where $A = H_{xx}/({H_{xx}}^2 + {H_{xy}}^2)$ and $B = H_{xy}/({H_{xx}}^2 + {H_{xy}}^2)$. As figure 6–2 (b) shows this hyperbola has its origin at (A, A) and as $r_1 \leq r_2$ the permissible values of r_1 and r_2 are restricted to its upper curve.

When uncertainties exist in the values obtained for H_{xx} and H_{xy} then r_1 and r_2 are restricted to a set of hyperbolæ. This is true of the graphs presented in chapter 7 so they show zones rather than lines of permissible values.

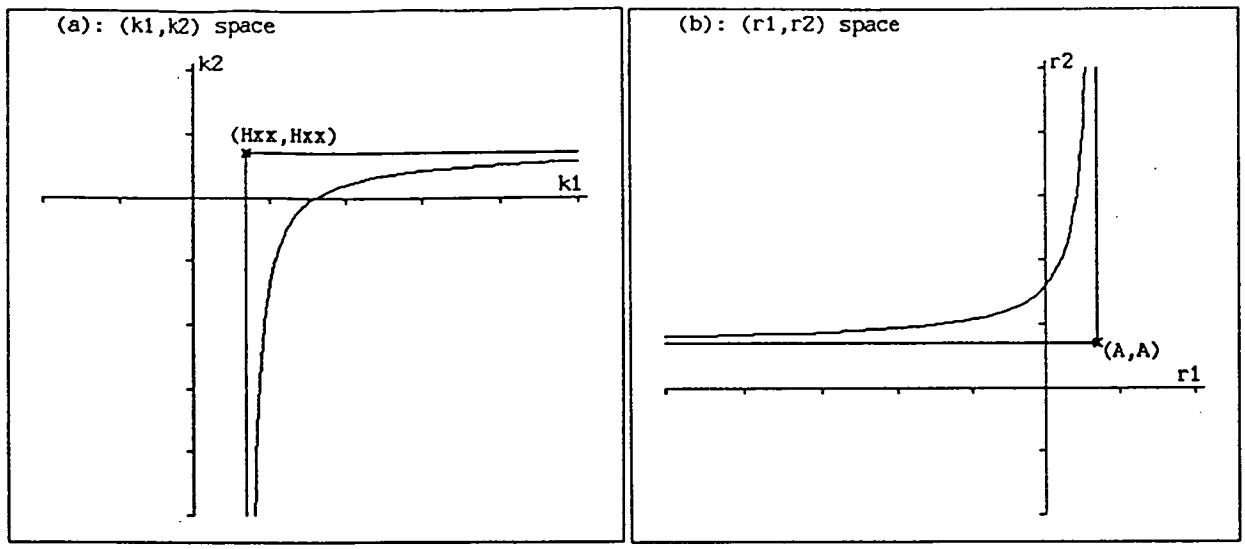


Figure 6–2: *(a) Stereo analysis constrains the principal curvatures k_1 and k_2 to one curve of a hyperbola in (k_1, k_2) space. Its origin is at (H_{xx}, H_{xx}). (b) Similarly the radii of curvature $r_1 = 1/k_1$, $r_2 = 1/k_2$ lie on the upper curve of a hyperbola in (r_1, r_2) space. Its origin is at (A, A) where $A = H_{xx}/(H_{xx}{}^2 + H_{xy}{}^2)$.*

6.2.3 Locally hyperbolic surface ($k_1 = -k_2$)

Two new ways of interpreting the curvature constraints (6.6) are considered in the following sections. Curvature can be obtained in full for surfaces known to be locally spherical or cylindrical—see section 5.6.7. This is also true for a locally hyperbolic surface with principal curvatures k_1 and k_2 such that $k_1 = k_h$, $k_2 = -k_h$. When monocular analysis suggests that $|k_1| = |k_2|$ both this hyperbolic case and the locally spherical case should be considered as possible. In this hyperbolic case the hessian matrix H, expressed using equation (5.3), simplifies to

$$\mathsf{H} = \mathsf{R}(\theta_0) k_h \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \mathsf{R}(-\theta_0) \tag{6.17}$$

which expands (see appendix G.2) as:

$$\mathsf{H} = k_h \begin{pmatrix} \cos 2\theta_0 & -\sin 2\theta_0 \\ -\sin 2\theta_0 & -\cos 2\theta_0 \end{pmatrix} \tag{6.18}$$

and this expression in conjunction with the constraint equation (6.6) provides an estimate of the orientation $\theta_h = \theta_0$:

$$\theta_h = \tan^{-1} \frac{1}{2}\left(\frac{\delta n_2 R_1 + \delta n_1 R_2}{\delta n_2 R_2 - \delta n_1 R_1}\right) \tag{6.19}$$

and two estimates of the magnitude k_h:

$$k_h = \delta n_1/(R_2 \sin 2\theta_h - R_1 \cos 2\theta_h), \tag{6.20}$$

$$k_h = \delta n_2/(R_1 \sin 2\theta_h + R_2 \cos 2\theta_h). \tag{6.21}$$

A least squares combination of the latter expressions produce a more robust estimate:

$$k_h = \frac{\delta n_1(R_2 \sin 2\theta_h - R_1 \cos 2\theta_h) + \delta n_2(R_1 \sin 2\theta_h + R_2 \cos 2\theta_h)}{R_1^2 + R_2^2}. \tag{6.22}$$

6.2.4 Non-spherical test

The non-spherical test is again based on equation (6.6)—it determines when a surface is not locally spherical. Whenever the test indicates a surface is not locally spherical the interpretation as a sphere (via equation (5.28), page 149) need not be considered.

If a surface is locally spherical then the hessian matrix is $\mathsf{H} = k_s \mathbb{I}$, where k_s is the magnitude of curvature. In this case $-k_s \mathbf{x} = (\delta n_1, \delta n_2)^T$, i.e. $\mathbf{x}$ and $(\delta n_1, \delta n_2)^T$ are parallel. So when measured in any coordinate frame: $-k_s = \delta n_1/R_1 = \delta n_2/R_2$. If this equality holds (to within calculated error bounds) then the surface might be locally spherical. If it does not then the surface is **not** locally spherical. It is simplest to check if the equality is broken in the frame in which $\mathbf{x}$ and $(\delta n_1, \delta n_2)^T$ are readily available—i.e. the local coordinate frame.

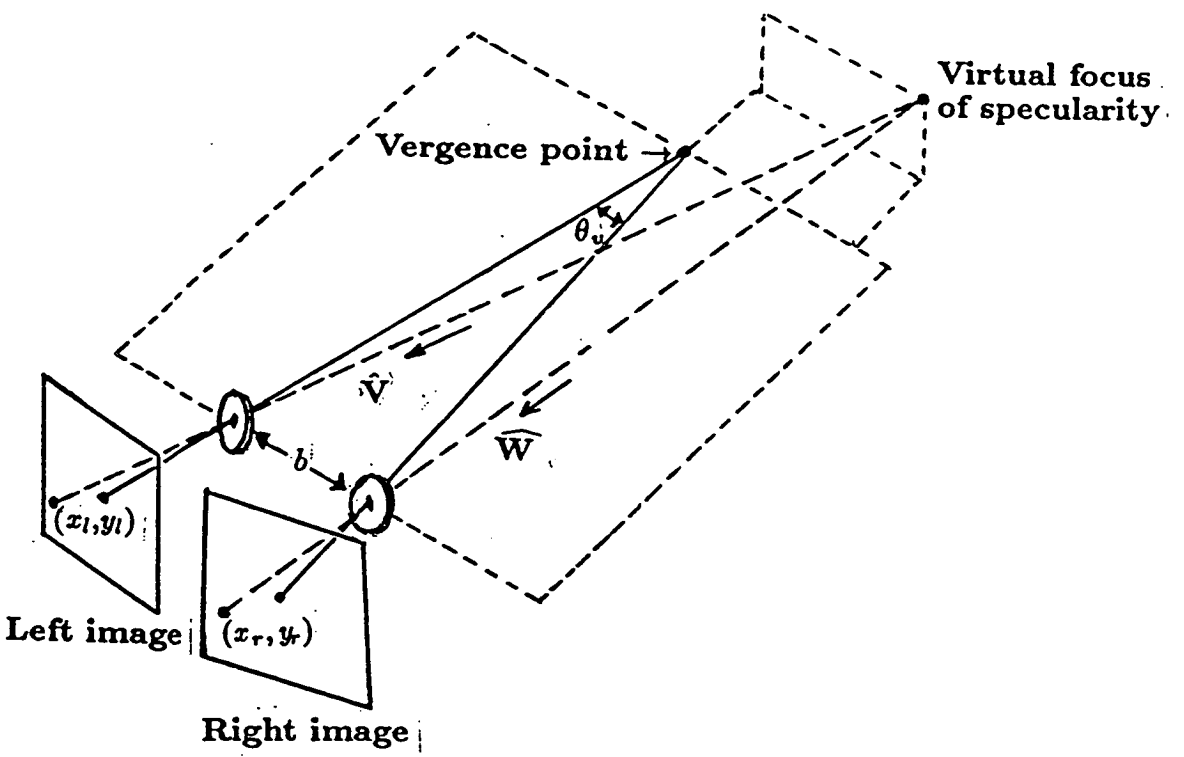


Figure 6–3: *Stereo viewing geometry: In the left and right images the specularity appears at points* (x_l, y_l) *and* (x_r, y_r) *respectively. In general the optical axes of the camera need not verge at a point equidistant from their focal point—even though they do for the geometry of the implemented system.*

6.3 Stereo measurement

6.3.1 Stereo viewing geometry

Before the constraint equation (6.6) can be used to infer curvature, the four quantities listed in section 6.2.1: $\mathbf{V}$, $\hat{\mathbf{W}}$, $\mathbf{d}$ and $\mathbf{S}$ must be supplied. Methods to determine these quantities depend on the particular stereo viewing geometry. Although the stereo analysis, described above, does not preclude other geometries, that used in this thesis was chosen as it is both convenient and fairly general. It is as shown in figure 6–3: the cameras are oriented so that their optical axes verge at a point equidistant from their focal points while their vertical axes are perpendicular to the plane containing the optical axes[2]. The position of the light source $\mathbf{S}$ is likely to be known with a degree of uncertainty. How uncertainties in the quantities $\mathbf{V}$, $\hat{\mathbf{W}}$, $\mathbf{d}$ and $\mathbf{S}$ results in uncertainty in the constraint equation is discussed in section 6.3.4. In the next two sections, the methods used to obtain $\hat{\mathbf{W}}$ and $\mathbf{V}$ are described.

6.3.2 Obtaining $\hat{\mathbf{W}}$ and $\hat{\mathbf{V}}$

To obtain $\hat{\mathbf{W}}$ and $\hat{\mathbf{V}}$ involves four basic stages:

1. Detecting specularities in both left and right images: the scheme described in chapter 4 is used for this purpose.

2. Matching specularities detected in the left image to those in the right. A method to do this is described below.

---

[2]In this case the appropriate baseline vector $\mathbf{d}$ and transformation matrix $\mathsf{T}_{r\to l}$, that takes a vector in the right camera coordinate frame into the left, are given in appendix I.1.

3. Using the location of the specularities in the left and right image: (x_l, y_l) and (x_r, y_r) respectively, to provide the two directions :

$$\hat{\mathbf{V}} = \frac{1}{\sqrt{x_l^2+y_l^2+f_l^2}}(x_l,\ y_l,\ f_l)^T,\quad \hat{\mathbf{W}} = \frac{1}{\sqrt{x_r^2+y_r^2+f_r^2}}(x_r,\ y_r,\ f_r)^T \quad (6.23)$$

where f_l and f_r are the known focal lengths of the left and right caméras respectively.

4. Transforming $\hat{\mathbf{W}}$ into the left camera coordinate system: $\hat{\mathbf{W}} = \mathsf{T}_{r\to l}\hat{\mathbf{W}}$.

In theory $\hat{\mathbf{V}}$ and $\hat{\mathbf{W}}$ can be evaluated and the constraint equation applied for every appropriate *specular pair* of image points, (x_l, y_l) and (x_r, y_r). However, rather than attempting to identify every available specular pair it is more useful to use a few readily identifiable pairs—those with constituent points lying at the central peaks of specular blobs. Descriptions of specular blobs are delivered by the specularity detector. The central peaks are particularly suitable because:

- They are the most noise-resistant specular features and thus the most likely to be detected.

- Their image locations are identifiable and matchable. Points at the periphery of specular blobs formed by distributed light sources have ambiguous matches—in general there is no way to distinguish which source point matches to which specularity point.

- They are the most likely points to give specular reflection at a mirror-like orientation as required by stereo analysis. Points at the periphery of specular blobs are often part of an extended spread function—see section 2.4.4—where reflection is not mirror-like. In the rare circumstance when the central peak is an off-specular glint it should not be used.

When a specular blob is distorted by the focusing effects discussed in section 5.6.4, it should not be used to infer curvature. However, there is no way to tell whether focusing has occurred. The scheme in chapter 7 infers curvature for all available blobs and accepts an occasional false inference due to focusing.

The uncertainty in locating (x_l, y_l) and (x_r, y_r) depends upon the size and shape of the matched pair of specular blobs:

- The specularity matcher—described below—matches only blobs of approximately the same size, shape and orientation. This simplifies the task of locating (x_l, y_l) and (x_r, y_r). Gross errors only occur if bad matches are made. No such gross errors occur in the results in chapter 7.

- When a blob is small its centroid and peak tend to coincide and (x_l, y_l) and (x_r, y_r) can be located to within half a pixel [75].

- When a small blob contains rival peaks the scheme in chapter 7 chooses those closest to the centroid. So again (x_l, y_l) and (x_r, y_r) can be located to within half a pixel.

- Elongated blobs present problems. For example the elongated specularity in figure 7–31 page 216 has several peaks along its spine. Ideally each peak along the spine should be matched in two images to produce a set of (x_l, y_l), (x_r, y_r) pairs. Then local stereo analysis could be applied to each. Scope exists to extend the scheme in chapter 7 to do this. The current implementation treats all blobs alike: applying the stereo analysis only at the peak closest to the centroid. However the ends of an elongated specularity are difficult to locate—as figure 7–31 demonstrates: it is hard to pin-point exactly where the specularity stops and conventional shading starts when irradiance falls off gradually. So the centroid—somewhere mid way between the two ends—has an added uncertainity in its location along the axis of elongation. The uncertainty tends to increase with blob length h. This is incorporated into the scheme by assigning an uncertainty to the location of the centroid, ϵ. Accurately estimating ϵ is impractical—both the surface curvature and light source characteristics are influencing factors. A very rough estimate is selected in chapter 7 where $\epsilon = \max(1/2, 5\% \text{ of } h)$. This uncertainty is then propagated through the scheme to determine the uncertainty in the results—as described later.

Matching specularities

The blob-like descriptions provided by the detector (section 4.6.3) are used to match specularities into pairs suitable for stereo analysis. Each pair consists of one blob from the right image and one from the left. The following matching criteria are employed:

- **Size and shape:** The images of the same light source reflected by two adjacent points on a smooth surface tend to be similar in size and shape. So a pair of similar specularities are likely to be a good match. The height h, width w and orientation θ of blobs in the left and right images are used as a measure of their likeness. In particular the absolute difference in their heights[3] $D_{l,r} = |h_l - h_r|$, aspect ratios $A_{l,r} = |h_l/w_l - h_r/w_r|$ and orientations $O_{l,r} = |\theta_l - \theta_r|$ are used.

- **Relative position:** A gross disparity occurs between the image locations of a surface obtained for the left and right view points. Surfaces at different depths result in different disparities. Relative to the surface it lies on, a specularity has an additional disparity determined by the local curvature. Convex and concave regions result in relative disparities in opposite directions. In theory, the disparity of each pair of specularities could be entirely independent. In practice, it is profitable to match sets of pairs that have similar disparities because:

  - Specularities tend to cling to small regions where the local curvature is high. Thus their relative disparities are usually small compared to the gross disparities of the surface regions.

  - The surfaces of interest usually lie at similar depths and so tend to have similar gross disparities.

---

[3]Again subscripts l and r are used to denote quantities in the left and right images respectively

The difference in centroid locations: $P_{l,r} = \sqrt{(x_{c,l}-x_{c,r})^2 + (y_{c,l}-y_{c,r})^2}$ provides the measure of the disparity employed below.

- **Unpaired specularities are rare:** A specularity in the right image need not correspond to one in the left (and vice versa). There are various reasons for this:

  - Local surface orientation may be at a mirror-like alignment for only one view point[4]. The specularity on the handle of the mug in figure 7–31, page 216 provides a typical example.
  - Large disparities in image location may displace a specularity out of one of the images. For cylinders specularities often traverse very large distances in the direction of zero curvature.
  - Failure of the specularity detector can result in individual unpaired specularities.

  When there are only a few unpaired specularities, it is beneficial to favour sets of matched pairs over unmatched sets. The matching scheme below allows this possibility through its penalty term U.

Global solutions to the matching problem encompassing all three of the above criteria can be obtained by minimising the following quantity (over all possible matches):

$$\omega_U U + \sum_{r\in R}\sum_{l\in L} M_{l,r}$$

where $M_{l,r} = \omega_D D_{l,r} + \omega_A A_{l,r} + \omega_P P_{l,r} + \omega_O O_{l,r}$,

L is the set of specularities in left image,

R is the set of specularities in right image and,

U is the no. of unmatched specularities in both images.

(6.24)

[4]E.g: two entirely different sets of specularities can be seen by closing each eye in turn when viewing a sunlit choppy sea.

The ω's are weighting constants[5] and D, A, P and O are as defined above. A uniform set of constants were applied to match each stereo-pair shown in chapter 7: $\omega_D = 100/\textit{image-size}$, $\omega_P = 100/\textit{image-size}$, $\omega_O = 50/\pi$, $\omega_U = 10,000/\textit{number-of-blobs}$. However, these images represent fairly simple matching problems—in more difficult circumstances different weightings may be appropriate. The minimisation of equation (6.24) is implemented as a pruned[6], depth-first search.

6.3.3 Obtaining V: depth estimation

The direction of the view vector, $\hat{\mathbf{V}}$ is obtained from equation (6.23). Its magnitude (the depth) remains to be determined. Binocular stereo can be applied to estimate surface depth but not at specularities (section 5.2.2). Instead the depth of a surface marking near the specularity is determined. The vector of the marking, F, thus obtained, is used as a first approximation to V. This approximation is then refined using an iterative process. In all, estimating V consists of eight stages:

1. **Edge detection:** Edges of both surface markings and specularities are extracted from the left and right images. The same process is used by the specularity detector—a Canny operator followed by hysteresis thresholding.

2. **Pruning:** Any edges corresponding to specularities are removed. The masking technique described in section 4.7.7 achieves this.

---

[5]In fact ω_O is not strictly constant: orientation is made to carry less weight when it is known to be uncertain—i.e. for specularities with small aspect ratios.

[6]By maintaining a minimum score so far, any search paths exceeding the current minimum are pruned.

3. **Rectification:** The remaining edge segments are rectified to obtain the standard parallel geometry—shown in figure 5–4, page 131. An epipolar based matcher can now by used.

4. **Matching:** The PMF matcher [91,92] is used to match corresponding edge segments in the left and right images. The horizontal disparity of each match is provided.

5. **Triangulation:** The disparity is converted into an estimate of depth—using equation (5.7), page 132.

6. **De-rectification:** The rectification process is reversed so that each depth estimate is associated with its location in the original left image. The estimate is then adjusted so that it applies to the unrectified left eye.

7. **Extraction:** The edge segments of the left image are searched. Those closest to each specularity are extracted. Their depth and image location are used to estimate the view vector $\mathbf{F}$. Horizontal edge segments provide unreliable depth estimates so they are not used.

8. **Iterative refinement:** An estimate of $\mathbf{V}$ is made from $\hat{\mathbf{V}}$, $\mathbf{F}$ and $\mathbf{S}$—the position of the light source.

Iterative refinement

The view vector, $\mathbf{F}$ to a near-by surface point C, is used to estimate the view vector, $\mathbf{V}$ at the specularity, A. Figure 6–4 shows the geometry involved. The problem is: to use the known quantities $\mathbf{F}$, $\hat{\mathbf{V}}$ and $\mathbf{S}$ to obtain reasonable and consistent estimates of depth, $V = |\mathbf{V}|$ and surface orientation, $\hat{\mathbf{n}}$. For any value of V the orientation must be consistent with mirror-like reflection. If C is close to A then it will almost lie on the tangent plane of the surface at A. Here, an estimate is obtained by assuming that C lies *in* the tangent plane. Under this assumption two equalities hold:

$$\mathbf{F}\cdot\hat{\mathbf{n}} = \mathbf{V}\cdot\hat{\mathbf{n}} \quad \text{and} \quad \hat{\mathbf{n}} = (\hat{\mathbf{V}} + \hat{\mathbf{L}})/|\hat{\mathbf{V}} + \hat{\mathbf{L}}|. \tag{6.25}$$

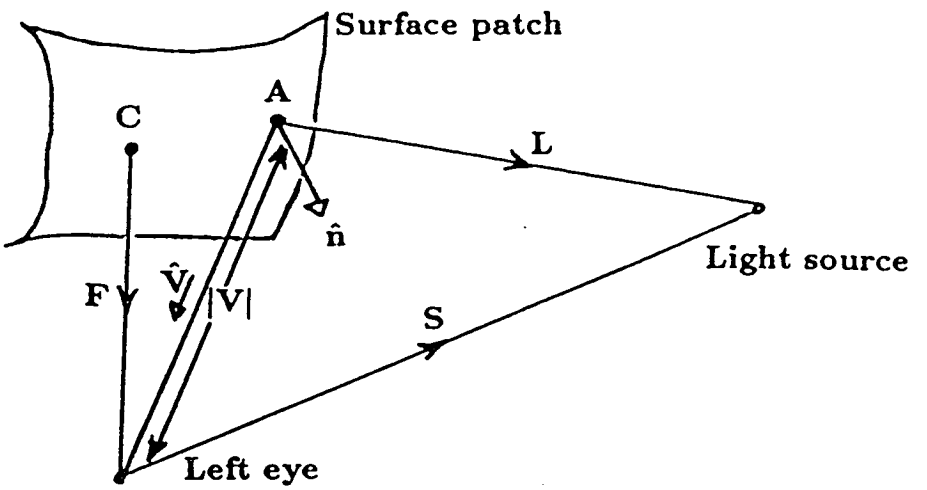


Figure 6–4: *Geometry for iterative refinement. The unknown quantities:* $V = |\mathbf{V}|$, $\hat{\mathbf{n}}$, $\mathbf{L}$ *are obtained from the knowns:* $\mathbf{F}$, $\hat{\mathbf{V}}$, $\mathbf{S}$. *The specularity lies at A and the surface marking at C.*

The first expresses the tangent-plane-assumption and the second the law of mirror-reflection. They are simultaneously satisfied by a unique pair of the quantities V and $\hat{\mathbf{n}}$. These quantities are estimated using the iterative[7] scheme shown in the box below.

$i=0$, $V^{(i)}=|\mathbf{F}|$.
repeat

$$
\begin{aligned}
i &= i+1, \\
\mathbf{V}^{(i)} &= V^{(i)}\hat{\mathbf{V}}, \\
\mathbf{L}^{(i)} &= \mathbf{V}^{(i)}+\mathbf{S}, \\
\hat{\mathbf{n}}^{(i)} &= \left(\hat{\mathbf{V}}+\hat{\mathbf{L}}^{(i)}\right)/|\hat{\mathbf{V}}+\hat{\mathbf{L}}^{(i)}|, \\
V^{(i)} &= \left(\hat{\mathbf{n}}^{(i)}.\mathbf{F}\right)/\left(\hat{\mathbf{n}}^{(i)}.\hat{\mathbf{V}}\right).
\end{aligned}
$$

while $|V^{(i)}-V^{(i-1)}| > 0.01|\mathbf{F}|$.
$\mathbf{V}=\mathbf{V}^{(i)}$, $\hat{\mathbf{n}}=\hat{\mathbf{n}}^{(i)}$, $\mathbf{L}=\mathbf{L}^{(i)}$.

[7]Superscript i denotes the step of iteration.

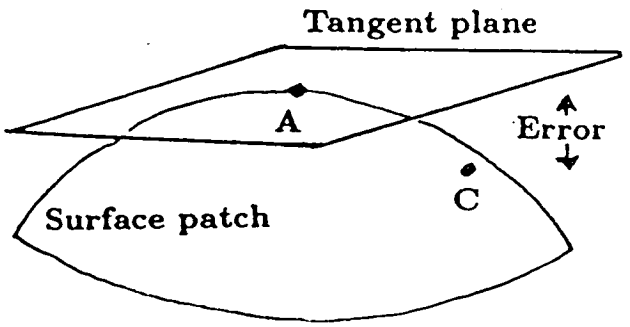


Figure 6–5: *Point C lies on the surface at a little distance from A. Along this distance the surface curves gently out of the tangent plane. So the "tangent plane assumption" is only an approximation.*

As the scheme[8] iterates V moves from C to A along a curve lying in the tangent plane at A. In the absence of other information—the assumption that the point C lies in the tangent plane is a useful first approximation. However, it is in fact inaccurate for any curving surface. Consider the example shown in figure 6–5: the point C lies on the surface at a little distance from A. Along this distance the surface curves gently out of the tangent plane. The departure of C from the tangent plane introduces an error into the curvature information inferred by stereo analysis. The magnitude of this error is proportional to the magnitude of curvature of the actual surface. This is of course unknown, before the analysis is applied. In consequence uncertainty exists in the inferred curvature information.

---

[8]Simply checking that the *relative error* $\epsilon_i = |V^{(i)} - V^{(i-1)}|$ has dropped to 1% of $|\mathbf{F}|$, as done above, does not necessarily guarantee 1% accuracy in the estimated value of V. A more rigorous condition for terminating the scheme is to check when an estimate of *absolute error* η_i drops below a specified value. Often the relative error decays such that $\epsilon_i = \epsilon_0 a^{-i}$ so that $\eta_i = \epsilon_i(1 - a)$ (see [99]). So if a is first estimated from the first few iterations an estimate of η_i is readily available for the rigorous test. Note however, the test on relative error proved sufficient to analyse the images in this thesis.

The scheme shown in chapter 7 does not cater for this uncertainty—it assumes the error is small. However, an intriguing possibility exists: the error might be estimated from the inferred curvature as follows. The patch of surface around A might be modelled using the inferred curvature. This model could then be used to estimate the actual position of C. Then the departure of C from the tangent plane could be estimated and corrected for before running the stereo analysis again.

6.3.4 Error treatment

Each measured quantity input to the stereo analysis has uncertainty associated with it. This uncertainty can be quantified using error bounds. By combining the errors it is possible to quantify the uncertainty of the quantities in the constraint equation (6.6), i.e. R_1,R_2,δn_1 and δn_2. This can be done by applying the appropriate error combination formula at each step in the analysis. Porrill [93] and Durrant-Whyte [33] provided some principles for doing this. The scheme in chapter 7 it is done using gaussian error combination [100]. This method strictly only applies when combining independent sources of uncertainty. As the uncertainties involved here are unlikely to be completely independent, room exists for improving the treatment of errors by the scheme. Table 6–1 summarises the sources of uncertainty and indicates how they are quantified by the scheme. Table 6–2 lists the six applications of the constraint equation, previously discussed in isolation. All are used by the scheme in chapter 7. The first three estimate the curvature which would be consistent if the surface were locally spherical, cylindrical or hyperbolic ($k_1 = -k_2$). The fourth decides whether the constraints are inconsistent with a spherical surface. The fifth estimates parameters used by the constraint graph. The last combines the constraint equation with the known direction of curvature in order to obtain both principal curvatures. The uncertainties in the constraint equation, i.e. in R_1,R_2,δn_1 and δn_2, create uncertainties in the outcomes of all six applications. These uncertainties

Source of uncertainty	Input	Status	Used by scheme
Image measurement $\hat{\mathbf{V}}$	(x_l, y_l)	Known	ϵ (page 162)
Image measurement $\hat{\mathbf{W}}$	(x_r, y_r)	Known	ϵ (page 162)
Camera base-line	$\mathbf{d}$	Known	0
Light source position	$\mathbf{S}$	Unknown*	0
Depth estimation (PMF)	V	Known	1% (PMF)
Tangent plane assumption	V	Unknown+	0
* Depends on other knowledge sources. + Depends on curvature.			

Table 6–1: Uncertainties *of the quantities input to the stereo analysis.*

Application	Outcome	Reference
If locally spherical	k_s	(5.28)
If locally cylindrical	(k_c, θ_c)	(5.29)
If locally hyp.($k_1 = -k_2$)	k_h, θ_h	(6.19) (6.22)
Non-spherical test	pass/fail	Sec 6.2.4
Constraint graph	H_{xx}, H_{xy}	(6.13)
Known direction, θ_0	k_1, k_2	(5.34)

Table 6–2: *The* six applications *of the constraint equation.*

are estimated by using the same step-by-step error treatment as is applied to the stereo analysis itself.

6.4 Monocular analysis

6.4.1 Introduction

In most circumstances specularities are formed by distributed light sources so that the monocular approach introduced earlier can be used to infer curvature. Although the monocular approach can, by itself, provide useful curvature information when applied to a single image, it is most profitable to use it in conjunction with the stereo approach, described above. Often the two approaches provide complementary information and full curvature information can be inferred by combining them. For example: the monocular approach can provide the principal curvature direction so that equation (5.34), page 150, can be used to estimate both principal curvatures from the stereo analysis. Below the monocular analysis, reviewed in section 5.6 is expanded. As indicated in that section, an analysis based on two known points on the source requires rather too detailed a knowledge of the lighting conditions to be usefully applicable in general circumstances. Nevertheless, a formulation of this analysis, which avoids the short baseline assumption, is given in appendix H. Two more useful questions of monocular analysis are tackled below:

1. How is curvature inferred, up to a fourfold ambiguity when the light source is circular?

2. When are conditions suitable for inferring a principal curvature direction from a specularity's axis of elongation?

In both cases the linear transformation, equation (5.23), on page 144, is used and the five assumptions listed in section 5.6 are required.

6.4.2 Monocular analysis for a circular source

Points on the contour of a circular source satisfy:

$$\alpha^T \alpha = \rho^2, \tag{6.26}$$

where ρ is the source radius and α specifies an arbitrary contour point—as defined on page 144. A specularity formed by this source has an elliptical contour. Figure 5–6, on page 145, illustrates this case. The position of an arbitrary point on this contour, specified by δ_m (also defined on page 144) is determined by $\mathsf{T}\delta_m = \alpha$ where T is the transformation defined (in equation (5.27)) in terms of the lighting and viewing geometry and the surface curvature. Equation (6.26) now becomes:

$$\delta_m^T \mathsf{T}^2 \delta_m = \rho^2, \tag{6.27}$$

where $\mathsf{T}^2 = \mathsf{T}^T\mathsf{T}$ as T is symmetric.

When the source radius ρ is known, T^2 can be fully determined by measuring the ellipse. (Its eigenvalues Λ_1 and Λ_2 are determined by ρ and the lengths of principal radii of the ellipse and its eigenvectors are aligned with the direction of ellipse's principal axes.) T^2 then determines T itself, up to a fourfold ambiguity. The eigenvectors of T and T^2 share the same directions. The fourfold ambiguity occurs because the signs of the two eigenvalues of T, λ_1 and λ_2 are unknown, i.e:

$$(\lambda_1, \lambda_2) = \left(\pm\sqrt{\Lambda_1}, \pm\sqrt{\Lambda_2}\right). \tag{6.28}$$

For each of the four possible transformations T the interpretation in terms of the surface curvature can be obtained by inverting equation (5.27), on page 148, to compute H

$$\mathsf{H} = \mathsf{M}^{-1}\mathsf{P}\left((1/2V)\mathsf{T} + \kappa_{VL}\mathsf{I}\right)\mathsf{P}^{-1} \tag{6.29}$$

In general, four different interpretations are possible, each may have principal curvatures of different magnitudes and in different directions. The four possible directions can collapse down to two. This occurs as slant, $\sigma \to 0$—in this case both M and P are I and (6.29) is simply $\mathsf{H} = ((1/2V)\mathsf{T} + \kappa_{VL}\mathsf{I})$. It also occurs, as

illustrated in Figure 6–6, when principal curvatures are not small and are not of approximately the same magnitude, i.e. $||k_1|-|k_2|| \gg 2\kappa_{VL}\sec\sigma$.

When the source radius is unknown T can still be determined but only up to an arbitrary multiplicative constant. In this case either the surface is locally approximately planar—i.e. both principal curvatures are of the order of κ_{VL}' or less—or else at least one principal curvature k_1 is large ($k_1 \gg \kappa_{VL}\sec\sigma$) and H can be computed up to a fourfold ambiguity and an arbitrary multiplicative constant.

Allowing for polar projection

Equation (5.26) describes the transformation of a circular contour into an elliptical specularity in polar projection. In polar projection the image is formed on a plane, shown in figure 6–7, perpendicular to $\hat{\mathbf{V}}$ (the direction of mirror reflection for the centre of the source), at a distance V from the surface. For conventional cameras the image plane is perpendicular to the direction of the optical axis $\hat{\mathbf{K}}$ not $\hat{\mathbf{V}}$. Nevertheless, the analysis can still be applied, as long as the points on the specularity are transformed into polar projection before applying (5.26). Transformation Q takes an arbitrary image point X into polar projection:

$$\delta_m = \mathsf{Q}\mathbf{X}. \tag{6.30}$$

Q accounts for both foreshortening (via F) and rotation (via R, see equation (5.4)):

$$\mathsf{Q} = \mathsf{R}(-\beta)[\mathsf{F}]. \tag{6.31}$$

Angle β and the foreshortening transformation F are defined in terms of two directions that lie in the image plane: $\hat{\mathbf{u}}$ and $\hat{\mathbf{v}}$. The former is the projection of $\hat{\mathbf{V}}$ onto the image plane and the latter is the orthogonal direction, i.e:

$$\begin{aligned} \hat{\mathbf{u}} &= \hat{\mathbf{K}}\times(\hat{\mathbf{K}}\times\hat{\mathbf{V}}), \\ \hat{\mathbf{v}} &= \hat{\mathbf{K}}\times\hat{\mathbf{u}}. \end{aligned} \tag{6.32}$$

Now $\beta = \cos^{-1}\left(\hat{\mathbf{v}}\cdot\hat{\mathbf{j}}\right)$ where $\hat{\mathbf{j}}$ is the direction defined in equation (6.7). Also

$$\mathsf{F} = \begin{pmatrix} u_x\cos\gamma & u_y\cos\gamma \\ v_x & v_y \end{pmatrix} \tag{6.33}$$

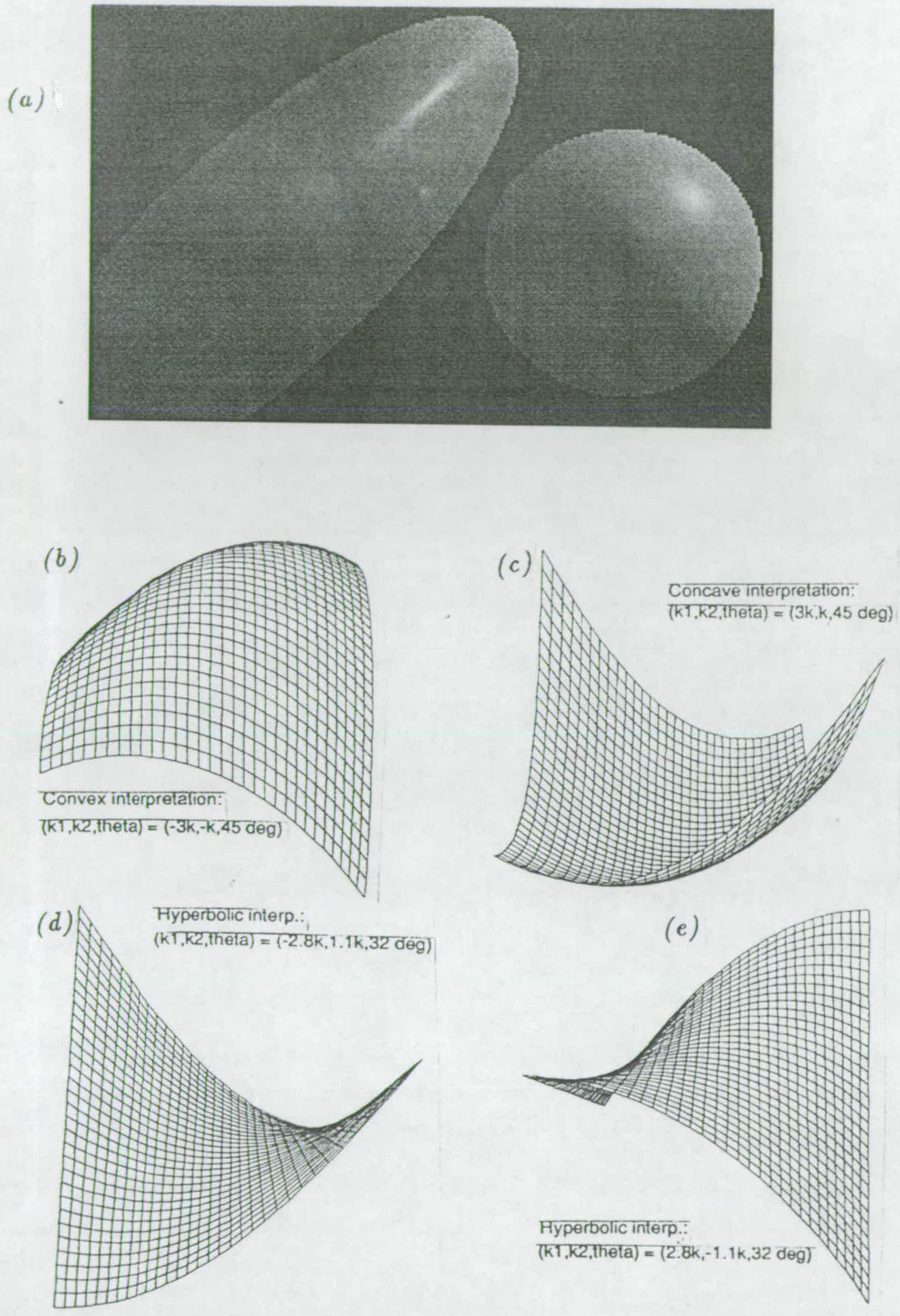


Figure 6–6: *Fourfold ambiguity: (a) A monocular image of a convex sphere and a convex ellipsoid. Observation of the specularity on the sphere determines the light source direction ($\sigma = 45^0$). Observation of the specularity on the ellipsoid fixes the curvature parameters (k_1, k_2, θ_0) up to a fourfold ambiguity and an arbitrary multiplicative constant k. Plots (b)(c)(d)(e) depict the four possible local surface interpretations.*

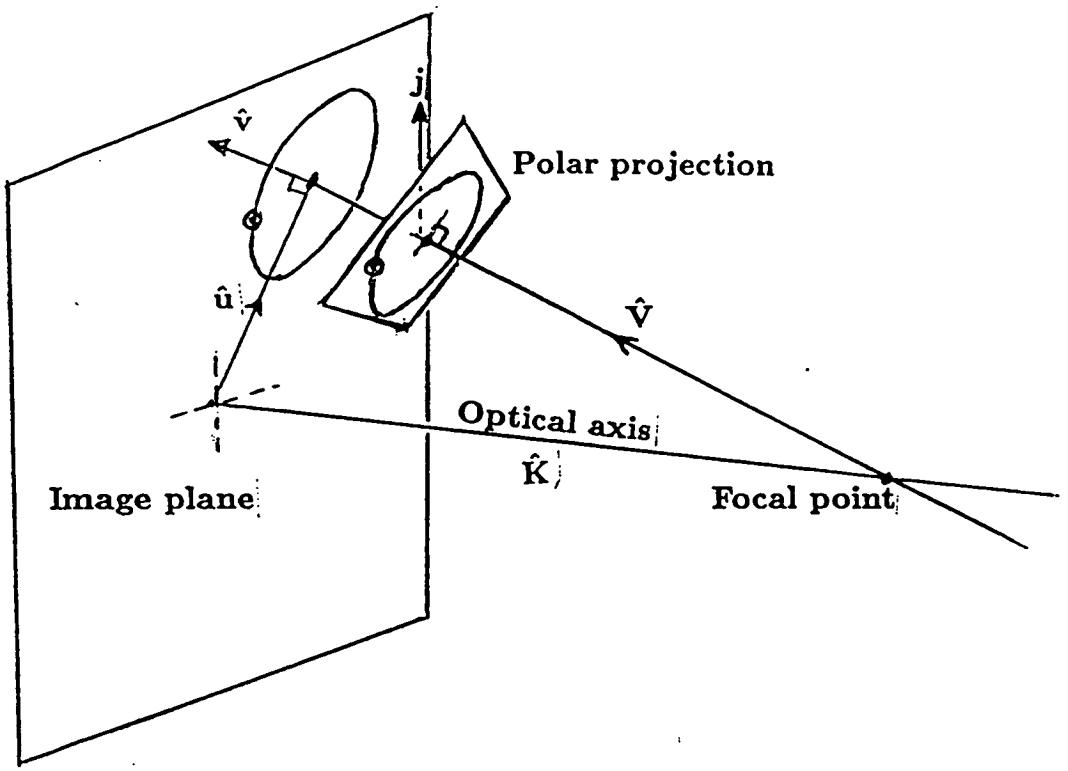


Figure 6–7: *By transforming each point on the contour of the specularity in the image into polar projection the correct shape of the ellipse can be measured. Transformation* **Q** *takes a point on the image plane into polar projection, for a particular specularity.*

where $\hat{\mathbf{u}} = (u_x, u_y)^T$, $\hat{\mathbf{v}} = (v_x, v_y)^T$ and angle $\gamma = \cos^{-1}(\hat{\mathbf{V}} \cdot \hat{\mathbf{K}})$ is the off-axis angle.

The shape on the image plane is obtained using the transformation (6.30): $\delta_m = \mathbf{Q}\mathbf{X}$

$$\mathbf{X}^T\mathbf{Q}^T\mathbf{T}^2\mathbf{Q}\mathbf{X} = \rho^2. \tag{6.34}$$

Measurements made within the image plane can provide $\mathbf{T}_Q = \left(\mathbf{Q}^T\mathbf{T}^2\mathbf{Q}\right)$, the shape transformation so that $\mathbf{T}^2$ can be computed:

$$\mathbf{T}^2 = \left(\mathbf{Q}^T\right)^{-1}\mathbf{T}_Q\mathbf{Q}^{-1}. \tag{6.35}$$

6.4.3 Monocular inference of principal curvature directions

When the source is known only to be compact—i.e. not elongated—then if it forms an elongated specularity the direction of least[9] curvature can be estimated in the following manner. Identify the specularity's axis of elongation and project it back onto the surface tangent plane. This estimate might then be used to complement stereoscopic constraints to estimate full curvature—as described in section 5.6.7. The estimate of least curvature direction obtained by this method is only an approximation. For locally almost flat surfaces[10] the approximation is unreliable. As the surface curvature increases the approximation improves. As it does also as slant $\sigma \rightarrow 0$, as the source becomes more compact and as the blob becomes less compact.

Three factors influence the quality of the approximation:

---

[9]The least absolute principal curvature.

[10]Stereoscopic measurements, when available can be used to distinguish such surfaces—as very small values of $|(\delta n_1, \delta n_2)^T|$ will be computed via equation (6.6).

1. The eigenvectors of $\mathbb{T}$ do not lie exactly in the same direction as the specularity's axis of elongation. However, it has been shown[11] [17] that the angular error in assuming the two directions are the same does not exceed:

$$\tan^{-1}\left\{K\left(\frac{K^2-1}{a^2-K^4}\right)^{1/2}\right\}, \tag{6.36}$$

where K quantifies the compactness of the source as the ratio of the maximum to minimum distance from the source's centre to its bounding contour. The aspect ratio of the specularity is a. Note: $a > K^2$.

2. Equation (5.27) on page 148 can be restated as:

$$\mathbb{T} = 2V\mathbb{P}^{-1}\mathbb{M}\mathbb{H}^*\mathbb{P} \quad ; \quad \mathbb{H}^* = \mathbb{H} - \kappa_{VL}\mathbb{M}^{-1}. \tag{6.37}$$

Now if $\mathbf{u}$ is an eigenvector of $\mathbb{T}$, then in the limit when its eigenvalue tends to zero, it can be seen that $\mathbb{P}\mathbf{u}$ is an eigenvector of $\mathbb{H}^*$. But, when its eigenvalue is non-zero $\mathbf{u}$ is only approximately an eigenvector of $\mathbb{H}^*$. However, it has been shown [17] that the angular error in assuming $\mathbb{P}\mathbf{u}$ to be in the direction of an eigenvector of $\mathbb{H}^*$ does not exceed:

$$\min\left[\sin^{-1}\left\{\sin\sigma(1-\sec^2\sigma\frac{K}{a})^{-1/2}\right\},\ \sin^{-1}\left\{\tan\sigma(\cos^2\sigma\frac{a}{K}-1)^{-1/2}\right\}\right]. \tag{6.38}$$

3. Even if the eigenvectors of $\mathbb{H}^*$ are correctly estimated, they are only approximately equal to the eigenvectors of $\mathbb{H}$. However, it has been shown [17] that the angular error in assuming that $\mathbb{H}^*$ is $\mathbb{H}$ does not exceed:

$$\sin^{-1}\left\{\sin\sigma\left(3\frac{|k_{min}|}{|k_{max}|}+8\frac{K}{a}\sec^2\sigma\right)^{1/2}\right\}. \tag{6.39}$$

---

[11] In the three expressions below the surface is assumed locally non-planar such that $|k_{max}| > 5\kappa_{VL}\sec\sigma$ where the principal curvatures are k_{max} and k_{min}, defined so that $|k_{max}| > |k_{min}|$.

6.5 Combined inference

More curvature information can be inferred by combining stereo and monocular measurements than can be inferred from them individually. Several methods to combine the measurements were given in sections 5.6.7 and 6.2. In this section these methods are brought together in a single inference scheme. It infers the fullest possible curvature information from the available stereo and monocular measurements, while ensuring an interpretation consistent with both—as shown in the table below.

COMBINED INFERENCE				
MONOCULAR			STEREO	
Source[12] shape	**Specularity shape[13]**		**Yes[14]**	**No**
Circle	Circle or ellipse (not elongated)	Source radius known	A	B
		Source radius unknown	C	D
	Other shape		E	None
Compact (including circle)	Elongated	Conditions reliable[15]	F (Full)	G
		Conditions unreliable	E	None
	Not elongated		E	None
Other shape			E	None

For each pair of monocular and stereoscopic conditions a key letter *A* to *G* provides a reference to the appropriate combined inference in the list below. For instance, if the light source is a circle of known radius and the specularity

---

[12]For all monocular analysis the source dimension is assumed to be small, i.e. $s \ll L\cos\sigma$.

[13]Measured in polar projection.

[14]On occasions, special alignments occur, e.g. when there is no vertical disparity. In such cases one of the two stereoscopic constraints has an infinite uncertainty associated with it. Consequent combined inferences may have a large or infinite, uncertainty associated with them.

[15]Reliable enough to estimate the direction of least curvature in accordance with the bounds (6.36) and (6.38).

is an ellipse and stereoscopic evidence is available then A is the appropriate combined inference.

Key to table:

A (Fourfold ambiguity) Full curvature cannot be inferred. However, monocular analysis constrains the surface to be one of only four interpretations, (section 6.4.2). Each interpretation can be checked in turn to see if it is compatible with the two constraints that stereo has provided. This is done by extracting the direction of least curvature for the monocular interpretation and using it in conjunction with the stereo constraints to re-estimate the principal curvatures, (equation (5.34)). If this re-estimate disagrees with the original monocular interpretation's estimate of the principal curvatures then the interpretation can be discounted. All this can be done in addition to the first five "stereo-only" applications listed in table 6–2 (page 170), i.e. the locally spherical, cylindrical and hyperbolic ($k_1 = -k_2$) interpretations, the non-spherical test and the constraint graph.

B (Fourfold ambiguity) In the absence of stereoscopic constraints, monocular evidence alone constrains the surface to be one of only four interpretations, (section 6.4.2).

C (Fourfold ambiguity) Full curvature cannot be inferred. However, monocular analysis constrains the surface to be one of only four interpretations. In each interpretation the principal curvatures is only determined up to an arbitrary multiplicative constant, because the source radius is unknown, (section 6.4.2). This constant k can be usefully eliminated using the stereoscopic constraints as follows. Extract the principal curvature directions for each monocular interpretation. Rotate the stereoscopic measurements, $(r_1, r_2)^T$ and $(\delta n_1, \delta n_2)^T$ in 2-D, so that the r_1 component is aligned along the direction of greatest curvature. In the new frame the measurements are $(\acute{r}_1, \acute{r}_2)^T$ and $(\acute{\delta} n_1, \acute{\delta} n_2)^T$. Now the constant $k = -\acute{r}_1 / \acute{\delta} n_1$.

All this can be done in addition to the first five "stereo-only" applications listed in table 6–2.

D (Fourfold ambiguity and an arbitrary multiplicative constant) In the absence of stereoscopic evidence, monocular analysis alone constrains the surface to be one of only four interpretations. In each interpretation the principal curvatures is only determined up to an arbitrary multiplicative constant, because the source radius is unknown, (section 6.4.2).

E (Stereo constraints only) In the absence of monocular evidence, the first five "stereo-only" applications listed in table 6–2 can be used.

F (Full curvature) Full curvature can be inferred. Monocular analysis provides an estimate of the direction of least curvature (section 6.4.3), which can be used in conjunction with the stereoscopic constraints to obtain both principal curvatures (equation (5.34)). The stereo measurements can also be used to check that the surface is not locally planar—in this case the above procedure is invalid. The check is simply whether $|(\delta n_1, \delta n_2)^T|$ is not too close to zero. Finally another check should be made: that the bound (6.39) is not exceeded when evaluated using the newly inferred principal curvatures. Only if the bound is exceeded do the first five "stereo-only" applications listed in table 6–2 provide more information.

G (Direction of least curvature) In the absence of stereoscopic evidence, monocular analysis alone provides an estimate of the direction of least curvature (section 6.4.3).

6.5.1 Implementation of combined inference

The next chapter presents a computational system that partially implements the combine inference scheme described above. The system applies a full stereo analysis and all the "stereo only" interpretations listed in table 6–2. The monocular analysis for a compact source is fully implemented with the exception of

checking the bounds (6.36), (6.38) or (6.39). Uncertainties in the computed curvatures are calculated from the uncertainties in the stereoscopic and monocular measurements using the step-by-step gaussian error treatment, as was done for the stereoscopic analysis (section 6.3.4). The monocular analysis for a circular source is implemented in a reduced form: no fourfold ambiguities are considered. However, a circular source is still exploited by estimating the absolute ratio of principal curvatures $|k_1/k_2|$ when slant is close to zero. In this case it is easy to show (from equation (5.27), page 148), that $|k_1/k_2| \approx \sqrt{a}$, where a is the aspect ratio of the specular ellipse in polar projection and where $k_1 \gg \kappa_{VL}$ and $k_2 \gg \kappa_{VL}$. The system delivers the ratio without checking the latter condition. When the system indicates that $|k_1/k_2| \approx 1$ the stereoscopic interpretation for a locally spherical and hyperbolic ($k_1 = -k_2$) surfaces are presented. The former being omitted if the non-spherical test was passed. Further details of the implementation of the combined analysis are given in appendix I.2.

Chapter 7

Shape-from-specularity: a computational scheme

7.1 Introduction

An implementation of the shape-from-specularity inference scheme described in the previous section, is presented in this chapter. It consists of two parts:

1. A suite of programs that extract the low-level image descriptions that are used by the inference scheme.

2. An interactive, user friendly program (tool) that applies the inference scheme to the low-level descriptions and infers curvature information.

Figure 7–1 is a diagram showing the way in which information flows through the entire system: starting from a stereo image pair of a 3-D scene and ending with resultant curvature information. The thick line on the diagram marks the dividing line between the suite of programs and the interactive tool. The former all lie above the line: edge detection, specularity detection, specularity description, edge pruning, specularity matching, binocular stereo and depth extraction. These processes have all been discussed at length earlier in this thesis—details of how the programs are executed are given in appendix J. Below the line on the diagram are the constituents of the tool: an interactive choice of light source and specularity, iterative refinement of the depth data, and the

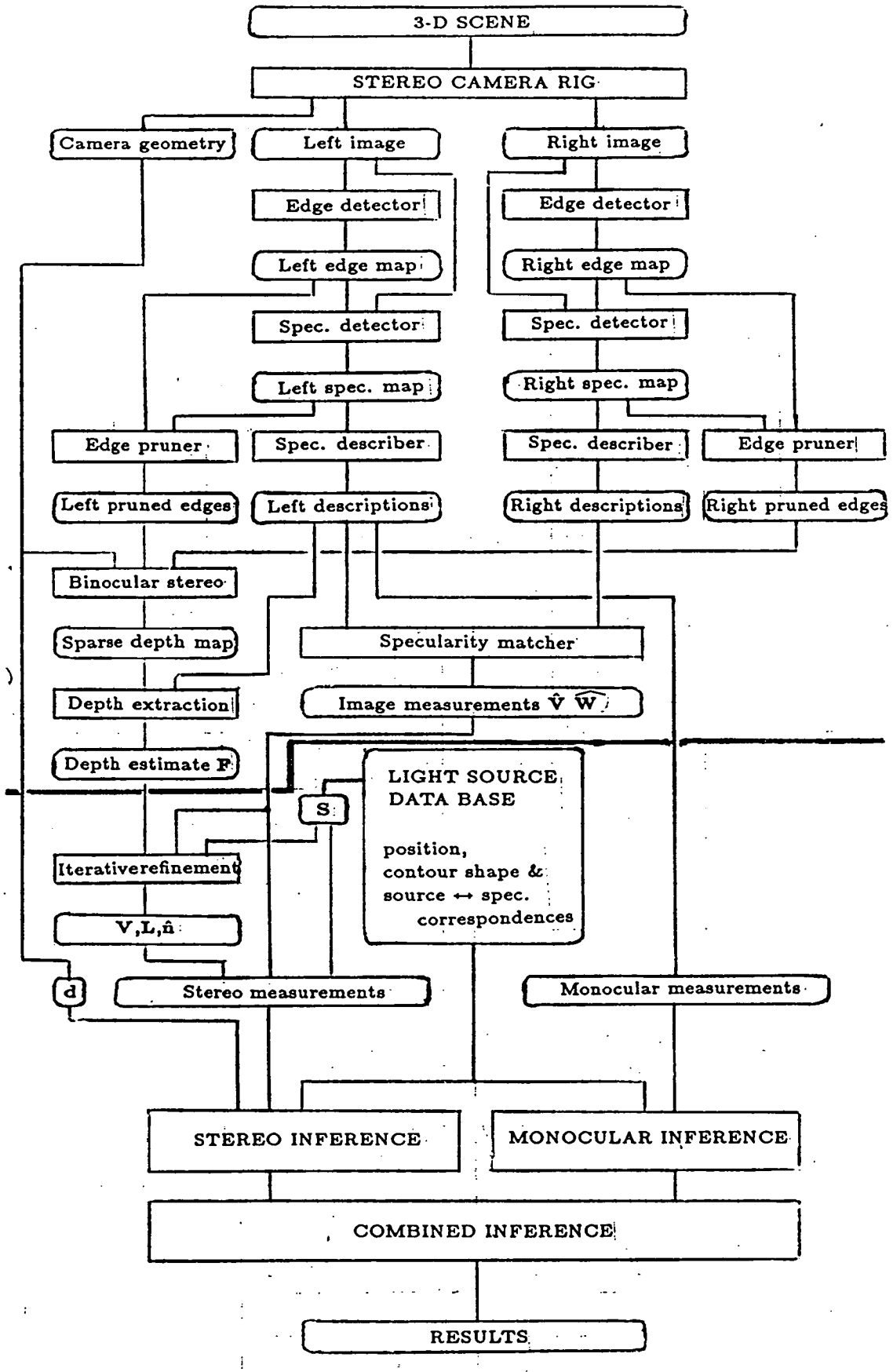


Figure 7–1: *Information flow in the entire system. Beneath the thick line lie the elements of the interactive tool.*

stereo, monocular and combined inference schemes. All these processes were described in detail in the preceding chapter.

Both the suite of programs and the interactive tool are written in the "C" programming language and run on a SUN-3 mini-computer under the UNIX 4.2 operating system. A description of the tool and an example of it in action are given below. Following that, are a presentation and evaluation of the results provided by the scheme, when it is applied to a set of stereo image pairs. The scheme performs robustly and well for a wide variety of shapes.

7.2 The interactive tool

The interactive tool runs in the SunView colour graphics environment, so the Suntools window system must be invoked before it can be used:

```
% suntools
```

Once the window system is ready and a `shelltool` is available, the user runs the tool by typing the following into the `shelltool`:

```
%tool picture/tv/scene
```

The program `tool` takes as its argument one or more Unix directories. Each directory contains all the files—created by the suite of programs—necessary to apply the inference scheme. In this case the directory `picture/tv/scene` is specified.

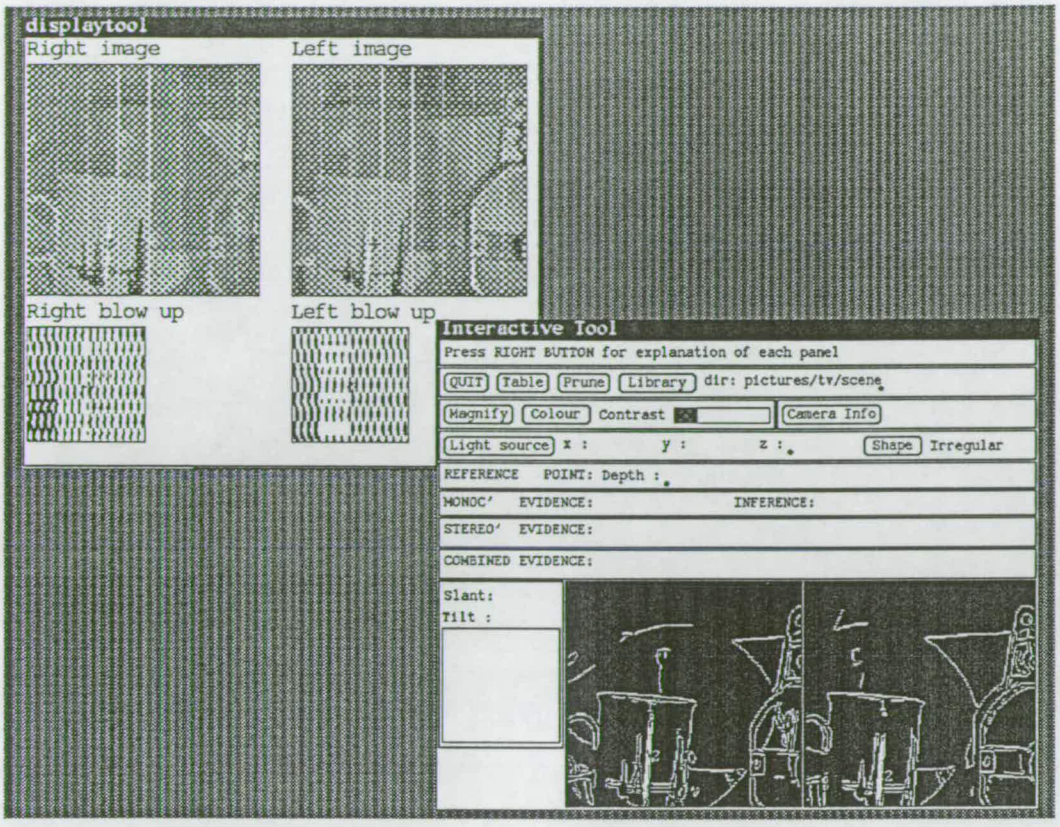


Figure 7–2:

After a while two new tools appear on the window system:

1. A "`displaytool`".

2. The shape-from-specularity `tool` itself.

The `displaytool` displays the original stereo image pair from which the specularities have been detected and from which the inferences are to be made.

Three different modes[1] of display are available:

1. The grey-level images themselves[2].

2. The grey-level images with red blobs superimposed where specularities have been detected.

3. An anaglyph image that facilitates direct depth perception when viewed through special glasses—a red filter is placed over the right eye and green over the left.

A single region of the `tool` provides control over the display options. The panel

Figure 7–3:

titled *colour* allows any of the three display modes to be selected. The slider titled *contrast* allows the distribution of grey tones in the images to be altered—in case the images appear too light or too dark. In the first two display modes a green cursor is drawn on each image to show the current point of interest. A blow-up of the region around each cursor is also drawn. The user can adjust the magnification of the blow-up using the *magnify* panel.

The interactive `tool` is simple to use. For example, the mouse is placed within the box in the `tool`'s lower right corner, clicked (using the left button) on top of a marked specularity (here number 2 is used) and the *combined inference* is selected (through the right button).

---

[1]Users with only a monochrome screen can switch off the `displaytool` using the option `-n`—the binary images at the bottom of the `tool` provide sufficient spatial information by themselves..

[2]The right image is placed to the left of the left image to facilitate "cross-eyed" stereo depth perception. The same is done for the binary images in the `tool`.

Immediately the results are displayed:

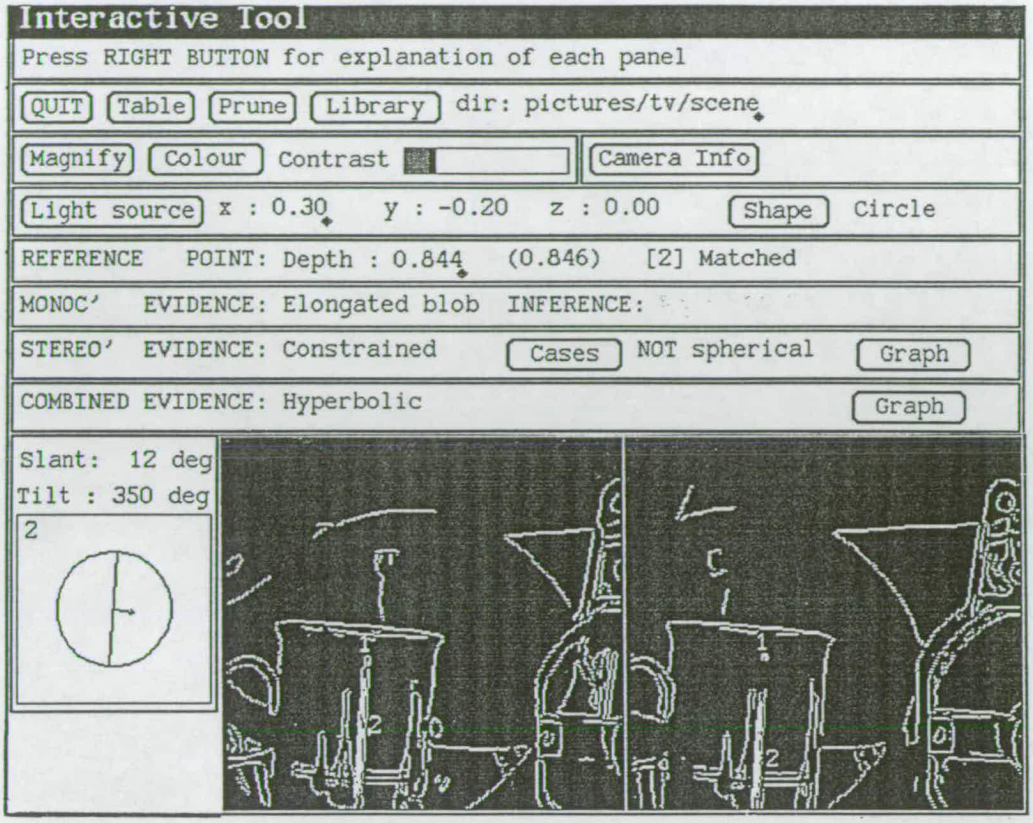


Figure 7–4:

- The word *matched* is written to confirm that the chosen specularity matches with one in the opposite image. The numbers next to each marked specularity indicate a matching pair.

- The nearby surface marking (reference point), at which depth was extracted is marked on the `displaytool` as a yellow square. In addition the *depth* is written out.

- The depth estimate after iterative refinement is added in brackets.

The iterative refinement was applied using the light source position with the (x,y,z) coordinates indicated. By clicking the right mouse button on top of the *light source* panel, lights of different positions and shapes can be selected from the data-base. Iterative refinement also estimates the local surface orientation.

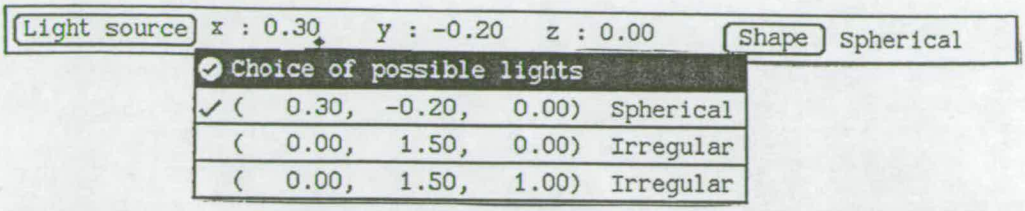


Figure 7–5:

The box at the lower left of the `tool` depicts this orientation. The ellipse-and-arrow diagram provide the user with a direct perception of the computed values of slant and tilt. The monocular inference depends on the shape of the

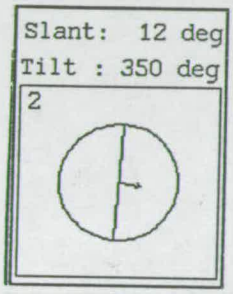


Figure 7–6:

specularity. A brief description is given: *elongated blob*—pressing the right mouse button provides the details: The outcome of the monocular inference is

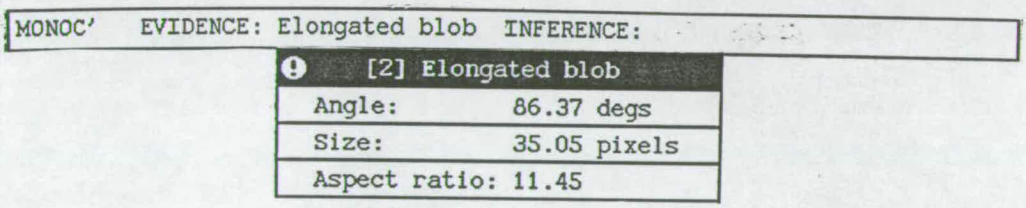


Figure 7–7:

indicated in two ways:

1. The long straight line on the arrow-ellipse diagram marks the projection of the direction of least curvature (magnitude) onto the (left) image.

2. A rough estimate of the absolute ratio of the principal curvatures is written as the *ratio*.

Stereo inference provides its own constraints. Their existence is indicated by the word *constrained*. Use of the mouse reveals the estimated values of two of

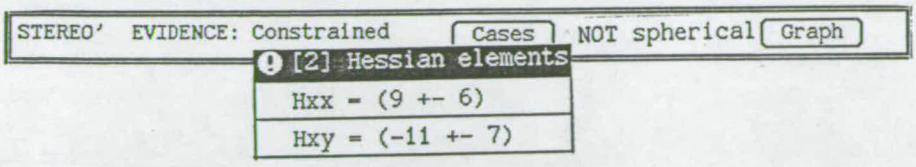


Figure 7–8:

the hessian's elements. The panel marked *cases* shows the parameters obtained if the local surface shape is known. In this case only the cylindrical case is relevant. The words *not spherical* show that the non-spherical test was passed.

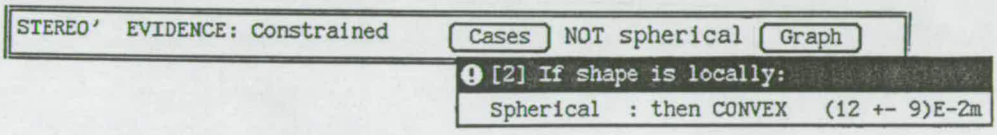


Figure 7–9:

The *graph* panel initiates a `graphtool` which depicts, in grey, the region of (r_1, r_2) space[3] in which stereo inference constrains the surface to lie.

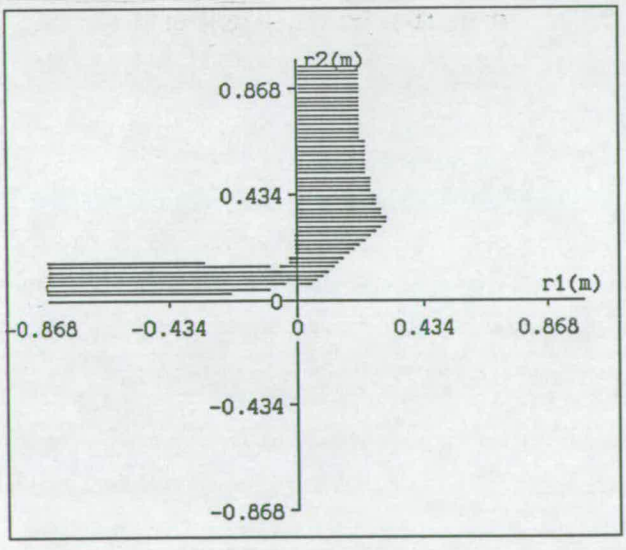


Figure 7–10:

[3]Recall r_1 and r_2 are simply the reciprocals of the of principal curvatures k_1 and k_2.

Applying the combined inference scheme constrains the surface to lie in a small zone within the existing region. The lower *graph* panel creates a `graphtool` that shows this zone in solid black. The word *hyperbolic* provides

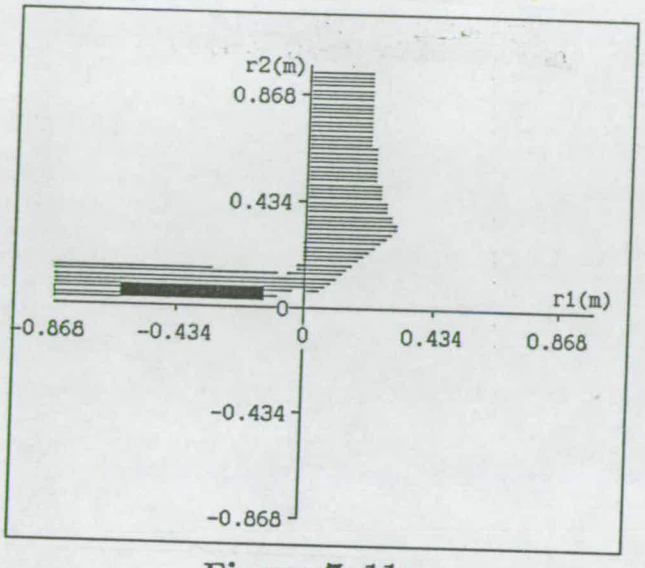


Figure 7–11:

a general description of the local shape of the surface. The estimates of the principal radii of curvature are also available: they can be either read off the `graphtool` or revealed using the right mouse button. These radii are consis-

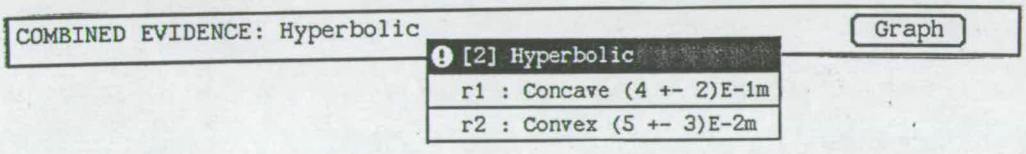


Figure 7–12:

tent with the real 3-D surface: the mug has a radius (r_2) of about 4 cm. The other radius of curvature, r_1 is estimated to be an order of magnitude larger—so although technically the surface is locally hyperbolic, it is locally close to cylindrical.

Running the system on the other pair of specularities (number 1), provides another example. This time, monocular inference for a circular source fails[4]

[4]A fourfold ambiguity would be obtained in this case.

because the slant is too large. Monocular inference for a compact source fails as the specularity is not elongated—as indicated by the inscribed ellipse on the ellipse-and-arrow diagram.

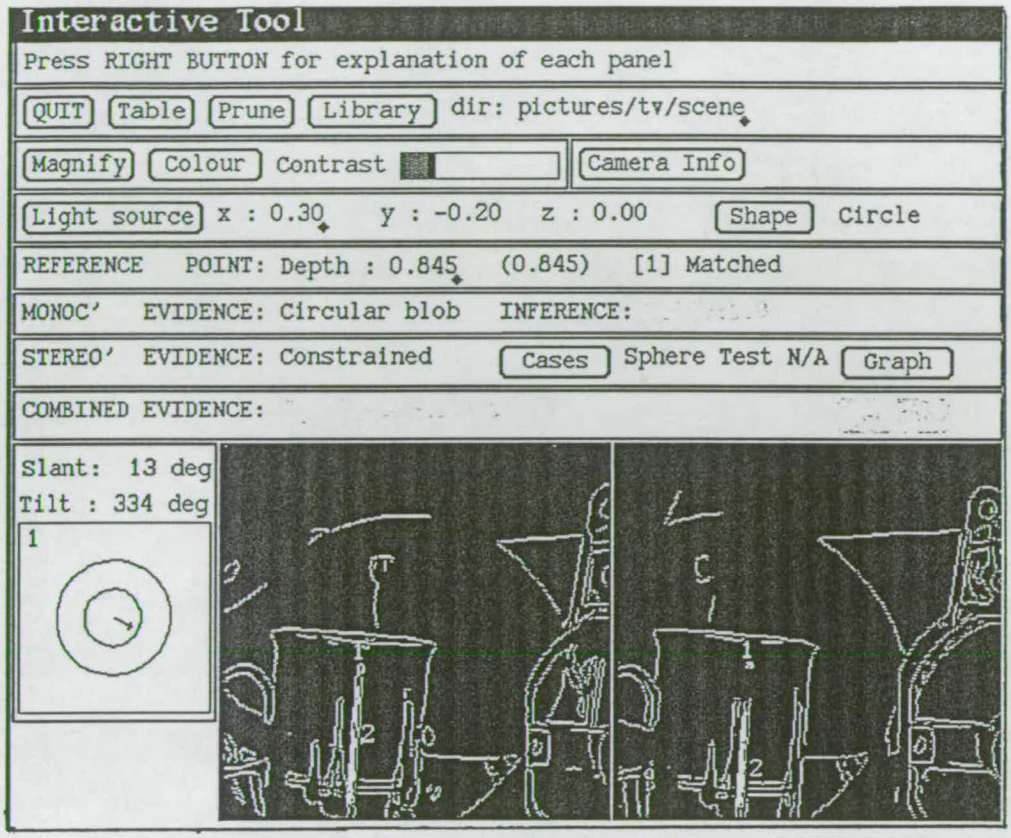


Figure 7–13:

However, the stereo analysis constrains the allowable (grey) region in (r_1, r_2) space.

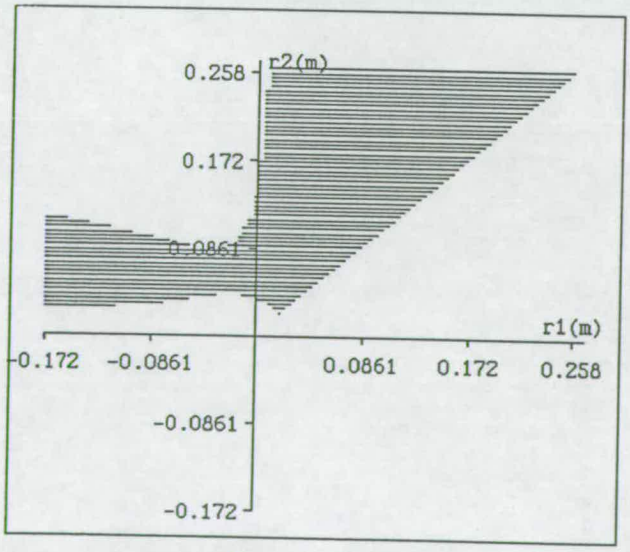


Figure 7–14:

As the source is circular and the specularity has an aspect ratio around 1, the user might suspect that $|r_1| \approx |r_2|$. This would certainly be true if slant were zero. Pressing the *Cases* button reveal the two possible surface interpretation when $|r_1| \approx |r_2|$: locally spherical or hyperbolic ($r_1 = -r_2$). However the parameters

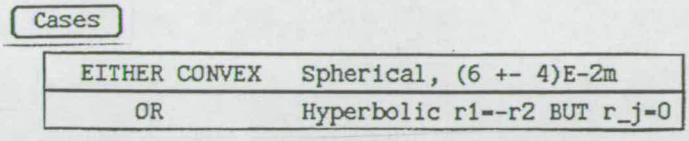


Figure 7–15:

accompanying the latter are unavailable because the vertical disparity is too small to permit accurate computation. For the same reason the non-spherical test could not be applied. In fact the surface is part of the lip of the mug and is hyperbolic.

A few panels on the `tool` remain to be explained. *Camera info* when pressed

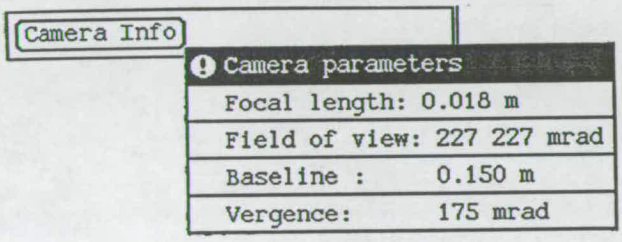


Figure 7–16:

gives the stereo viewing geometry. At the top of the `tool` is the black name banner. Below that, messages are written to aid the user. Through the *library*

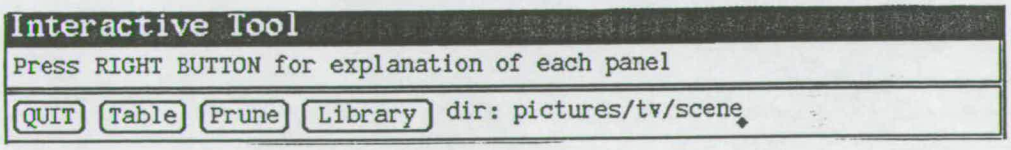


Figure 7–17:

panel new stereo pairs can be selected for inference. The *dir* panel indicates the current directory from which input data is read. The *prune* panel allows the

user to switch the binary images in the boxes at the bottom of the tool into the edge maps, pruned of any specular edges. By pressing *table* the `tool` goes into "auto-pilot" applying the inference scheme to all available specular pairs and prints out a table of the results.

7.3 Results: presentation

The inference scheme was applied to the set of images shown of figures 7–19–7–31. These figures also show the results obtained in the form of a constraint graph for each specularity. A constraint graph represents (r_1, r_2) space. It indicates the regions in which the stereo and combined stereo and monocular analyses compute the radii of curvature to lie. Also shown on each figure are the slant and tilt values and an edge map upon which the detected specularities are labelled. Both "real" and "artificial" images were processed:

- Images of specular surfaces of realistic appearance, artificially generated using the IBM CAD body modeller Winsom [95], are shown in figures 7–19–7–26. These images are useful for verifying the results of the inference scheme because:
  - They are noiseless so the scheme can be assessed independently of the noise characteristics of any particular imaging device.
  - The position and shape of the light sources, corresponding to each specularity, are precisely known and can be easily controlled.
  - The stereo viewing geometry is also precisely known. The same geometry—indicated in figure 7–18—applies for all these images.
  - Accurate values of the radii of curvature, extracted from the surface models, can be checked against the results obtained. On each (r_1, r_2) graph, the principal radii extracted from the model surface are marked with a white cross. An indication of the scheme's success is a white cross marked in the region it has inferred. Of course, if either r_1 or r_2 are

Camera parameters	
Focal length:	10.000 m
Field of view:	100 100 mrad
Baseline :	1.100 m
Vergence:	105 mrad

Figure 7–18: *Camera parameters for the CAD model generated images.*

infinite (as for cylinders), the white cross cannot be marked. Instead a white box is marked at the closest available spot.

Note: figures 7–23–7–26. do not show the surface markings at which depths were estimated. For these images estimates were extracted directly from the CAD model.

- Images of real surfaces, acquired using a stereo camera rig, are shown in figures 7–27–7–31. Figure 7–18 shows the camera parameters corresponding to these images. The cameras are as described in chapter 4. The curvatures inferred by the scheme are compared with estimates measured from the real surfaces. Again these measurements are marked as white crosses or boxes on the (r_1, r_2) graph.

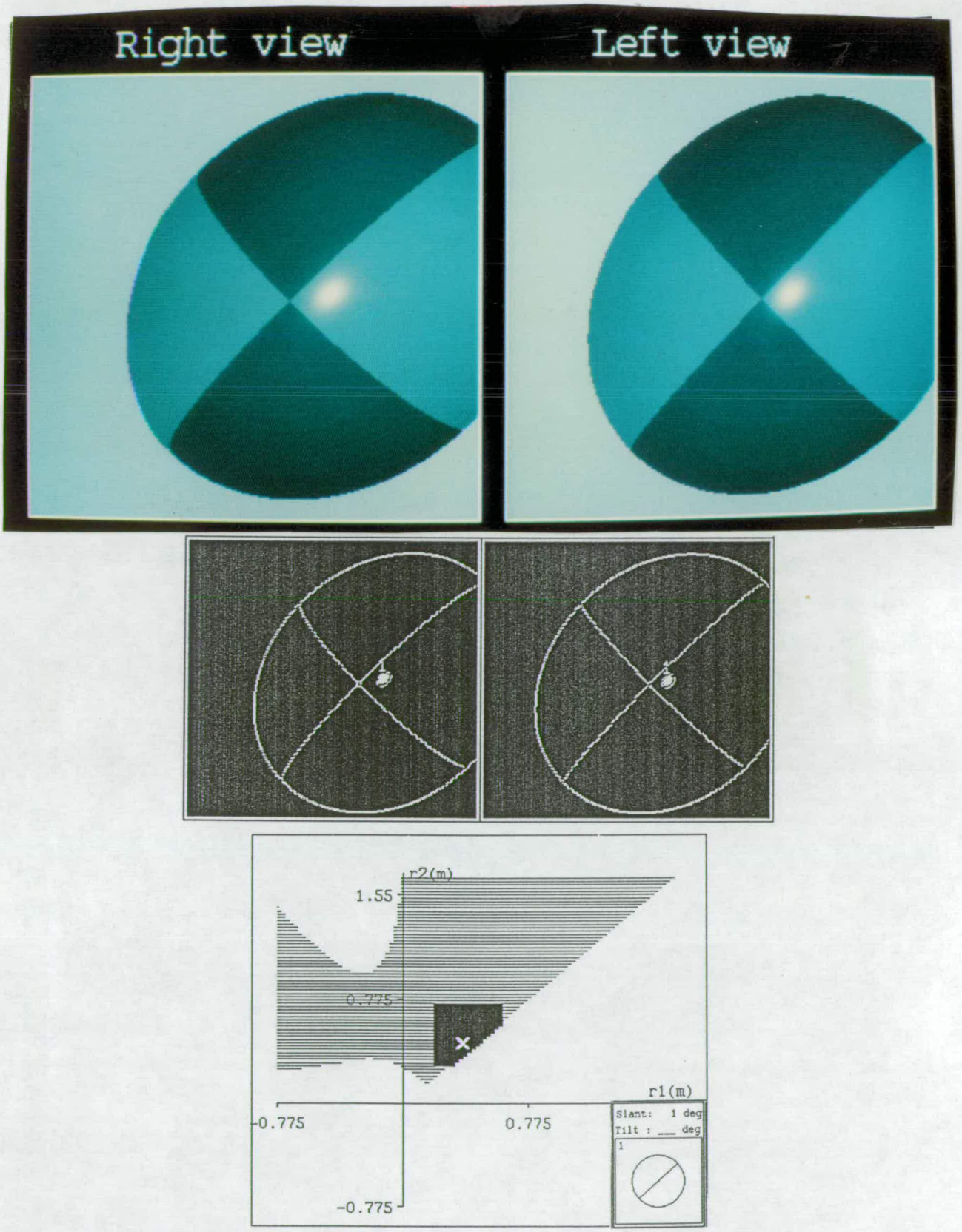


Figure 7–19: Top: *Stereo-pair of convex ellipsoid (CAD model)—arranged for binocular fusion.* **Centre:** *Edge map with superimposed, labelled specularity.* **Bottom:** *The system restricts principal curvatures to the grey region of the graph by stereo analysis alone and to the black region by combining stereo and monocular analyses. Actual curvatures are indicated by the white cross.*

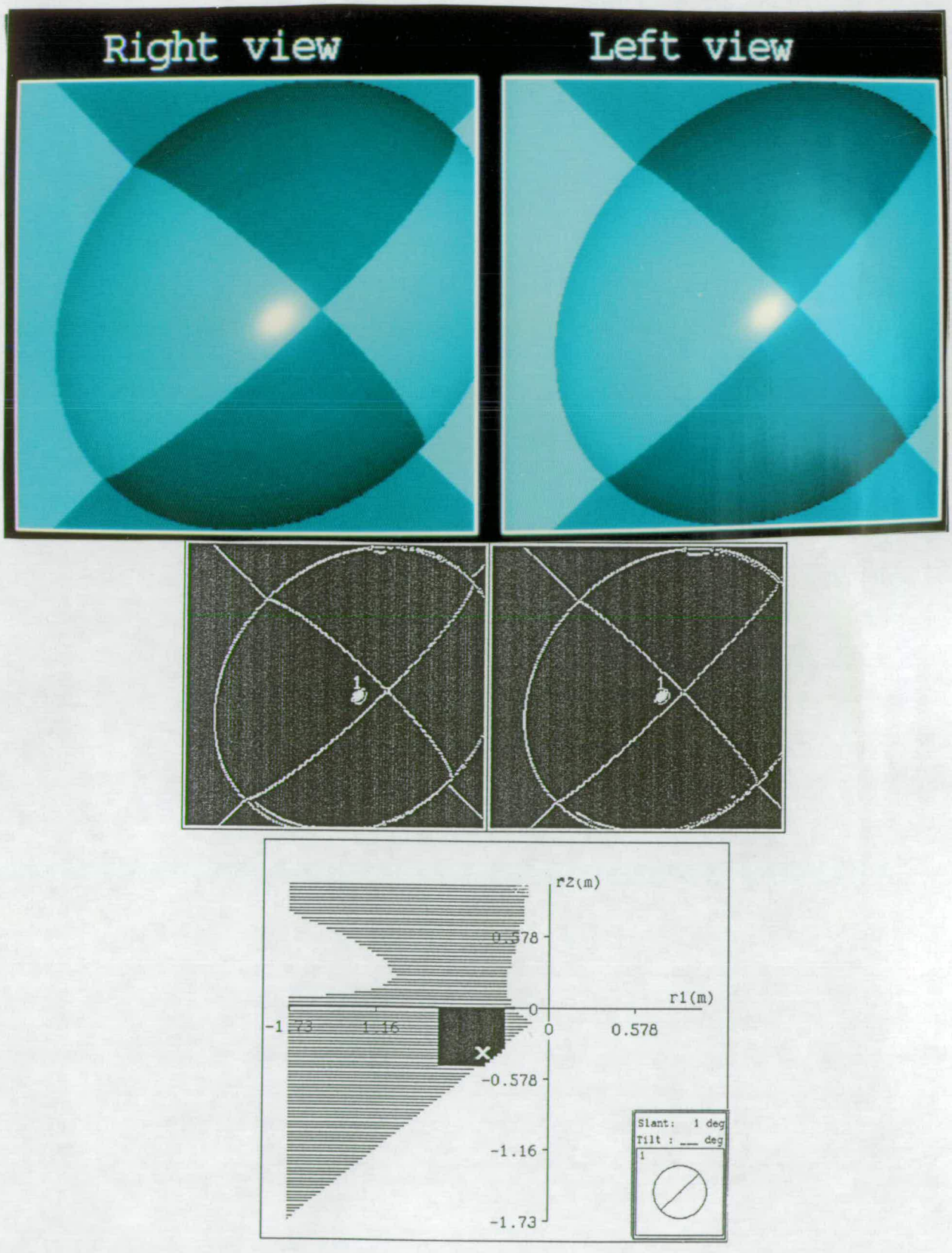


Figure 7–20: Top: *Stereo-pair of concave ellipsoid (CAD model)—arranged for binocular fusion.* **Centre:** *Edge map with superimposed, labelled specularity.* **Bottom:** *The system restricts principal curvatures to the grey region of the graph by stereo analysis alone and to the black region by combining stereo and monocular analyses. Actual curvatures are indicated by the white cross.*

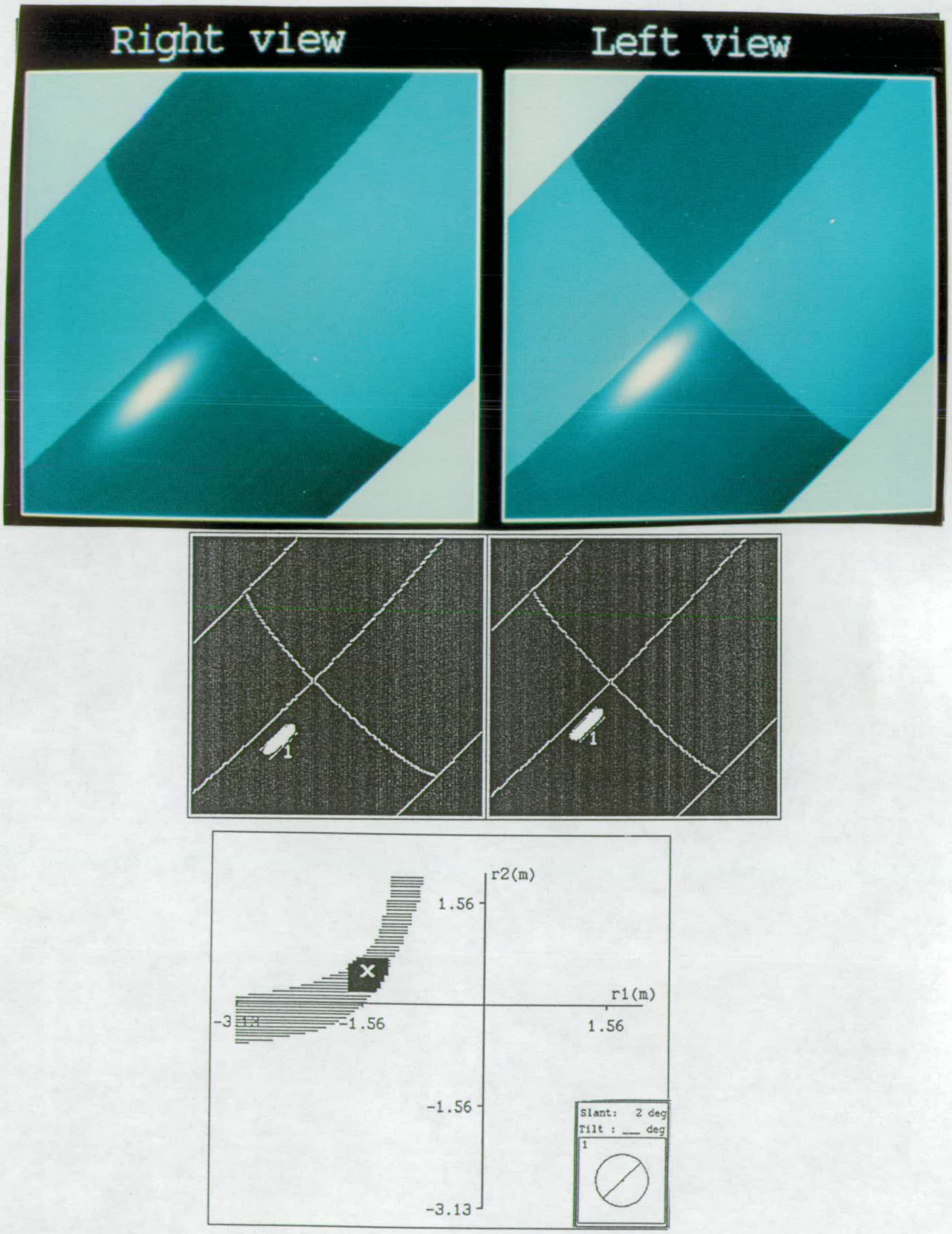


Figure 7–21: **Top:** *Stereo-pair of inner surface of torus (CAD model)—arranged for binocular fusion.* **Centre:** *Edge map with superimposed, labelled specularity.* **Bottom:** *The system restricts principal curvatures to the grey region of the graph by stereo analysis alone and to the black region by combining stereo and monocular analyses. Actual curvatures are indicated by the white cross.*

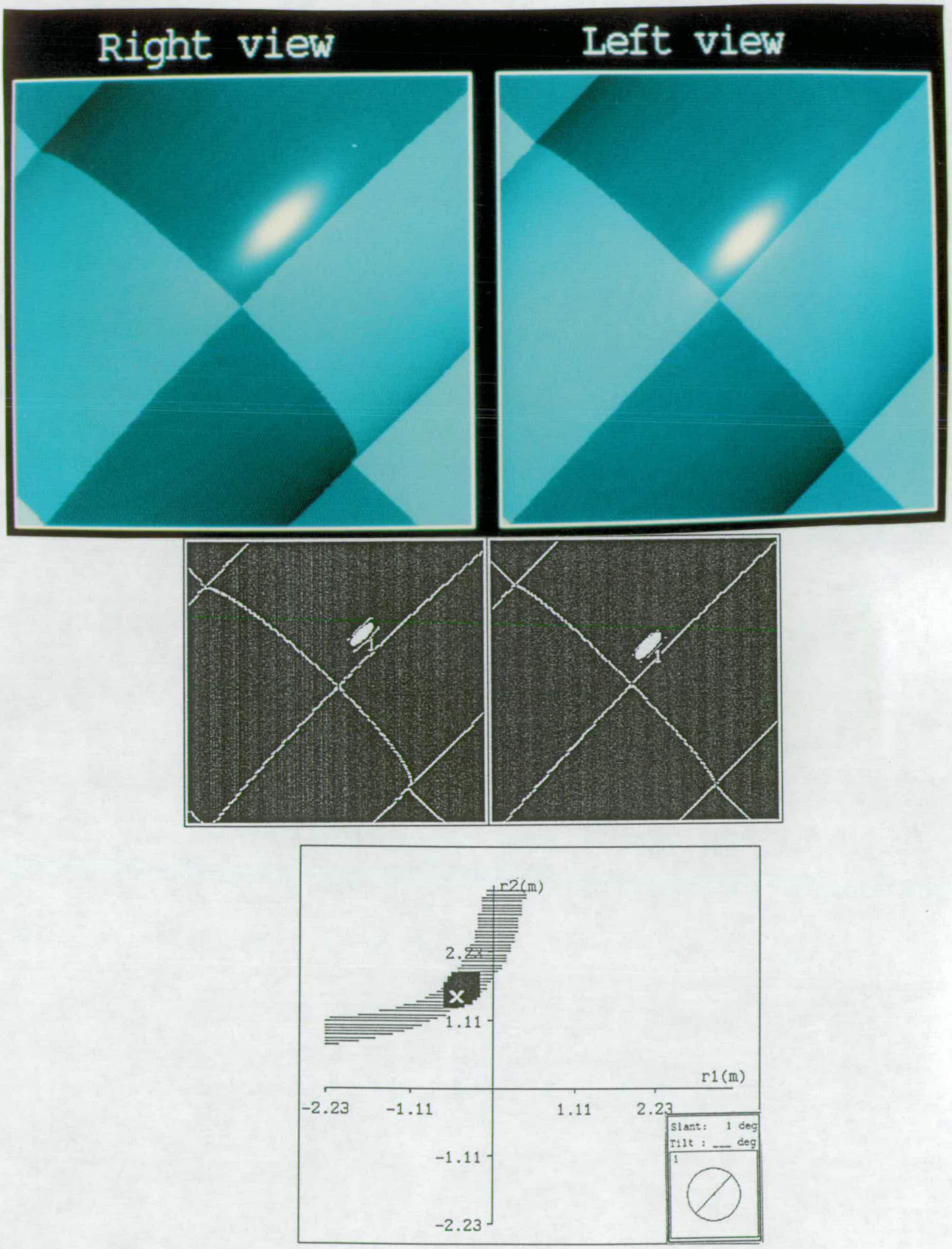


Figure 7–22: Top: *Stereo-pair of hyperbolic surface (CAD model)—arranged for binocular fusion.* **Centre:** *Edge map with superimposed, labelled specularity.* **Bottom:** *The system restricts principal curvatures to the grey region of the graph by stereo analysis alone and to the black region by combining stereo and monocular analyses. Actual curvatures are indicated by the white cross.*

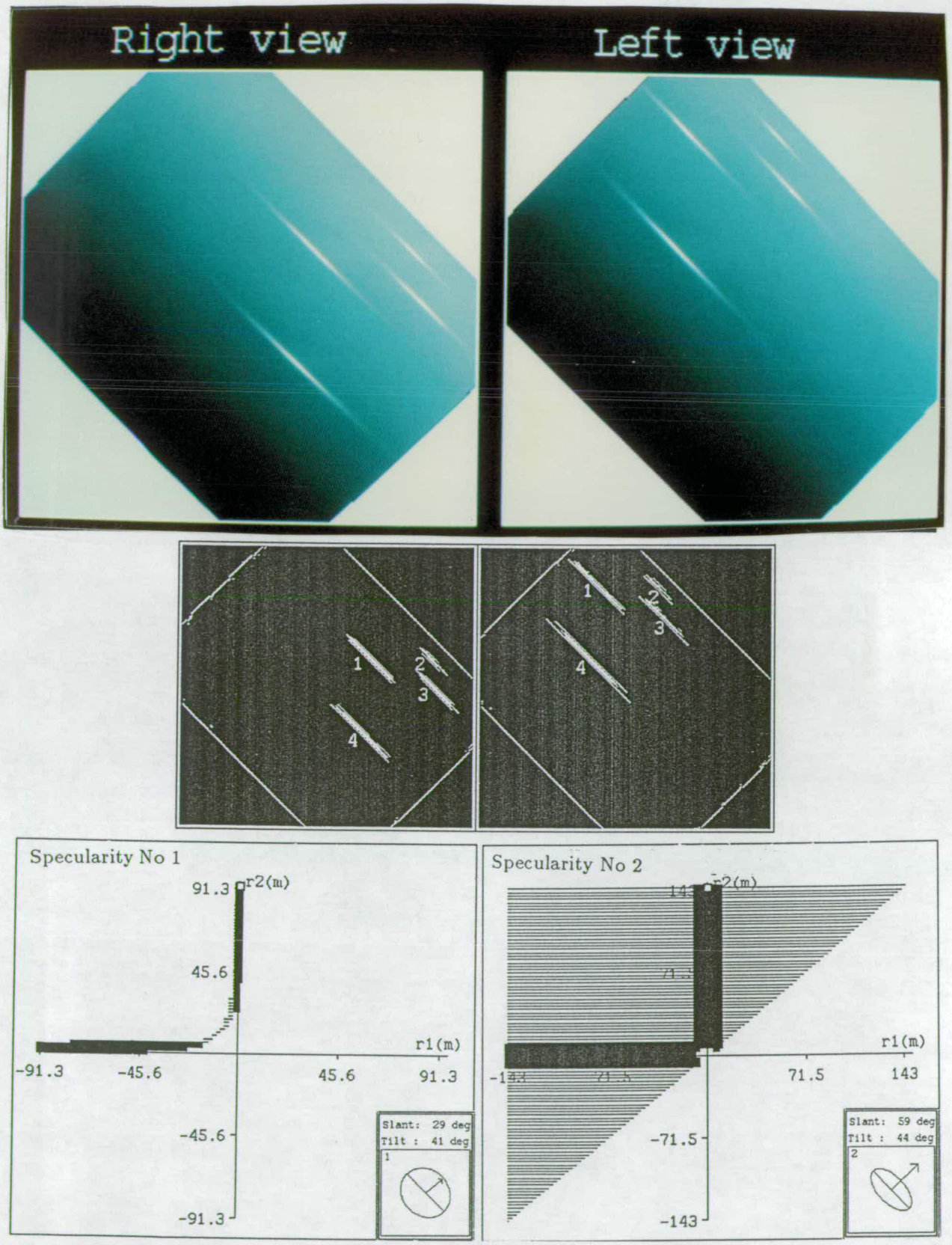


Figure 7–23: Top: *Stereo-pair of convex cylinder (depth estimated from CAD model)—arranged for binocular fusion.* **Centre:** *Edge map with superimposed, labelled specularities.* **Bottom:** *The system restricts principal curvatures to the grey regions of the graphs by stereo analysis alone and to the black regions by combining stereo and monocular analyses. Actual radii of curvature are* $(0.4, \infty)$. **Continued over page.**

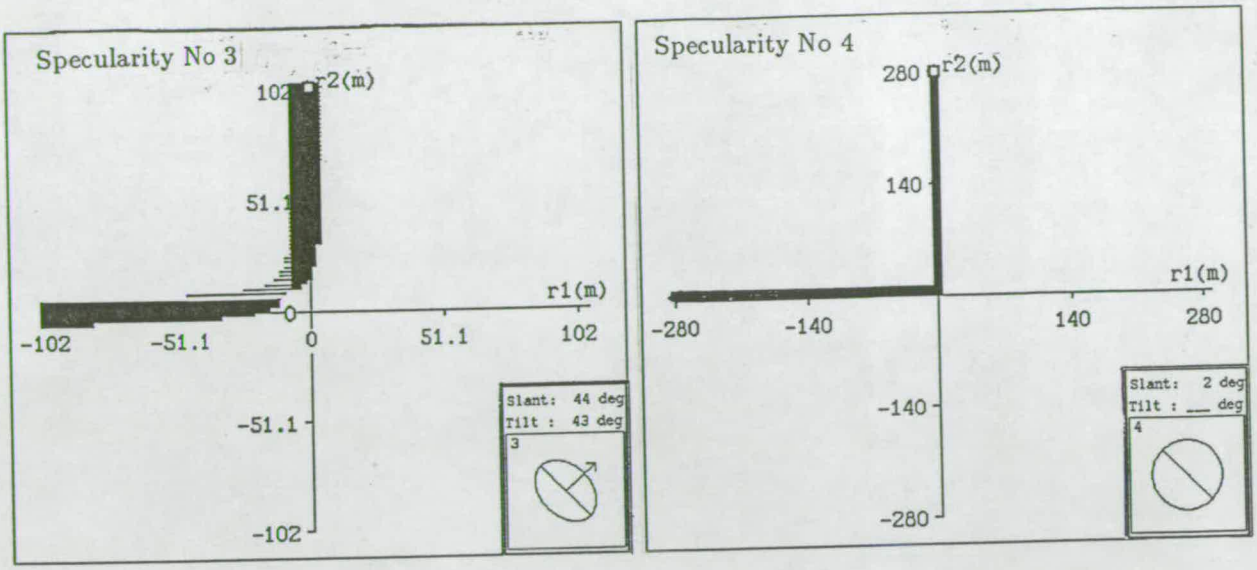


Figure 7–23 cont. *Convex cylinder. The system restricts principal curvatures to the grey regions of the graphs by stereo analysis alone and to the black regions by combining stereo and monocular analyses. Actual radii of curvature are* $(0.4, \infty)$.

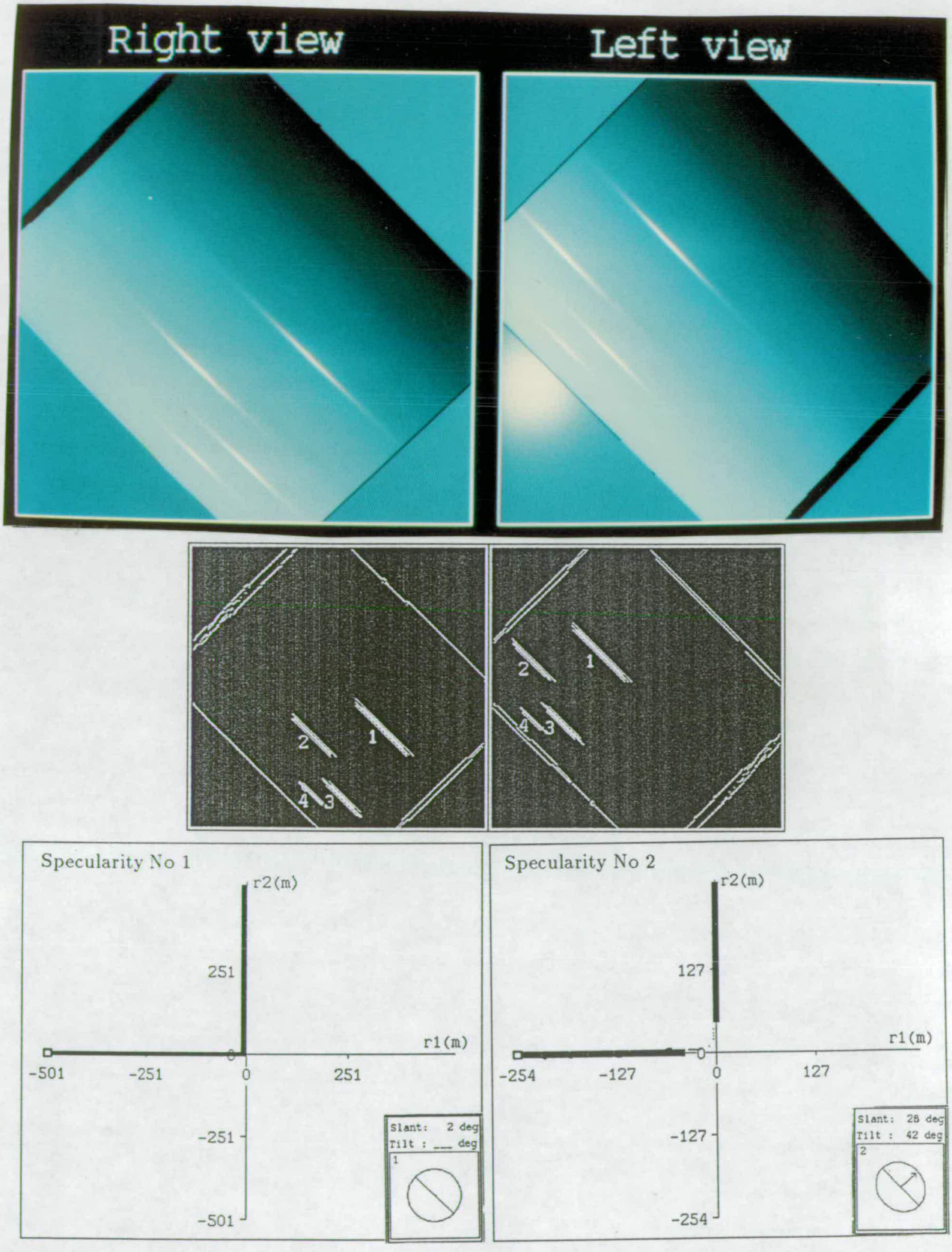


Figure 7–24: Top: *Stereo-pair of concave cylinder (depth estimated from CAD model)—arranged for binocular fusion.* **Centre:** *Edge map with superimposed, labelled specularities.* **Bottom:** *The system restricts principal curvatures to the grey regions of the graphs by stereo analysis alone and to the black regions by combining stereo and monocular analyses. Actual radii of curvature are* $(-0.4, \pm\infty)$. **Continued over page.**

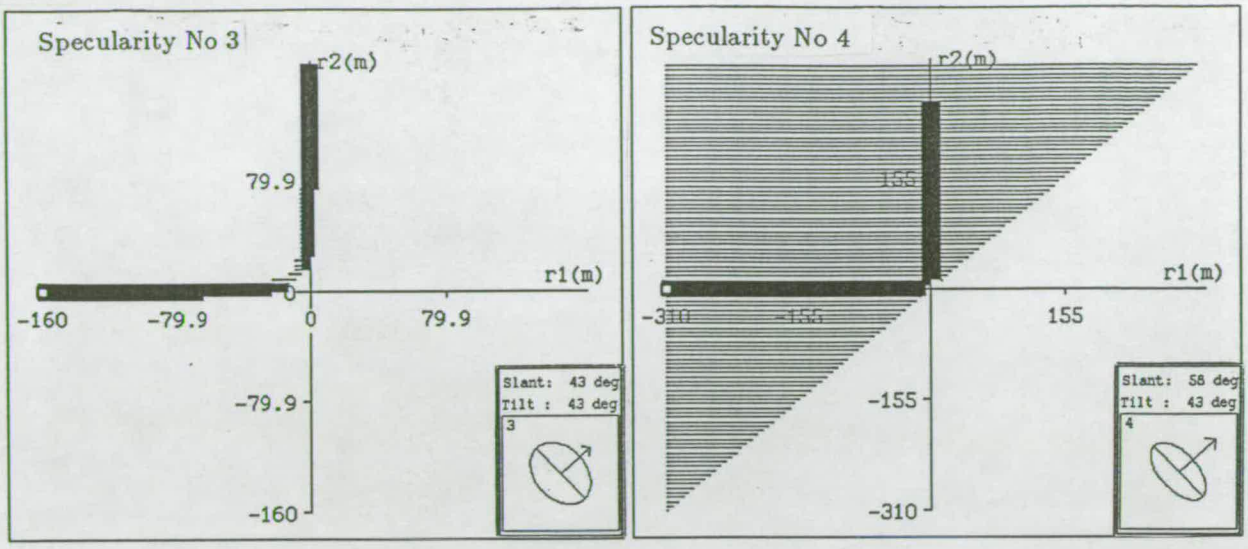


Figure 7–24 cont. *Concave cylinder. The system restricts principal curvatures to the grey regions of the graphs by stereo analysis alone and to the black regions by combining stereo and monocular analyses. Actual radii of curvature are* $(-0.4, \pm\infty)$.

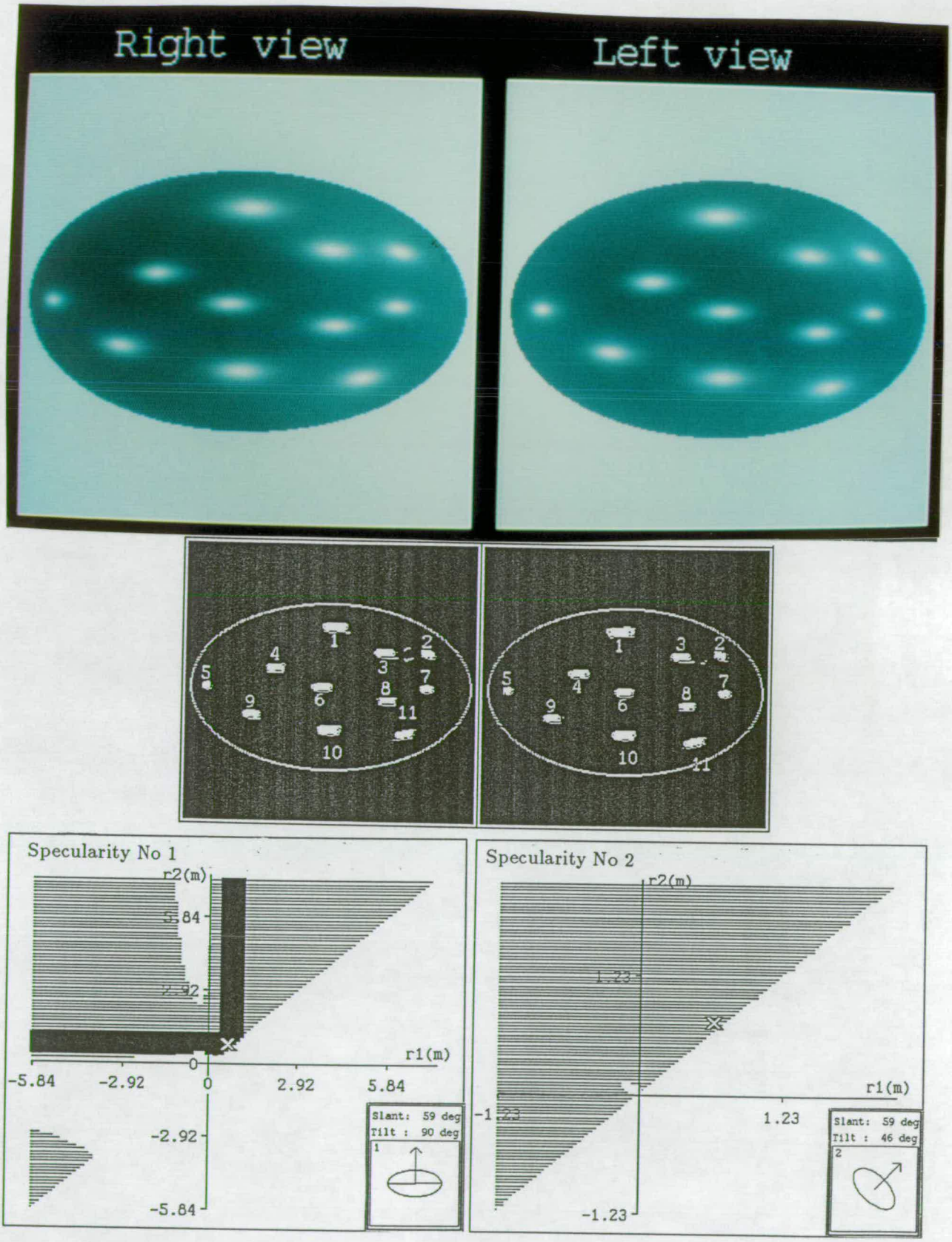


Figure 7–25: Top: *Stereo-pair of convex ellipsoid (depth estimated from CAD model)—arranged for binocular fusion.* **Centre:** *Edge map with superimposed, labelled specularities.* **Bottom:** *The system restricts principal curvatures to the grey regions of the graphs by stereo analysis alone and to the black regions by combining stereo and monocular analyses. Actual radii of curvature are indicated by white crosses.* **Continued over page.**

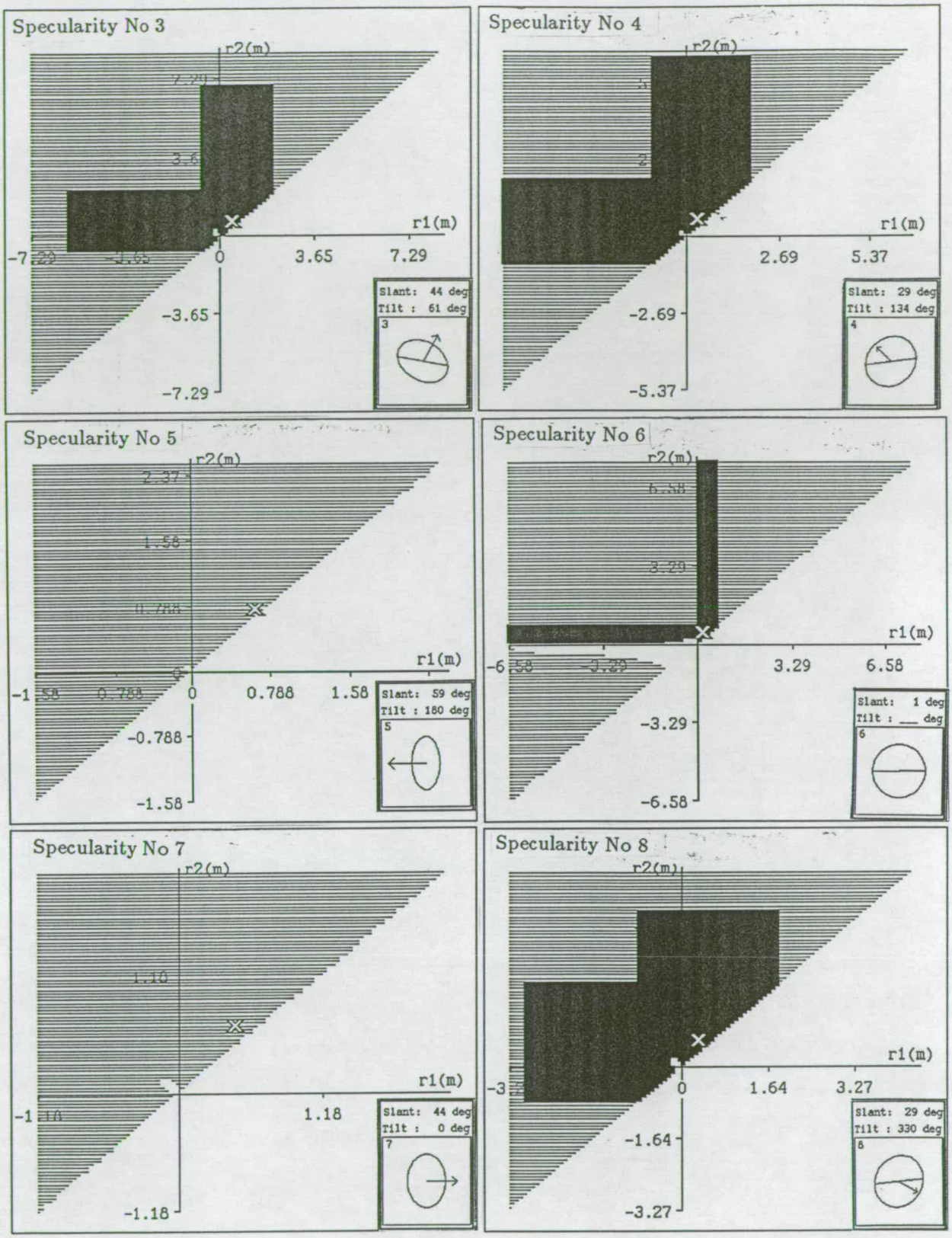


Figure 7–25 cont. *Convex ellipsoid. The system restricts principal curvatures to the grey regions of the graphs by stereo analysis alone and to the black regions by combining stereo and monocular analyses. Actual radii of curvature are indicated by white crosses.*
Continued over page.

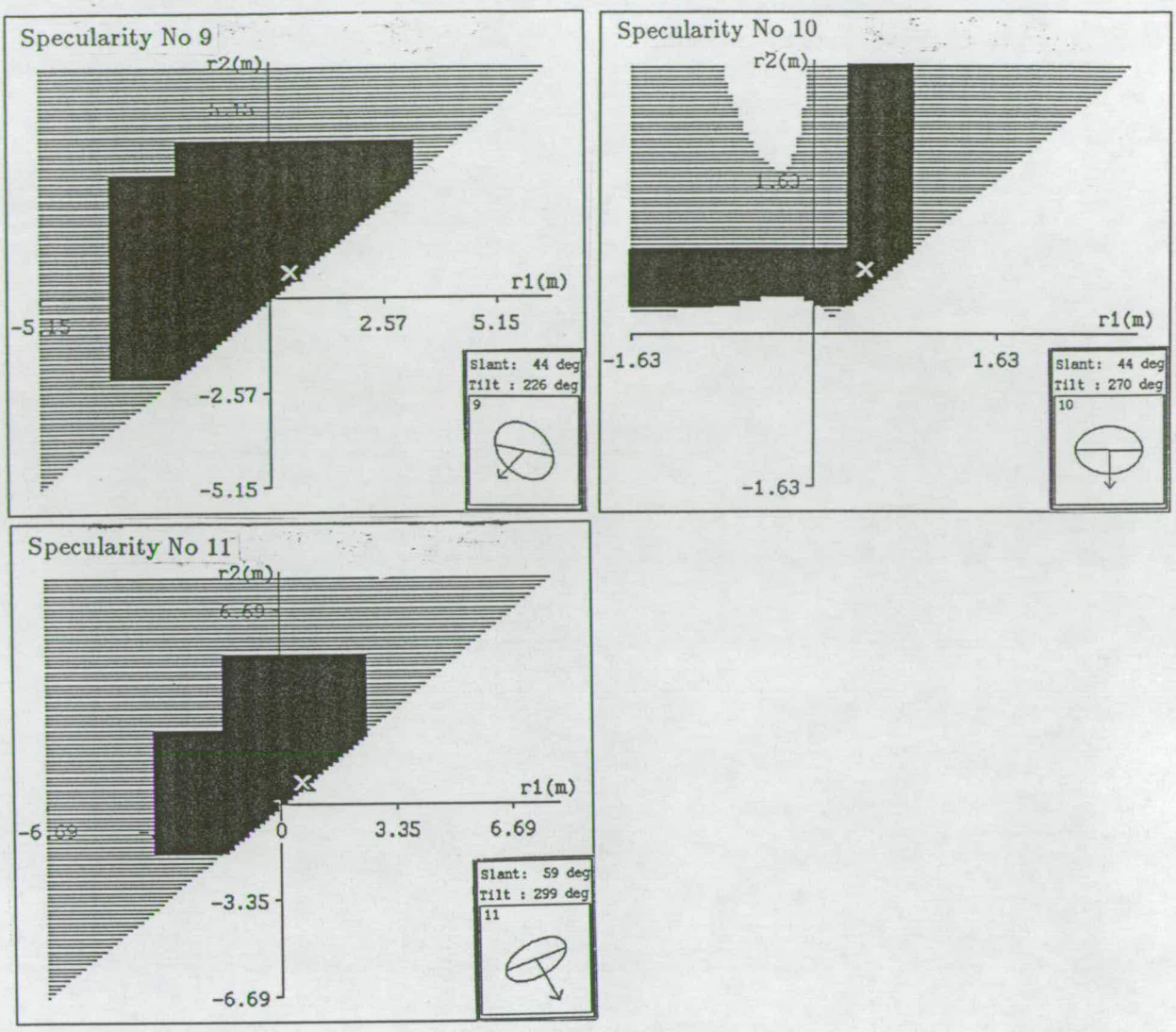


Figure 7–25 cont. *Convex ellipsoid. The system restricts principal curvatures to the grey regions of the graphs by stereo analysis alone and to the black regions by combining stereo and monocular analyses. Actual radii of curvature are indicated by white crosses.*

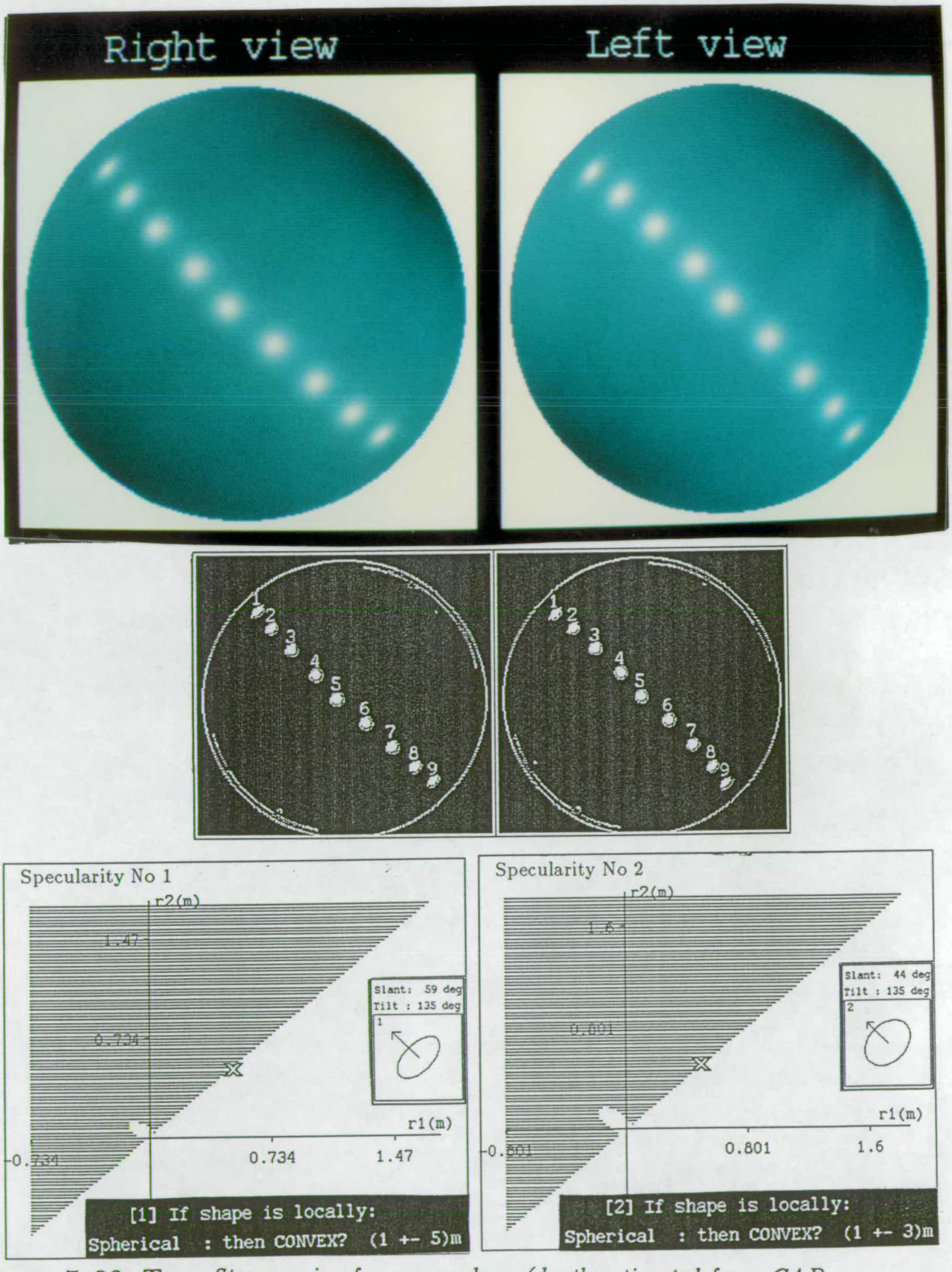


Figure 7–26: Top: *Stereo-pair of convex sphere (depth estimated from CAD model)—arranged for binocular fusion.* **Centre:** *Edge map with superimposed, labelled specularities.* **Bottom:** *The system restricts principal curvatures to the grey regions of the graphs by stereo analysis alone. Monocular analysis fails. Actual radii of curvature are indicated by white crosses.* **Continued over page.**

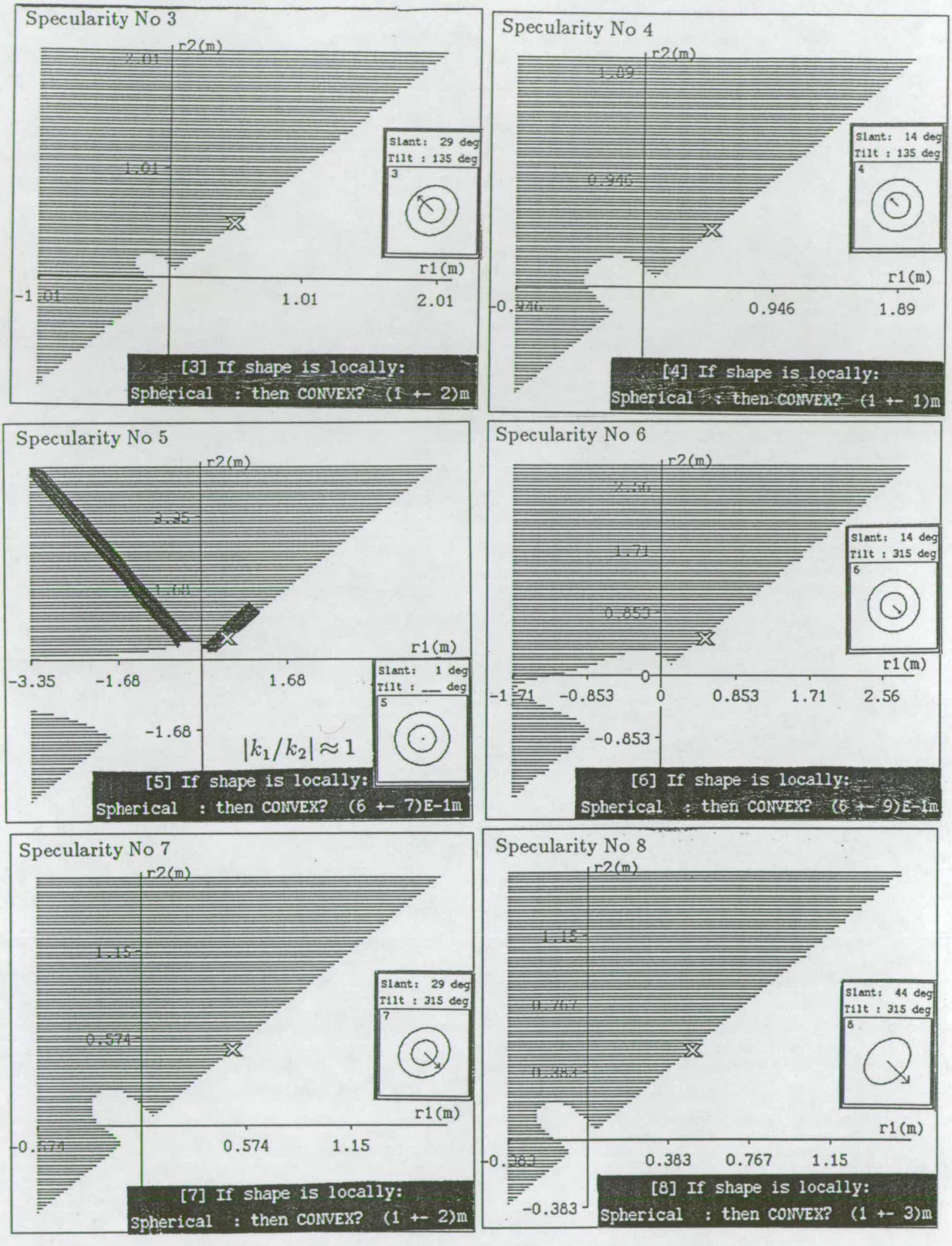


Figure 7–26 cont. *Convex sphere. The system restricts principal curvatures to the grey regions of the graphs by stereo analysis alone. Monocular analysis fails, except for specularity No 5, where combining it with stereo restricts curvatures to the black regions. Actual radii of curvature are indicated by white crosses.* **Continued over page.**

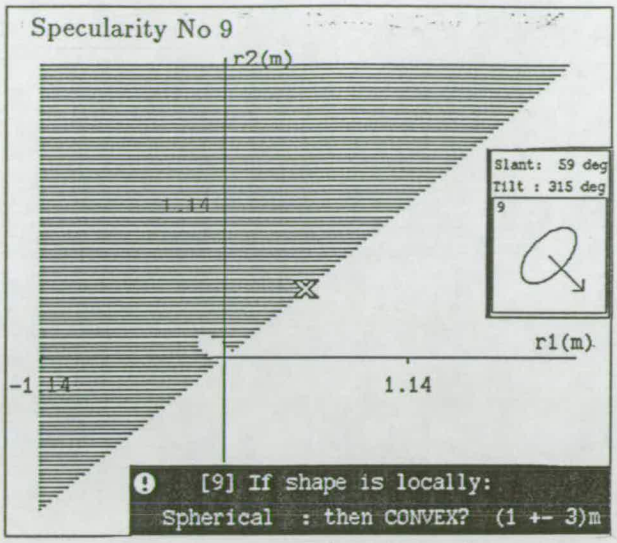


Figure 7–26 cont. *Convex sphere. The system restricts principal curvatures to the grey region of the graph by stereo analysis alone. Monocular analysis fails. Actual radii of curvature are indicated by white cross.*

7.4 Results: evaluation

The results in figures 7–19–7–31 show the scheme successfully detecting, matching and using specularities to infer curvature information, for a variety of surface shapes, at a variety of surface orientations. For these figures the detection and matching tasks were not difficult—so the evaluation below chiefly concerns the quality of the inferences. Some conditions prove more favourable for making inferences, than others, e.g:

- When the disparity of the specularity between left and right images, relative to the gross disparity of the surface, can be accurately measured.

- When the specularity has a shape suitable for monocular inference.

In such conditions both radii of curvature, r_1, and r_2 can often inferred with relatively little uncertainty—e.g. to within the black box in the (r_1, r_2) graph shown in figures 7–19–7–22. In adverse conditions the scheme's performance degrades gracefully—it assigns appropriate uncertainities to the inferred values of r_1 and r_2, to account for the prevailing uncertainties. This is made possible by the propagations of uncertainty through the scheme—as described in chapter 6. For example:

- There is almost no vertical disparity of the specularity No 1 on the ellipsoid in figure 7–25. Thus the uncertainty in its measurement is very large. As a result, one of the inferred radii of curvature is assigned a very large uncertainty. For this reason the (r_1, r_2) graph contains an elongated black line[5].

---

[5]The switch of direction of the black line at the origin is only due to definition of r_1 and r_2 being such that $r_1 \leq r_2$.

- In the same figure, specularity No 2 is the wrong shape to make an monocular inference: it is neither circular nor sufficiently elongated. In this case, stereo analysis alone restricts the curvatures to lie within a region of (r_1, r_2) space, depicted in grey on the graph. This is a much weaker curvature constraint because a much larger region of (r_1, r_2) space is indicated. However, it has the benefit of *not* being wrong.

There are uncertainities present in the measurements used for both the monocular and stereo inferences. Those in stereo measurements produce the most significant effect because—as mentioned above—monocular inferences are avoided when large uncertainties exist. Table 6–1 on page 170 lists the various uncertainties in the stereo measurements. Of these, the most significant are the uncertainties in of the positions of the specularity in the left and right images. The uncertainty assigned to both of these measurements is ϵ as specified on on page 162. Consequently the uncertainty assigned to the relative disparity of a specularity is approximately one pixel—except for an elongated blob when the motion in the direction of elongation is assigned an uncertainty of roughly 5% of the blob's length. What really matters is the uncertainty of the relative disparity along the principal directions of curvature (as projected onto the image):

- When the relative disparity significantly exceeds ϵ along both principal directions then strong constraints on curvature are obtained. This is the case for **figures 7–19–7–22.**

- When the relative disparity is insignificant in either direction then a large uncertainty is obtained along that direction.

  - This is apparent for the following specularities:
  figure 7–25 Nos :6,10, and to a lesser extent for **figure 7–25** Nos 1,3,4,8,11.

  - It also occurs for specularities that are elongated along one of the principal directions. E.g: **figure 7–23**, **figure 7–24**, and **figure 7–30.**

- When the relative disparity is insignificant in both directions both principal radii are highly uncertainty. E.g: **figure 7–25**:Nos 9,10.

When the source is circular, slant is zero and a circular specularity has been detected the monocular inference $|r_1| \approx |r_2|$ is made. Specularity No 5 on the sphere in **7–26** provides an example. It is also instructive to apply the local spherical interpretation. Note that:

- The most precise estimates or spherical radius are obtained at small slants, e.g: figure 7–26 specularity No 5 the radius $= 0.6 \pm 0.7$ m, c.f. the actual value: 0.5 m.

- At large slants the shape of the specularities distort so that no monocular inference can be made.

- Despite the fact that the relative disparity of specularity No 5 in figure 7–26 is entirely in the horizontal direction, the local spherical interpretation still functions. This is a consequence of its robust, least-squares formulation—equation (5.28).

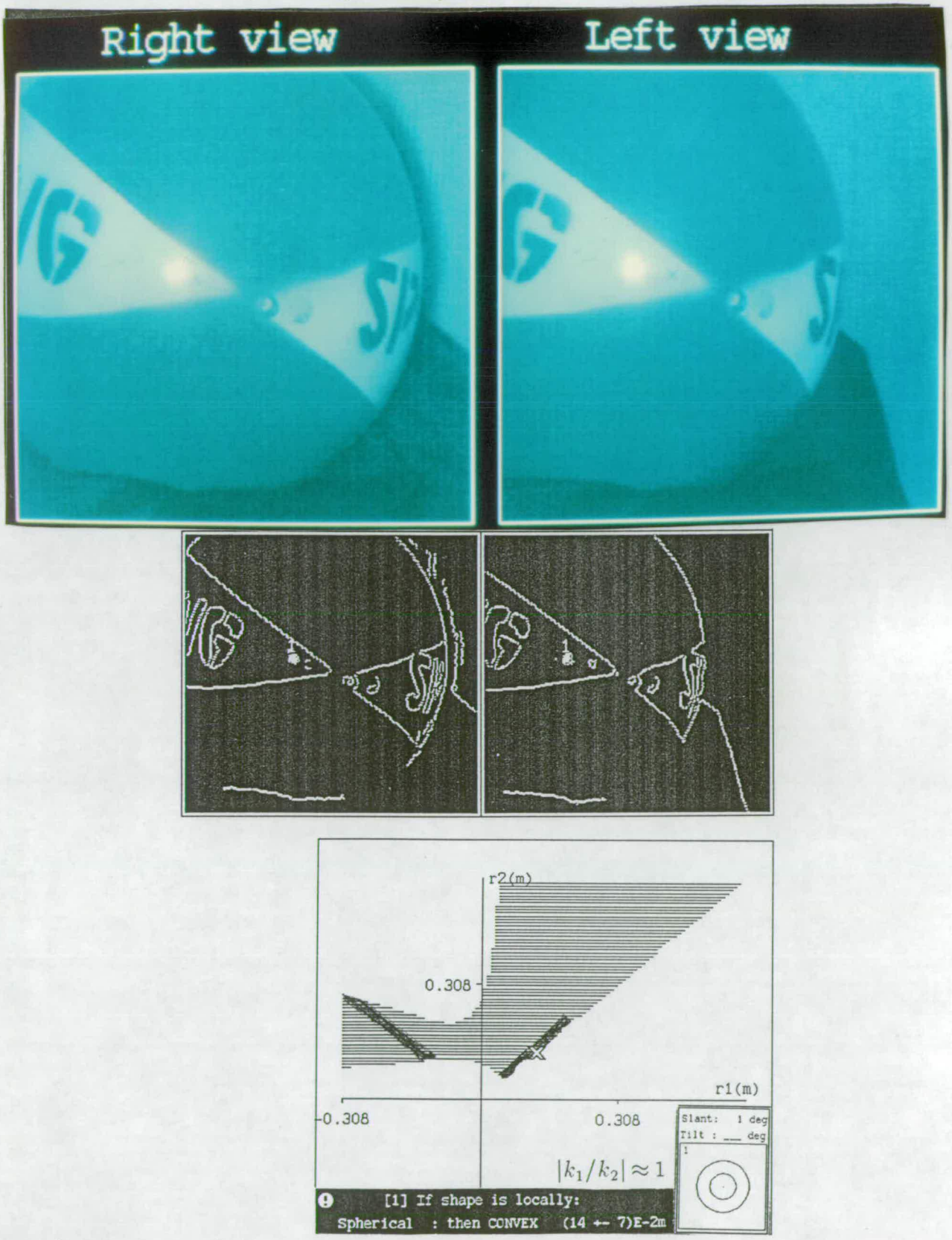


Figure 7–27: Top: *Real stereo-pair of a beach ball—arranged for binocular fusion.* Centre: *Edge map with superimposed, labelled specularities.* Bottom: *Stereo analysis alone restricts principal curvatures to the grey regions of the graph. Monocular evidence correctly indicates principal curvatures of similar magnitude. The locally spherical interpretation then restricts the curvatures to the black regions. Actual radii of curvature (both $\sim$ 12 cm) are indicated by a white cross.*

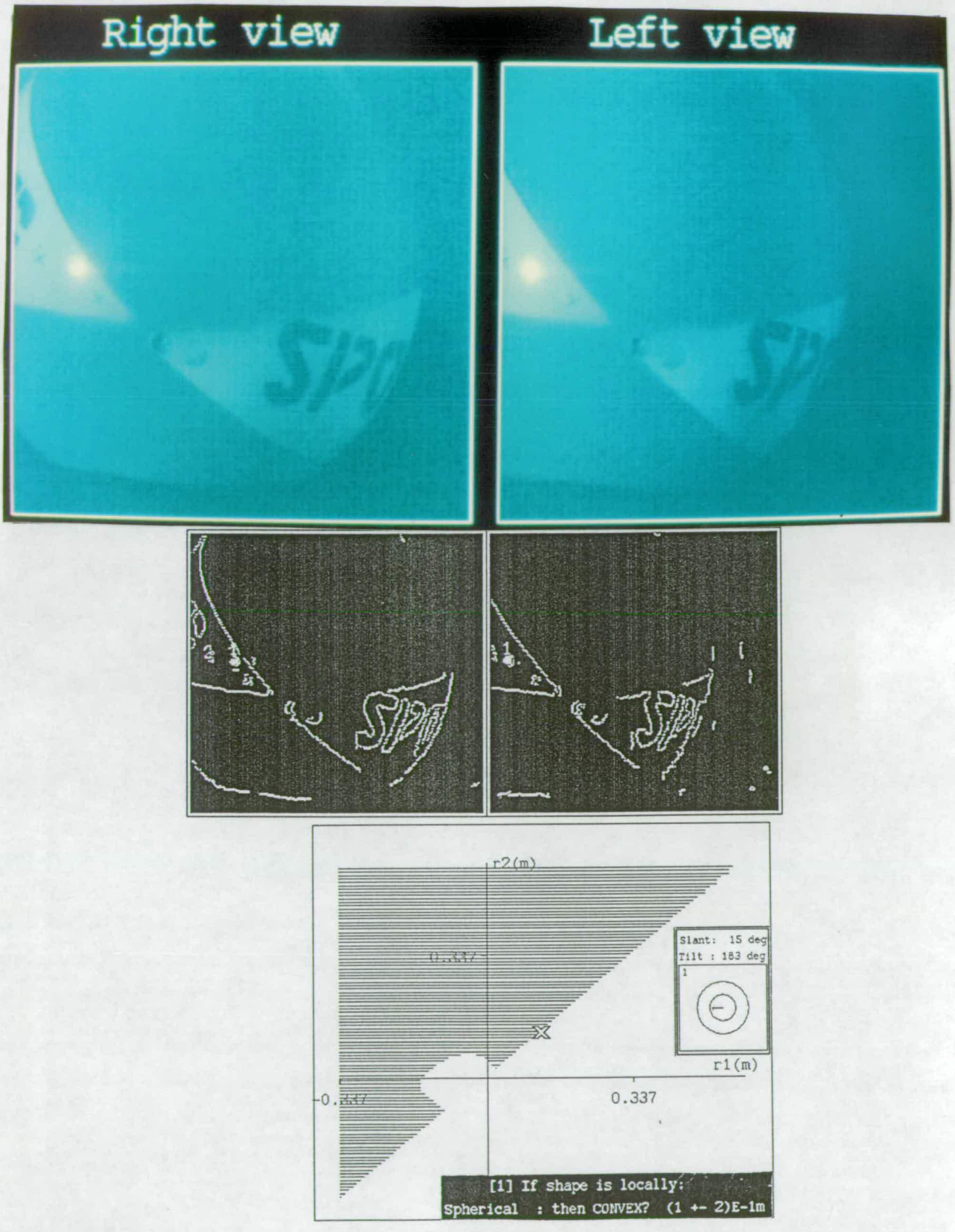


Figure 7–28: Top: *Real stereo-pair of a beach ball—arranged for binocular fusion.* **Centre:** *Edge map with superimposed, labelled specularities.* **Bottom:** *The system restricts principal curvatures to the grey regions of the graph by stereo analysis alone. Monocular analysis fails. Actual radii of curvature (both* ~ 12 *cm) are indicated by a white cross.*

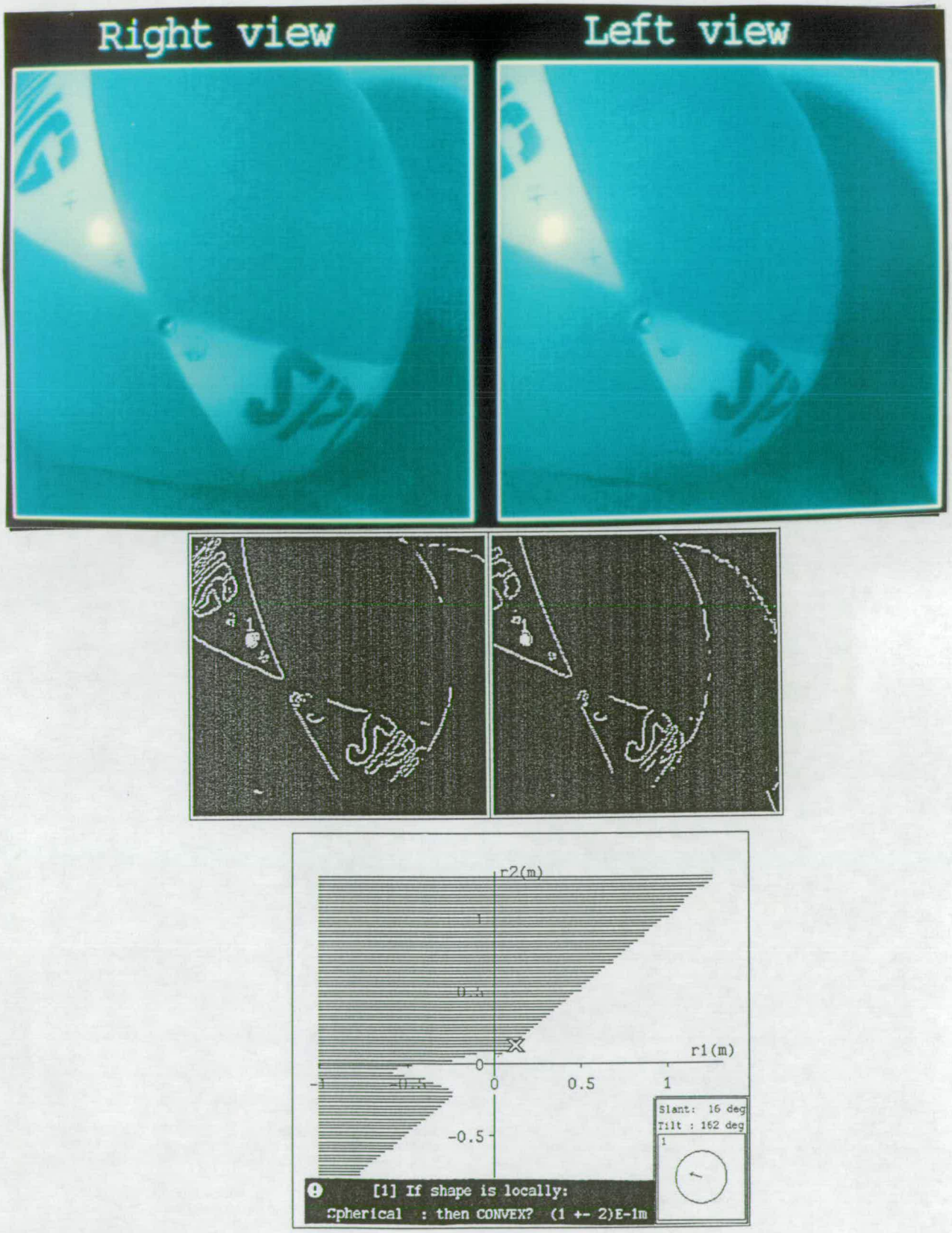


Figure 7–29: Top: *Real stereo-pair of a beach ball—arranged for binocular fusion.* **Centre:** *Edge map with superimposed, labelled specularities.* **Bottom:** *The system restricts principal curvatures to the grey regions of the graph by stereo analysis alone. Monocular analysis fails. Actual radii of curvature (both* $\sim$ *12 cm) are indicated by a white cross.*

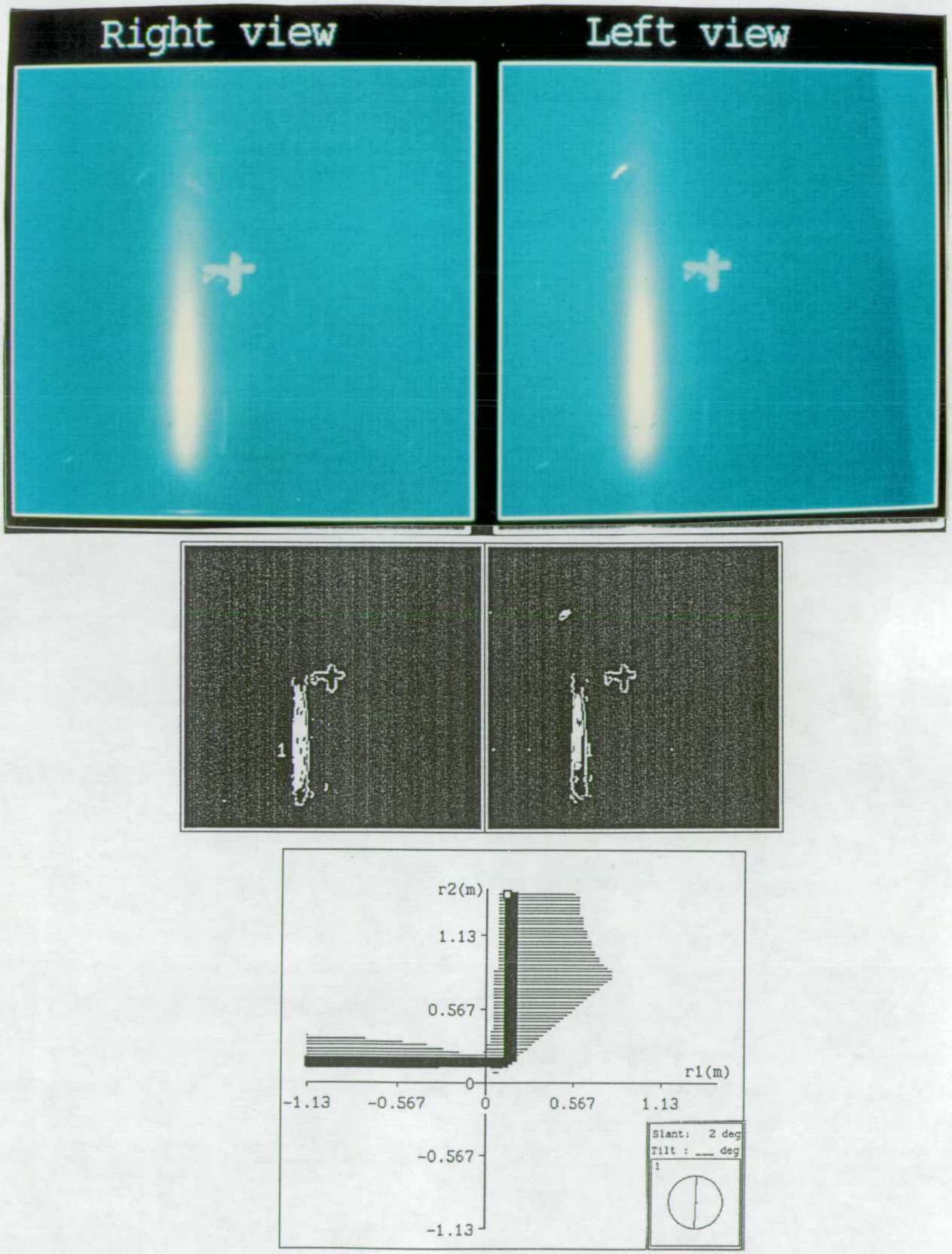


Figure 7–30: Top: *Real stereo-pair of a metal waste basket—arranged for binocular fusion.* **Centre:** *Edge map with superimposed, labelled specularities.* **Bottom:** *The system restricts principal curvatures to the grey region of the graph by stereo analysis alone and to the black region by combining stereo and monocular analyses. Actual radii of curvature are* $(13.5cm, \pm\infty)$*, i.e. locally cylindrical.*

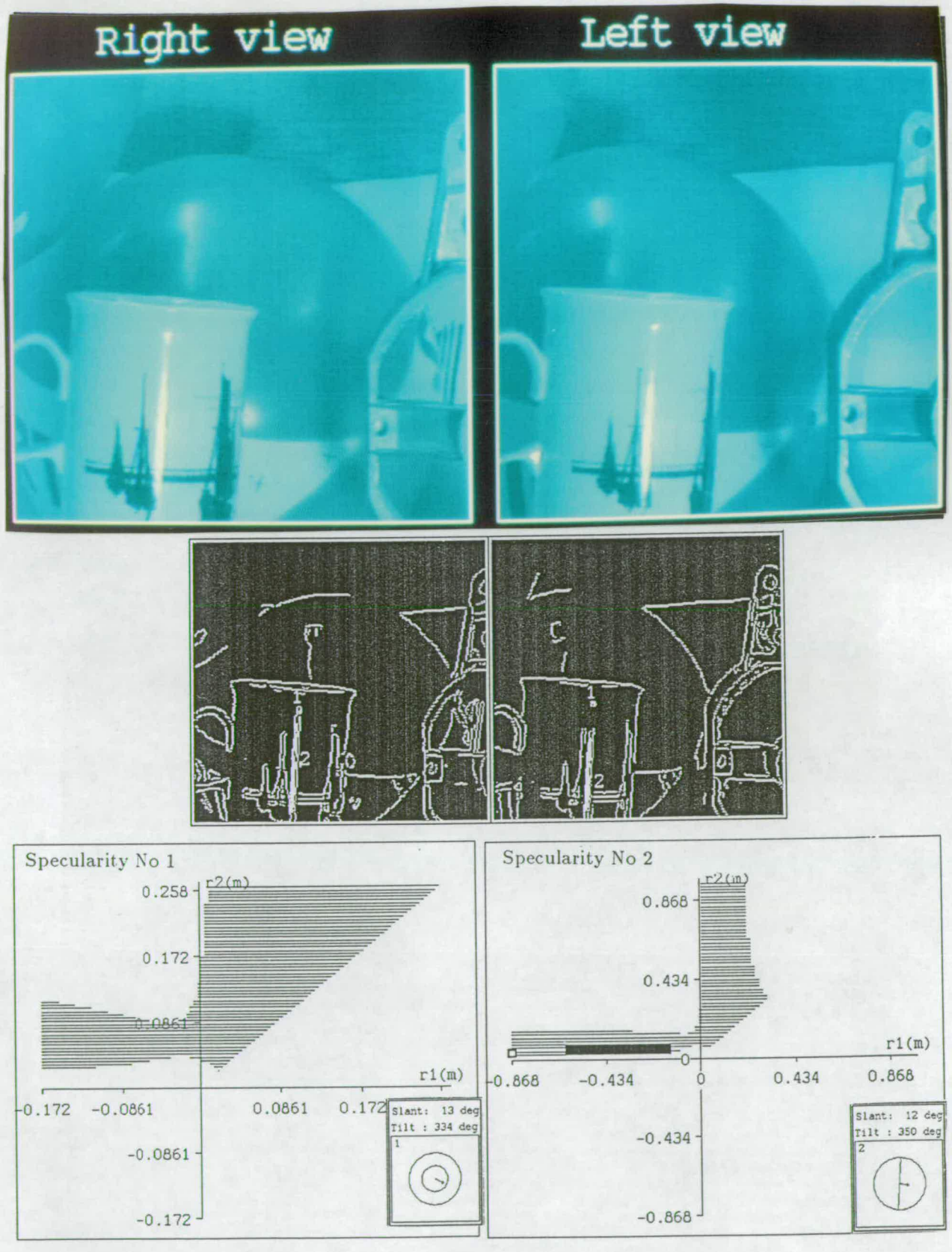


Figure 7–31: Top: *Real stereo-pair of a cluttered scene—arranged for binocular fusion. Mug has a 4 cm radius.* **Centre:** *Edge map with superimposed, labelled specularities.* **Bottom:** *The system restricts principal curvatures to the grey region of the graphs by stereo analysis alone and to the black region by combining stereo and monocular analyses. For specularity No 1 monocular analysis failed.*

Real images

Figures 7–27–7–29 show the curvature inferred from three stereo-pairs of a beach ball. In each figure the specularity lies at a different surface orientation. The depth is estimated at the nearest available point on the border of the ball's white logo-band. The stereo inference made at each orientation is similar to that made from the artificial sphere:

- For figure 7–27 slant is zero and so the (r_1, r_2) graph resembles that at specularity No 5 in figure 7–26.
- For figure 7–28 and figure 7–29 the (r_1, r_2) graph resembles that at specularities No 4 in figure 7–26.

For the real images the relative disparity is more certain than for the artificial ones. Consequently the relative uncertainties reported for the local spherical interpretations are also less. However, the monocular inferences made from the real images are worse than for the artificial ones. The monocular inference $|r_1| \approx |r_2|$ can be made only where slant is zero, i.e. for the specularity in figure 7–27.

Figure 7–30 shows a stereo pair of a convex cylindrical surface. The depth estimate is obtained from a point on the cross painted on the surface close to the elongated specularity. This specularity has a large horizontal relative disparity between the left and right images. Consequently the cylinder's radius 13.5 cm is accurately inferred to be 16 ± 4 cm. An infinite "radius" of curvature in the cylinder's axial direction is also correctly inferred. However, because the uncertainty in the vertical disparity is very large the inference has a very large uncertainty associated with it. This results in the long black line on the (r_1, r_2) graph, resembling that for specularity No 4 on the convex cylinder in figure 7–23).

Figure 7–31 shows a stereo pair of a cluttered scene containing a mug of radius 4 cm. For both specularities lying on the mug, the depth was estimated from near-by points on surface pattern. At the elongated specularity (No 2) one principal radius is correctly inferred to be 5 ± 3 cm. However, it appears

that the mug is not exactly cylindrical—but has large finite concave radius of curvature along its axial direction. A rigorous evaluation of this finding is not possible—without a special measuring device a large and an infinite radii of curvature can be indistinguishable. Note however, some doubt should be placed on this particular inference because the specularity covers such a large distance on the surface. It is likely that the local-surface-patch-assumption is being violated. At the other specularity (No 1) the monocular inference fails. and the only useful information is provided by the constraint graph.

Chapter 8

Conclusions and scope for further work

This thesis has shown that specular reflections can be successfully detected and used to infer surface curvature. The main conclusions that have been drawn are detailed below, with some suggestions for further research. From the review of surface reflection in chapter 2 the following conclusions were drawn:

- Image irradiance is related to local surface orientation through the local reflectance map. The form of the reflectance map depends on the surface material and both the directional and spatial distribution of illumination.

- In general reflectance maps are very complicated functions.

- Perfectly matt surfaces have Lambertian reflectance maps— which are much simpler to handle.

- Theory and experiment strongly suggest that the reflectance maps of real surfaces are roughly Lambertian—except in regions corresponding to specular reflection.

- An upper bound (30) on the dynamic range of Lambertian-like regions of real reflectance maps can be established—and later exploited to detect specularities. This is achieved by considering the two independent factors:

realistic distributions of illumination which account for a factor of 3, and realistic variations in (effective) albedo which account for a factor of 10.

From the specularity detector developed—in chapter 4—the following conclusions were drawn:

- Of the three tests for specularity—based on the Lambertian constraints—two proved usefully reliable: the retinex-based and cylinder tests. The third: the local contrast test is only marginally useful.

- A useful line of work would be the generalisation of the cylinder test, so that its initial assumption—that the surface is approximately cylindrical—can be relaxed.

- The retinex-based test could be made more secure by developing a retinex process which operates sensibly in the rare adverse circumstances mentioned in chapter 4.

- The results show that the scheme used to combine and propagate the evidence of the three tests, delivers useful descriptions of the prominent specularities in each image.

- The current method of propagation might be made more robust by developing a method based on function-fitting.

- Methods might also be developed to provide more comprehensive descriptions of the shapes of specular blobs that are detected.

From the work on shape-from-specularity developed and described—in chapters 6 and 7—the following conclusions can be drawn:

- Both monocular and stereoscopic cues present at specularities can be used to partially infer the local curvature.

- The new formulation of the stereoscopic analysis clarifies aspects of Blake's original (linearised approximate) scheme.

- By consistent treatment of the uncertainties occurring in the stereoscopic and monocular cues, the uncertainity of the inferred curvature can be assigned. It would be interesting to see how quantitative error bounds would change if a more principled error combination method [33,93] was used, in place of the current gaussian based system.

- Stereoscopic and monocular analysis of specularities are complementary and can be combined to improve the quality of the inferred curvature. Often local curvature can be inferred in full. Results with assigned error bounds were obtained by applying the combined analysis to a variety of test images.

- The incorporation of the (fourfold ambiguous) monocular analysis for a circular source into the inference scheme would be a useful extension.

Some interesting questions remain:

- Is it practical to use vertical disparity as an additional cue for detecting specularities? Indeed, is it used by human vision?

- Does the mathematical framework for machine inference of surface curvature have its parallel in human vision? For example, is the predicted fourfold ambiguity actually perceived?

- How can the inferences obtained from specularities be integrated with other sources of visual information? Is the scheme to do this outlined in chapter 1 feasible?

- Does the shape inference scheme easily extend to cope with extended viewer motion? Stereoscopic analysis is after all a special case of motion analysis.

Bibliography

[1] AIVRU, *GDB Release 1.0, User Documentation* AIVRU-014, AIVRU, University of Sheffield, (1984).

[2] Aloimonos, J. *Computing intrinsic images*, Ph.D thesis, University of Rochester, USA, (1986).

[3] Aloimonos, J., Weiss, I. and Bandyopadhyay, A. Active Vision, *Proc. First Int. Conf. on Computer Vision*, London, (1987), 35-54.

[4] American Institute of Physics *American Institute of Physics Handbook*, 3rd Ed, McGraw Hill, (1972).

[5] Ayache, N. and Lustman, F., Trinocular Stereovision: Recent Results, *Proc. IJCAI-87*, (1987), 826-828.

[6] Babu, M.D.R., Lee, C-H. and Rosenfeld, A. Determining Plane Orientation From Specular Reflectance, *Pattern Recognition*, 18, 1, (1985), 53-62.

[7] Baker, H.H. Depth from edge and intensity based stereo, *Proc IJCAI-81*, (1981), 583-588.

[8] Baker, H.H. and Binford, T.O. A System for Automated Stereo Mapping, *Symp. ISPRS Commission II*, Ottawa, Canada (1982).

[9] Ballard, D.H. and Brown, C.M. *Computer Vision*, Prentice Hall, (1982).

[10] Barnard, S.T. and Thompson, W.B. Disparity Analysis of Images, *IEEE, Trans. PAMI*, 2, 4, (1980), 333-340.

[11] Beckmann, P. and Spizzichino, A. *The Scattering of Electromagnetic Waves from Rough Surfaces*, MacMillan, New York, (1963).

[12] Billmeyer, F.W., Lewis, D.L. and Davidson, J.G. Goniophotometry of Pressed Barium Sulphate, *Color Engineering*, May/June, (1971) 31-36.

[13] Blake, A. *Inferring Surface Shape by Specular Stereo*, Internal Report, Dept. of Computer Science, University of Edinburgh, CSR-179-84, (1984).

[14] Blake, A. Boundary Conditions for the Lightness Computation in Mondrian World, *Computer vision, graphics, and image processing*, 32, (1985), 314-327.

[15] Blake, A. Specular Stereo, *Proc. IJCAI-85*, 2, (1985), 973-976.

[16] Blake, A. *On the Geometric Information Obtainable from Simultaneous Observation of Stereo Contour and Shading*, Internal Report, Dept. of Computer Science, University of Edinburgh, CSR-205-86, (1986).

[17] Blake, A. and Brelstaff, G.J. Geometry from Specularities, Forthcoming paper: submitted to *ICCV-88*, Florida, December, (1988).

[18] Blake, A. and Mayhew, J.E.W. *Alvey 2.5D Sketch Project*, AIVRU report 009, University of Sheffield, (1986).

[19] Blake, A. and Zisserman, A. *Visual Reconstruction*, M.I.T. Press, Cambridge, Mass, (1987).

[20] Blinn, J.F. Models of light reflection for computer synthesized pictures, *ACM SIGGRAPH*, 11, 2, (1977), 192-198.

[21] Brady, J.M., and Yuille, A. An extremum principle for shape from contour, *IEEE Trans. PAMI*, 6, (1984), 288-301.

[22] Brelstaff, G.J. and Blake, A. Computing Lightness, *Pattern Recognition Letters*, 5, 129-138, (1987). (Included as Appendix K.)

[23] Buchanan, C.S.B. *Determining Surface Orientation from Specular Highlights*, RCBV Tech. Report, RCBV-TR-87-19, University of Toronto, Aug (1987).

[24] Canny, J.F. *Finding Edges and Lines in Images*, Masters Thesis, AI-TR 720, (1983), A.I. Lab., M.I.T.

[25] Catmull, E. A tutorial on compensation tables, *ACM Trans. on Computer Graphics*, 4, (1979), 1-7.

[26] Chylek, P. Light Scattering by small particles in an absorbing medium, *J. Opt. Soc. Am.* 67, 4, (1977), 561-563. AI-TR 720, (1983), A.I. Lab., M.I.T.

[27] Chylek, P. and Wiscombe, W.J. Mie Scattering between any two angles, *J. Opt. Soc. Am.* 67, 4, (1977), 572-573. AI-TR 720, (1983), A.I. Lab., M.I.T.

[28] Coblentz, W.W. The diffuse reflecting powers of various substances, *Bulletin of the Bureau of Standards*, 9, (1912), 283-325.

[29] Cohen, M.F. and Greenberg, D.P. The Hemi-cube A Radiosity Solution For Complex Environments, *ACM SIGGRAPH*, 19, 3, (1985), 31-40.

[30] Coleman, E.N.Jr., and Jain, R. Obtaining 3-Dimensional Shape of Textured and Specular Surfaces Using Four-Source Photometry, *Computer graphics and image processing*, 18, (1982), 309-328.

[31] Cook, R.L. and Torrance, K.E. A Reflectance Model for Computer Graphics, *ACM Transactions on Graphics*, 1, 1, (1982), 7-24.

[32] do Carmo, M.P. *Differential geometry of curves and surfaces*, Prentice Hall, Englewood Cliffs, USA (1976).

[33] Durrant-Whyte, H.F. *Consistent Integration and Propagation of Disparate Sensor Observations*, Thesis, University of Pennsylvania, (1985).

[34] Elson, J.M. Light scattering from surfaces with single dielectric overlayer, *J. Opt. Soc. Am.* 66, 7, (1976), 682-684.

[35] Elson, J.M and Bennet, J.M. Relation between the angular dependences of scattering and the statistical properties of optical surfaces, *J. Opt. Soc. Am.* 69, 1, (1979), 31-47.

[36] Faugeras, O.D. and Herbert, M. *The representation, recognition and positioning of 3-D shapes from range data*, Research report IA336, (1985), INRIA, France.

[37] Forbus, K. *Light Source Effects* A.I. Memo 422, (1977), A.I. Lab., M.I.T.

[38] Forsythe W.E. *Smithsonian Physical Tables*, 9th ed., (1954), 549-559.

[39] Gershon, R., Jepson, A.D. and Tsotsos, J.K. Highlight Identification Using Chromatic Information, Proc. IJCAI-87, Milan, 2, (1987), 752-754.

[40] Gershon, R., Jepson, A.D. and Tsotsos, J.K. The Use of Color in Highlight Identification, *Proc. First Int. Conf. on Computer Vision*, London, (1987), 161-170.

[41] Gibson J.J. *The perception of the Visual World*, Houghton Miffin, Boston, (1950).

[42] Goral, C.M., Torrance, K.E., Greenberg, D.P. and Battaile, B. Modeling the diffuse interaction of light between diffuse surfaces, *Proc. ACM Computer Graphics*, (1984), 213-222.

[43] Gordon J.I. Optical Properties of Objects and Backgrounds, *Applied Optics*, 3, 5, (1964), 556-562.

[44] Gordon J.I. and Church, P.V. Sky Luminances and Directional Luminous Reflectances of Objects and Backgrounds for a Moderately High Sun, *Applied Optics*, 5, 5, (1966), 793-801.

[45] Gordon J.I. and Boileau, A.R. Atmospheric Properties and Reflectances of Ocean Water and other Surfaces for a Low Sun, *Applied Optics*, 5, 5, (1966), 803-813.

[46] Gordon J.I. and Church, P.V. Overcast Sky Luminances and Directional Luminous Reflectances of Objects and Backgrounds under Overcast Skies, *Applied Optics*, 5, 6, (1966), 919-923.

[47] Grimson, W.E.L. *From Images to Surfaces* M.I.T. Press, Cambridge Mass. U.S.A. (1981).

[48] Grimson, W.E.L. *Binocular Shading and Visual Surface Reconstruction* M.I.T. A.I.Lab. Memo No 697, August (1982).

[49] Grimson, W.E.L. and Lorenzo-Perez, T. Model-based recognition and localisation from sparse range or tactile data, *Int. J. Robotics Res.*, 3, 3, (1984), 3-35.

[50] Gubareff G.G., Janssen, J.E. and Torbourg, R.H. *Thermal Radiation Properties Survey: A Review of the Literature*, Honeywell Research Center, Minneapolis, USA, (1960).

[51] Haralick, R.M. Ridges and Valleys on Digital Images, *Computer vision, graphics and image processing*, 22, (1983), 28-38.

[52] Harrison, V.G.W. *Definition and Measurement of Gloss*, Monograph, The Printing and Allied Trades Research Assoc. (1945).

[53] Healey, G., and Binford, T.O. Local shape from specularity, *Proc. First Int. Conf. on Computer Vision*, London, (1987), 151-160

[54] Hecht, E., and Zajac, A. *Optics*, Addison-Wesley, Reading Mass., (1980).

[55] Hodgman, C.D. *Handbook of Chemistry of Physics*, 32nd Ed, (1950-51), 2443-2445.

[56] Hopkinson, R.G., Petherbridge, P. and Longmore, J. *Daylighting*, William Heinemann, London, (1966).

[57] Horn, B.K.P. *Shape from Shading*, MIT Project MAC Internal Report TR-79, and MIT AI Memo 232, (1970).

[58] Horn, B.K.P. Determining Lightness from an Image, *Computer Graphics and Image Processing*, 3 (1974) 277-299.

[59] Horn, B.K.P. *Robot Vision*, McGraw-Hill, 1986.

[60] Horn, B.K.P. and Brooks, M.J. The Variational Approach to Shape from Shading, *Computer Vision, Graphics and Image Processing*, 33 (1986) 174-208.

[61] Horn, B.K.P. and Sjoberg, R.W. Calculating the reflectance map, *Applied Optics*, 18, 11, (1979), 1770-1779.

[62] Hottel, H.C. and Sarofim A.F. *Radiative Transfer*, McGraw-Hill, 1967.

[63] Hunter, R.S. Methods of determining gloss, *J. Research Nat. Bur. Std.* 18, (1937) 19-39.

[64] Hurlbert, A. Formal connections between lightness algorithms, *J. Opt. Soc. Am. A*, 3, 10, (1986), 1684-1693.

[65] Ikeuchi, K. Shape from Regular Patterns, *Artificial Intelligence*, 22, (1984), 49-75.

[66] Ikeuchi, K. Determining Surface Orientation of Specular Surfaces by Using the Photometric Stereo Method, *IEEE Trans. PAMI*, 3, 6, (1981), 661-669.

[67] Ikeuchi, K. and Horn, B.K.P. Numerical Shape from Shading and Occluding Boundaries, *Artificial Intelligence*, 17, (1981), 141-184.

[68] Kajiya, J.T. Anisotropic Reflection Models, *ACM SIGGRAPH*, 19, 3, (1985), 15-21.

[69] Kanade, T. Determining the shape of an object from a single view, *Artificial Intelligence*, 17, (1981), 409-460.

[70] Kanatani, K. Detection of Surface Orientation and Motion from Texture by Stereological Technique, *Artificial Intelligence*, 23, (1984), 213-237.

[71] Kender, J.R. *Shape from Texture*, Ph.D. thesis, Dept of Comp. Science, CMU, Pittsburgh, PA, (1980).

[72] Klinker, G.J., Shafer, S,A., and Kanade, T. Using a color reflection model to separate highlights from object color, *Proc. First Int. Conf. on Computer Vision*, London, (1987) 145-150.

[73] Koenderink, J.J. and van Doorn, A,J. Photometric invariants related to solid shape, *Optica Acta*, 27, 7, (1980), 981-996.

[74] Krinov, E.L. *Spectral reflectance properties of natural formations*, National Research Council of Canada, Technical Translation, TT-439, (A translation from the Russian by G. Belkov, of the introduction and part I of chapter I of *Spektral'nain otrazhatel'naia sposobnost' prirodnykh obrazovanii*, Laboratoriia Aerometodov, Akad. Nauk SSSR. Moscow, (1947), 271 p.).

[75] Krotov, E.P. Visual hyperacuity: representation and computation of very high precision position information, *Computer Vision, Graphics and Image Processing*, 33, (1986), 363-370.

[76] Land, E.H. and McCann, J.J. Lightness and Retinex Theory, *J. Opt. Soc. Am.* 61, 1, (1971), 1-11.

[77] Land, E.H. *Personal communication.*

[78] Land, E.H. Recent advances in retinex theory and some implications for cortical computation: Color vision and natural images, *Proc. Natl. Acad. Sci. U.S.A.* 80, (1983), 5163-5169.

[79] Lipschutz, M.M. *Differential geometry*, McGraw Hill, New York, (1969).

[80] Lowe, D.G. and Binford, T.O. The Interpretation of Three-Dimensional Structure From Image Curves, *Proc. IJCAI-81*, (1981) 613-618.

[81] Luckiesh, J. *J. Franklin Inst*, 184, 79 (1917) - summarised in [38].

[82] Marinos, C.I. *Shape from Texture and Shape from Shading*, Robotics Research Group, Dept. of Engineering, Univ. of Oxford, (1988).

[83] Marr, D. *Vision*, Freeman, San Francisco, (1982).

[84] Mudgett, P.S. and Richards L.W. Multiple Scattering Calculations for Technology, *Applied Optics*, 10, 7, (1971), 1485-1502.

[85] Nicodemus, F.E., Richmond, J.C., Ginsberg, I.W., Hsia, J.J. and Limperist, T. *Geometrical Considerations and Nomenclature for Reflectance* Monograph 160, NBS, Washington D.C., (1977).

[86] Nishita, T. and Nakamae, E. Continuous Tone Representation of Three-Dimensional Objects Taking Account of Shadows and Interreflection, *ACM SIGGRAPH*, 19, 3, (1985), 23-30.

[87] Nishita, T. and Nakamae, E. Continuous Tone Representation of Three-Dimensional Objects Illuminated by Skylight, *ACM SIGGRAPH*, 20, 4, (1986), 125-132.

[88] Ohta, Y. and Kanade, T. Stereo by Intra- and Inter Scanline Search Using Dynamic Programming, *IEEE Trans. PAMI*, 7, 2, (1985), 139-154.

[89] Pentland, A. A new sense for depth of field, *Proc. IJCAI-85*, 2, (1985), 988-994.

[90] Phong, B.T. Illumination for computer generated pictures, *ACM Comms.* 18, 6, (1975), 311-317.

[91] Pollard, S.B. Mayhew, J.E.W. and Frisby, J.P. PMF: a stereo correspondence algorithm using the disparity gradient limit, *Perception*, 14, (1985), 449-470.

[92] Pollard, S.B., Porrill, J., Mayhew, J.E.W. and Frisby, J.P. Disparity gradient Lipschitz continuity: computing binocular correspondence, *3rd Int. Symposium on Robotics Research*, (1985), 19-26.

[93] Porrill, J., Pollard, S.B. and Mayhew, J.E.W. The Optimal Combination of Multiple Sensors including Stereo Vision, *Image and Vision Computing*, 5, 2, (1987), 175-180.

[94] Purdue University *Thermophysical Properties of Matter*, Vols 7, 8 and 9, Plenum, New York, (1970).

[95] Quarendon, P. *WINSOM User's Guide*, IBM UKSC 123, (1984).

[96] Sanderson, A.C., Weiss, L.E. and Nayar, S.K. Structured Highlight Inspection of Specular Surfaces, *IEEE Trans. PAMI*, 10, 1, (1988), 44-55.

[97] Sato, K. and Inokuchi, S. Range-Imaging System Utilizing Nematic Liquid Crystal Mask, *Proc, IEEE Int, Conf. on Computer Vision*, (1987), 657-661.

[98] Shafer, S.A. *Optical Phenomena In Computer Vision*, Technical Report, TR 135, (1984), Computer Science Dept., University of Rochester.

[99] Smith, G.D. *Numerical solutions of partial differential equations: finite difference methods*, Oxford University Press (1978). Technical Report, TR 135, (1984), Computer Science Dept., University of Rochester.

[100] Squires, G.L. *Practical Physics*, European Physics Series, McGraw-Hill, London (1968).

[101] Strat, T.M. *A Numerical Method for Shape from Shading from a Single Image*, S.M. Thesis, Dept Elec. Eng. and Comp. Sci. MIT, (1979).

[102] Terzopoulos, D. Multilevel computational processes for visual surface reconstruction,, *Computer Vision Graphics and Image Processing*, 24, (1985), 52-96.

[103] Thrift, P. and Lee, C-H. Using Highlights to Constrain Object Size and Location, *IEEE Trans. Sys, Man and Cybernetics*, 13, 3, (1983), 426-431.

[104] Torrance, K.E. and Sparrow, E.M. Theory of Off-Specular Reflection From Roughened Surfaces, *J. Opt. Soc. Am.* 57, 9, (1967), 1105-1114.

[105] Trowbridge, T.S. and Reitz, K.P. Average irregularity representation of a rough surface for ray reflection, *J. Opt. Soc. Am.* 65, 5, (1975), 531-536.

[106] Ullman, S. On Visual Detection of Light Sources, *Biol. Cybernetics*, 21, (1976), 205-212.

[107] Vincent, R.K. and Hunt, G.R. Infrared Reflectance from Mat Surfaces, *Applied Optics*, 7, 1, (1968), 53-59.

[108] Whitted, T. An improved illumination model for shaded display, *ACM Comms.* 23, 6, (1980), 343-349.

[109] Witkin, A.P. Recovering surface shape and orientation from texture *Artificial Intelligence*, 17, (1981), 17-45.

[110] Witkin, A.P. Intensity-Based Edge Classification, *Proc. A.A.A.I.* (1982), 36-41.

[111] Woodham, R.J. Photometric method for determining surface orientation from multiple images, *Optical Engineering*, 19, 1, (1980), 139-144.

[112] Woodham, R.J. Analysing Curved Surfaces, *Artificial Intelligence*, 17, (1981), 117-140.

Appendix A

Geometrical and Radiometrical concepts

A.1 Introduction

In 1977 the concepts of the geometrical specification of surface reflectance were rationalised and made into a single unified system [85]. This specification is in terms of both incident and reflected beam geometries. The standard nomenclature and symbols introduced by this system are adhered to within this thesis, wherever they prove convenient.

This appendix provides their precise definitions—as promised in section 2.3.2. The concepts are defined in sequence, with the later definitions depending on those that precede them. The most important concepts with respect to this thesis are the bidirectional reflectance distribution function, the incident radiance distribution, the reflectance map and the image irradiance equation.

A.1.1 Radiometrical concepts

The *polar coordinate system* shown in figure A–1 is used to define directions within the hemisphere above a surface point, O. The pair of angles (θ, ϕ)

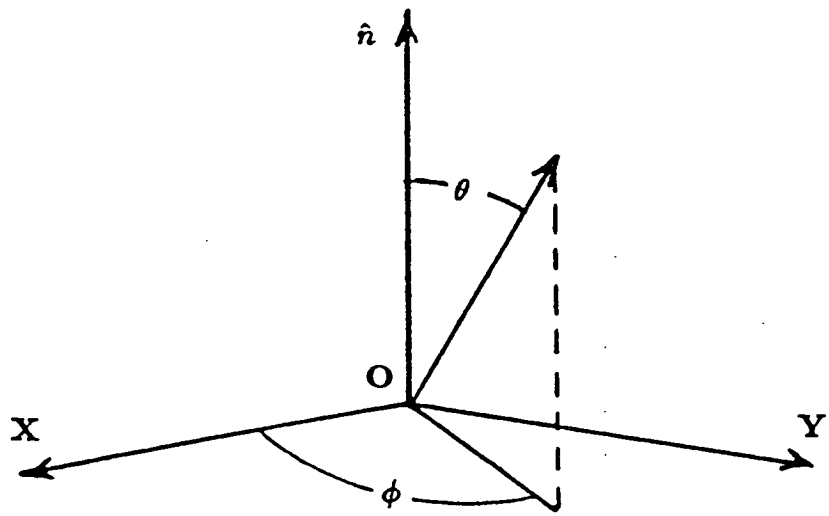


Figure A–1: *A polar coordinate system is used to define an arbitrary direction* (θ, ϕ) *with respect to the local surface normal* $\hat{\mathbf{n}}$. θ *is the polar angle and* ϕ *is the angle of azimuth.*

defines an arbitrary direction with respect to the local surface normal $\hat{\mathbf{n}}$. θ is the polar angle and ϕ the angle of azimuth relative to a fixed x-axis.

A beam emanating from point O to form an area A on the hemisphere that lies at unit distance above O occupies a *solid angle*

$$\omega = \int\int_A \sin\theta d\theta d\phi \tag{A.1}$$

as marked in figure A–2. Solid angle is measured in steradians (sr).

The power propagated by a beam of light is called *radiant flux* and denoted by Φ. Radiant flux is measured in watts (W).

Irradiance, denoted E is the density of flux incident upon a surface. It has units Wm^{-2}. *Image irradiance* is the quantity measured by a camera. Image irradiance is commonly referred to as intensity or brightness.

Radiance, denoted L, is defined with respect to an element of surface, and an arbitrary direction in the hemisphere above it. It is the flux propagated

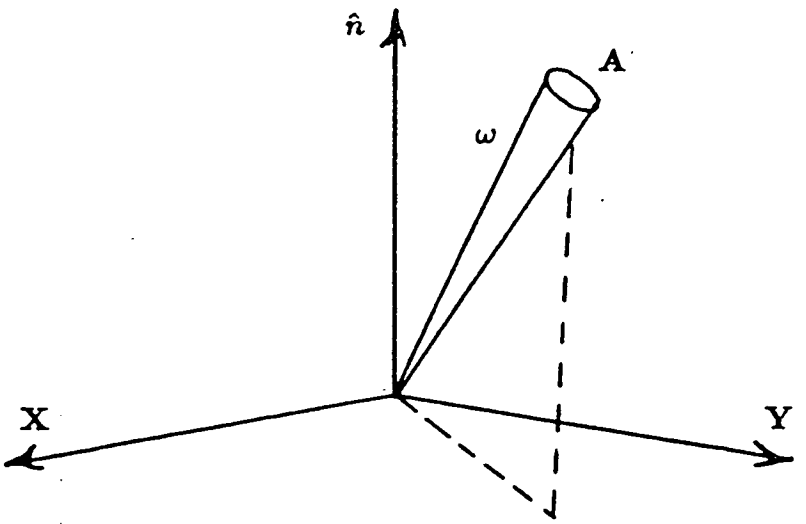


Figure A–2: *The solid angle occupied by a beam is defined in terms of an integral over an element of area A on the hemisphere at unit distance .*

along the direction per unit solid angle per unit foreshortened surface area. So it has units $Wsr^{-1}m^{-2}$. The radiance reflected by a surface L_r is an important quantity: it is proportional to the image irradiance measured by a camera—see section A.1.3.

Directional quantities in radiometry are denoted using the *infinitesimal notation.* This is in keeping with the notion that a ray in any single direction occupies only an infinitesimal solid angle $d\omega$. So for example the reflected radiance along an arbitrary direction is denoted dL_r.

Integration over a solid angle ω involves the infinitesimal notation. For a general quantity X the integration in terms of the polar coordinates is:

$$\int_\omega X d\omega = \int_{-\pi}^{\pi} \int_0^{\pi} X \sin\theta d\theta d\phi. \tag{A.2}$$

A.1.2 The BRDF and reflectance

The bi-directional reflectance distribution function (BRDF), denoted f_r describes, for a given surface the ratio of reflected radiance dL_r along an arbitrary

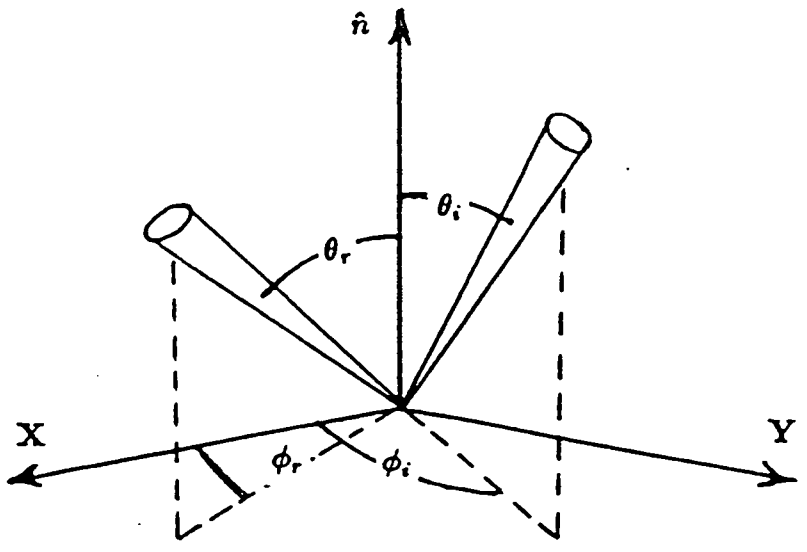


Figure A–3: *The BRDF is defined in terms of incident and reflected directions. These directions are defined using the local polar coordinate system. They are respectively* (θ_i, ϕ_i) *and* (θ_r, ϕ_r).

direction to the amount of incident irradiance dE_i along another arbitrary direction: i.e.

$$f_r(\theta_i, \phi_i; \theta_r, \phi_r) = dL_r(\theta_i, \phi_i; \theta_r, \phi_r; E_i)/dE_i(\theta_i, \phi_i) \qquad \text{(A.3)}$$

where (θ_i, ϕ_i) and (θ_r, ϕ_r) represent the directions of incidence and reflection respectively[1]. Figure A–3 helps to clarify the geometry of this definition.

The definition (A.3) is applicable wherever the interaction of light with the surface material is such that the scattering properties are locally uniform and isotropic. In broad terms this means all surfaces that do not change their reflective properties by merely rotating them. A few unusual materials such as diffraction gratings, iridescent feathers and special minerals (e.g. tiger eye) are not uniformly isotropic. These materials cannot be specified by a BRDF—in

[1]subscripts i and r denote the incident and reflected quantities, respectively.

terms of this thesis diffuse reflections from them might be mistaken for specularities.

The BRDF permits a precise expression for reflectance of a material in terms of arbitrary beam geometries. *Reflectance* is the ratio of reflected to incident flux. It is defined [85] as:

$$\rho(\omega_i, \omega_r; L_i) = \frac{d\Phi_r}{d\Phi_i} = \frac{\int_{\omega_r}\int_{\omega_i} f_r(\theta_i, \phi_i; \theta_r, \phi_r) L_i(\theta_i, \phi_i) \cos\theta_i \cos\theta_r d\omega_i d\omega_r}{\int_{\omega_i} L_i(\theta_i, \phi_i) \cos\theta_i d\omega_i} \quad \text{(A.4)}$$

where $L_i(\theta_i, \phi_i)$ represents the directional distribution of the *incident radiance.* For example it might represent the light from the whole sky. Equation (A.4) makes it clear that the ratio of reflected to incident flux is not a simple constant of the surface material—it does indeed depend on the illumination conditions, r.e. section 2.2.4.

A.1.3 Image irradiance

Horn and Sjoberg [61] derive an expression for image irradiance—the radiometrical quantity measured by a camera. The expression is derived using the radiometrical terms defined above. For a camera with focal length f and aperture diameter d, image irradiance E_P at a small area on the image P, is:

$$E_p = a_c L_r, \quad a_c = \frac{\pi}{4}\left(\frac{d}{f}\right)^2 \cos^4\gamma \quad \text{(A.5)}$$

where L_r is the reflected radiance passing on to area P from the scene, and γ is the angle between the camera's optical axis and the image forming light beam—as shown in figure A–4.

The derivation of (A.5) involves the following assumptions

- that the camera is properly focused.
- that there is no vignetting of the aperture.
- that the lens diameter is small relative to the viewing distance.

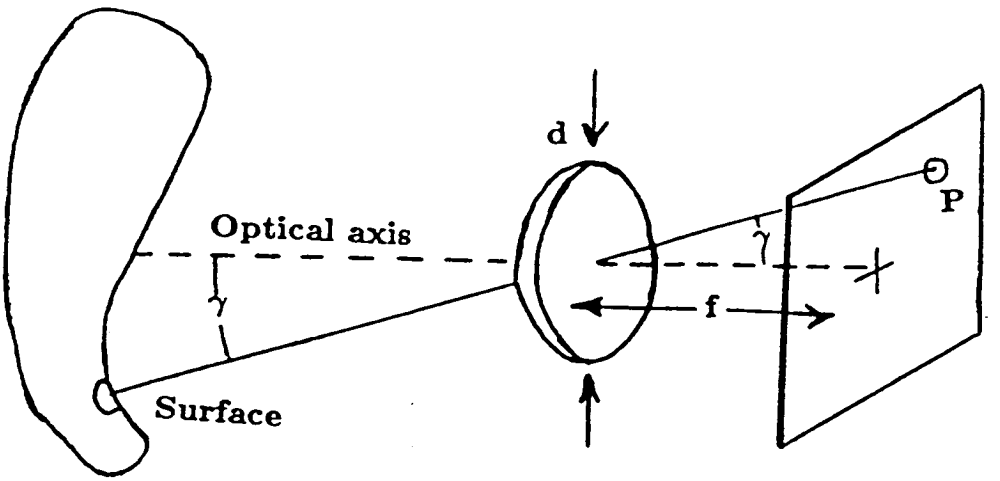


Figure A–4: *Image formation for a simple camera. Light reflected by a surface enters the lens aperture and is focused on to the image plane. The lens has diameter d, and focal length f. γ is the angle between the image forming beam and the optical axis.*

In most practical circumstances equation (A.5) provides a good approximation to image irradiance. Note than the camera's sensitivity is constant for a given image location, P, but is non-uniform over the image. This non-uniformity is due to the $\cos^4 \gamma$ off-axis term. Table A–1 shows that for fields of view less than twenty degrees the sensitivity varies by less than ten percent. For larger fields of view the camera should be calibrated to nullify the effects of these off-axis variations.

γ^o	$\cos^4 \gamma$
0	1.00
5	0.98
10	0.94
15	0.87
20	0.78
30	0.56

Table A–1: *The $\cos^4 \gamma$-attenuation for small off-axis angles γ.*

A.1.4 Reflected radiance

Equation (A.5) shows that image irradiance, E_P, is proportional to reflected radiance, L_r. The expression for L_r in terms of the BRDF [85] is:

$$L_r = \int_{\omega_i} f_r L_i \cos\theta_i d\omega_i, \tag{A.6}$$

which when expressed in terms of the local polar coordinates (via (A.2)) becomes:

$$L_r(\theta_r, \phi_r) = \int_{-\pi}^{\pi}\int_{0}^{\pi} f_r(\theta_i, \phi_i; \theta_r, \phi_r) L_i(\theta_i, \phi_i) \cos\theta_i \sin\theta_i d\theta_i d\phi_i. \tag{A.7}$$

This expression predicts the image irradiance measured by a camera along direction of reflection (θ_r, ϕ_r), from the BRDF, f_r and the incident radiance distribution L_i. The BRDF contains only information about the reflectance properties of the material, while L_i describes only the illumination falling on the surface. Thus equations (A.5) and (A.7) provide the complete geometrical model of the image formation of a surface point imaged by a camera. The advantage of this model is that it leaves nothing unspecified—unlike the ray optics model of section 2.2.

A.1.5 A perfectly diffuse (Lambertian) surface

The simple special case of a perfectly diffuse, or Lambertian surface is of particular importance in this thesis. It is used to develop an approach for detecting specularities.

In this special case all the light is scattered diffusely, resulting in an isotropic reflected radiance distribution—regardless of the form of the illumination. In terms of equation (A.7) this means that L_r is a constant over all directions (θ_r, ϕ_r). This corresponds again to the situation shown by figure 2–3. It can be shown [85] from equation (A.7) that this case can only occur if f_r is a constant (de-

noted $f_{r,d}$). For a lossless[2] Lambertian surface it transpires that

$$f_{r,d}(\theta_i, \phi_i; \theta_r, \phi_r) = 1/\pi, \tag{A.8}$$

and thus (via (A.6)) the reflected radiance is

$$L_{r,d} = \frac{1}{\pi} \int_{\omega_i} L_i \cos\theta_i d\omega_i. \tag{A.9}$$

This expression is a generalisation of the form of Lambert's cosine law, previously discussed in section 2.2.3—for the single incident ray $L_i(\theta_i, \phi_i)$ was only non-zero along one direction. Thus L_r was simply proportional to $\cos\theta_i$. And this in turn was proportional to the image irradiance (via (A.5)).

For a surface that is not lossless but still an isotropic scatterer, the reflected radiance is

$$L_{r,d} = \frac{\rho_d}{\pi} \int_{\omega_i} L_i \cos\theta_i d\omega_i, \tag{A.10}$$

where albedo ρ_d is the ratio of the reflected flux to that which is incident. For diffuse reflectors albedo is equivalent to reflectance. For other surfaces the more complicated expression (A.4) is needed to define reflectance. Note: albedo is defined only for isotropic scatterers.

A.1.6 A perfectly specular surface

The incident flux along any direction is reflected by a perfectly specular surface along the corresponding direction of mirror reflection. In terms of radiometric directions, this means that if (θ_r, ϕ_r) is the direction of reflection then the incident direction was $(\theta_i, \phi_i) = (\theta_r, \phi_r + \pi)$.

For a perfect lossless specular surface the reflected radiance is simply

$$L_{r,is}(\theta_r, \phi_r) = L_i(\theta_i, \phi_i + \pi) \tag{A.11}$$

---

[2]Lossless surfaces reflect all the incident light, absorbing none.

and it can be shown [85] that the BRDF must be expressed as a product of Dirac delta functions:

$$f_{r,s}(\theta_i, \phi_i; \theta_r, \phi_r) = \delta(\theta_i - \theta_r)\delta(\phi_i - \phi_r + \pi)/(\sin\theta_i \cos\theta_i). \quad \text{(A.12)}$$

Real surfaces, however, are never totally lossless. For an optically smooth surface this means that the expression (A.11) must be modified to:

$$L_{r,s} = F(\theta_i)L_i(\theta_i, \phi_i + \pi), \quad \text{(A.13)}$$

where $F(\theta_i)$ is the Fresnel attenuation term, introduced in section 2.2.2. The corresponding expression for the BRDF is

$$f_{r,s}(\theta_i, \phi_i; \theta_r, \phi_r) = F(\theta_i)\delta(\theta_i - \theta_r)\delta(\phi_i - \phi_r + \pi)/(\sin\theta_i \cos\theta_i). \quad \text{(A.14)}$$

In actuality, real specular surfaces spread the reflected light into a narrow beam, rather than a single pulse as is suggested by the delta functions in (A.14). The width of this beam is determined by properties of the surface material. Section 2.4 discusses this in more detail.

A.2 Scene radiance and reflectance maps

A.2.1 Introduction

Up to now the incident and reflected radiance distributions have been specified with respect to a locally defined direction—that of the surface normal $\hat{\mathbf{n}}$. So expression (A.7) is only in a useful form for local analysis. A new coordinate system is required in order to investigate the dependence of reflected radiance, L_r, on the direction of $\hat{\mathbf{n}}$. Horn and Sjoberg [61] coined the term *scene radiance* for L_r expressed in terms of the normal direction. The term *reflectance map* is also used—however this usually refers to the case where the parameters (p, q) are used to specify the normal direction. The definition of direction (p, q) is given in section A.2.3. As L_r is proportional to the image irradiance, via (A.5),

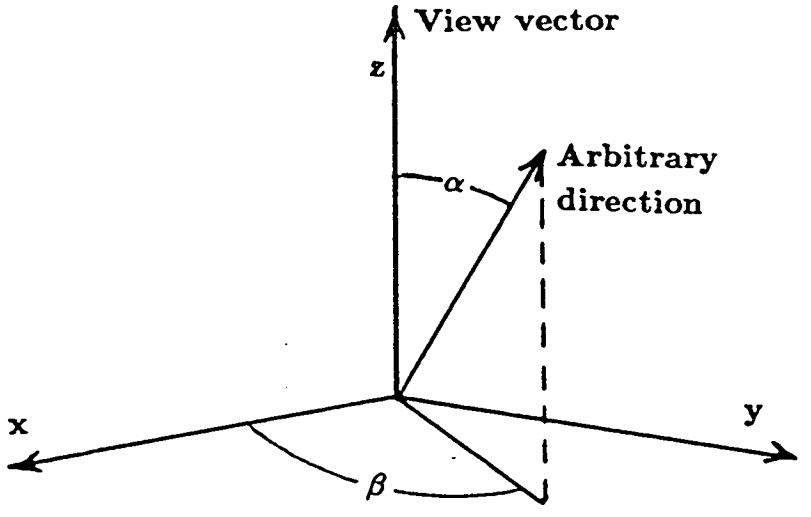


Figure A–5: *The new polar coordinate system in which an arbitrary direction* (α, β) *is defined with respect to the view vector.* α *is the polar angle and* β *is the angle of azimuth.*

the reflectance map is also. This relation is known as the *image irradiance equation*:

$$E_P = a_c R(\hat{\mathbf{n}}). \tag{A.15}$$

This is equation (2.5).

A.2.2 The viewer-centred coordinate system

Horn and Sjoberg [61] introduced a viewer centred coordinate system—as shown in figure A–5. They did this so that reflected radiance could be expressed in terms of the surface normal direction. In this system an arbitrary direction (α, β) is defined with respect to the view vector. α is the polar angle and β is the angle of azimuth. The surface normal direction is denoted as (α_n, β_n), and

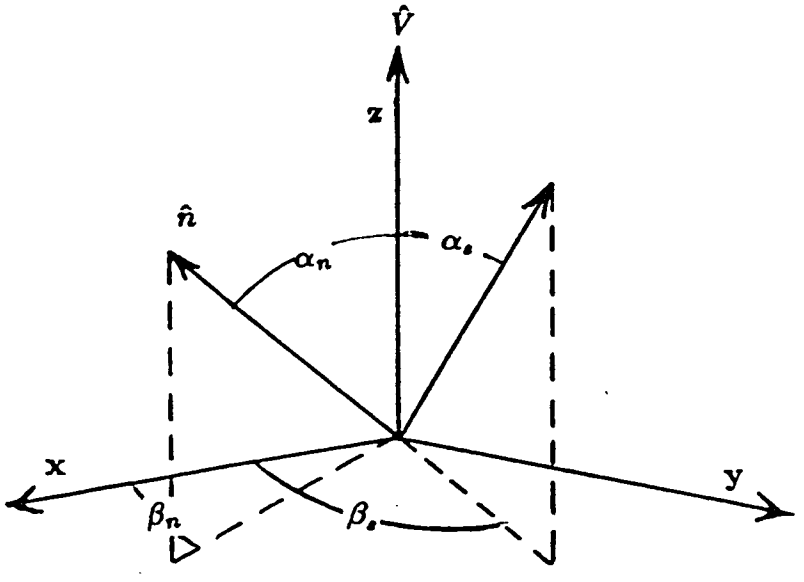


Figure A–6: *Scene radiance depends on the direction of the local surface normal vector* $\hat{\mathbf{n}}$*, with respect to the view vector. Its direction is defined as* (α_n, β_n)*.* (α_s, β_s) *is an arbitrary direction of incidence.*

an arbitrary direction of incidence is denoted as (α_s, β_s)[3]. In this new system the incident and reflected irradiance distribution are parameterised as $L_i(\alpha_s, \beta_s)$ and $L_r(\alpha_n, \beta_n)$ respectively. Figure A–6 shows the geometry. Now (A.7) can be rewritten to give an expression for the scene radiance or the reflectance map:

$$R(\hat{\mathbf{n}}) = L_r(\alpha_n, \beta_n) = \int_{-\pi}^{\pi} \int_{0}^{\pi} f_r(\theta_i, \phi_i; \theta_r, \phi_r) L_i(\alpha_s, \beta_s) \cos\theta_i \sin\theta_i d\theta_i d\phi_i. \tag{A.16}$$

The integration is still over all directions of incidence but the source direction (α_s, β_s) must be calculated for each of these directions. This can be done using the trigonometrical relations provided by Horn and Sjoberg [61]. These relations relate (α_s, β_s) and (α_n, β_n) to (θ_i, ϕ_i) and (θ_r, ϕ_r).

In the case of a Lambertian surface, f_r is constant, see section A.1.5 so that a Lambertian reflectance map R_L is:

$$R_L(\hat{\mathbf{n}}) = \frac{\rho_d}{\pi} \int_{-\pi}^{\pi} \int_{0}^{\pi} L_i(\alpha_s, \beta_s) \cos\theta_i \sin\theta_i d\theta_i d\phi_i, \tag{A.17}$$

or more compactly:

$$R_L(\hat{\mathbf{n}}) = \frac{\rho_d}{\pi} \int_{\omega_i} L_i \cos\theta_i d\omega_i. \tag{A.18}$$

A.2.3 Partial derivatives of smooth surfaces, (p, q)

Scene radiance is reflected radiance expressed as a function of the local surface normal direction. The distribution of scene radiance is called a reflectance map, strictly,[4] only when the normal direction is specified by the partial derivatives of the surface (p, q). Definitions of p and q ares given in section 5.1. Horn and Sjoberg [61] provide the an expression for (p, q) in terms of the (α_n, β_n):

$$(p, q) = (\cos\beta_n \tan\alpha_n, -\sin\beta_n \tan\alpha_n). \tag{A.19}$$

---

[3]This thesis introduces the notation (α, β) to avoid any unnecessary confusion. Horn and Sjoberg do not use this notation. In their notation $(\theta_n, \phi_n) = (\alpha_n, \beta_n)$, and $(\theta_s, \phi_s) = (\alpha_s, \beta_s)$.

[4]Reflectance map is also used loosely to mean scene radiance in this thesis.

Figure A–4 shows the particular (x, y) coordinates used to obtain this expression. Figure 2–7 (page 19) shows part of a typical reflectance map, $R(p, q)$. It represents a directionally lit surface that reflects light both diffusely and specularly.

Appendix B

The Fresnel relations

The Fresnel relations concern the fraction, F, of incident flux that is reflected by an optically smooth surface of a dielectric or a metal. This appendix gives these relations. These relations demonstrate—as was indicated in section 2.4—that it is not simple to predict the brightness of a specularity. F varies with both the angle of incidence and the polarisation of incident light. F_S is used to denote F for plane polarised incident light with its plane of polarisation normal to the plane of incidence. F_D denotes the orthogonal component, i.e. when the two planes are parallel. For light incident at angle θ_i, as shown in figure 2–2, then (quoted from [94])

$$F_S(\theta_i) = \frac{a^2 + b^2 - 2a\cos\theta_i + \cos^2\theta_i}{a^2 + b^2 + 2a\cos\theta_i + \cos^2\theta_i}, \tag{B.1}$$

$$F_D(\theta_i) = F_S(\theta_i)\frac{a^2 + b^2 - 2a\sin\theta_i\tan\theta_i + \sin^2\theta_i\tan^2\theta_i}{a^2 + b^2 + 2a\sin\theta_i\tan\theta_i + \sin^2\theta_i\tan^2\theta_i} \tag{B.2}$$

where

$$2a^2 = [(\acute{\eta}^2 - k^2 - \sin^2\theta_i)^2 + 4\acute{\eta}^2k^2]^{\frac{1}{2}} + (\acute{\eta}^2 - k^2 - \sin^2\theta_i), \tag{B.3}$$

and

$$2b^2 = [(\acute{\eta}^2 - k^2 - \sin^2\theta_i)^2 + 4\acute{\eta}^2k^2]^{\frac{1}{2}} - (\acute{\eta}^2 - k^2 - \sin^2\theta_i). \tag{B.4}$$

The optical constants of the material are described by the the complex index of refraction $\eta = \acute{\eta} + \mathrm{i}k$, in which the quantities are wavelength dependent.

Appendix C

Choosing a threshold for the retinex-based test

C.1 Introduction

In section 4.4.3 a six percent local contrast threshold is chosen for the retinex process. This choice is only appropriate for images not corrupted by noise. In all practical circumstances noise in the imaging process superimposes random gradients into the image irradiance signal—so applying the threshold to these noise corrupted gradients produces unreliable results. The thresholding operation can be made noise resistant if the signal is first smoothed. Smoothing not only counteracts the effects of noise, it also reduces the size of the genuine irradiance gradients in the signal (see figure 4–3, page 87). Consequently the value of the threshold ought to be reduced to compensate for this. As different imaging processes incur different noise levels they require different amounts of smoothing. In each case, the appropriate reduction in the threshold value should be determined by the amount of smoothing applied. A practical suggestion of how an appropriate threshold can be chosen is made below.

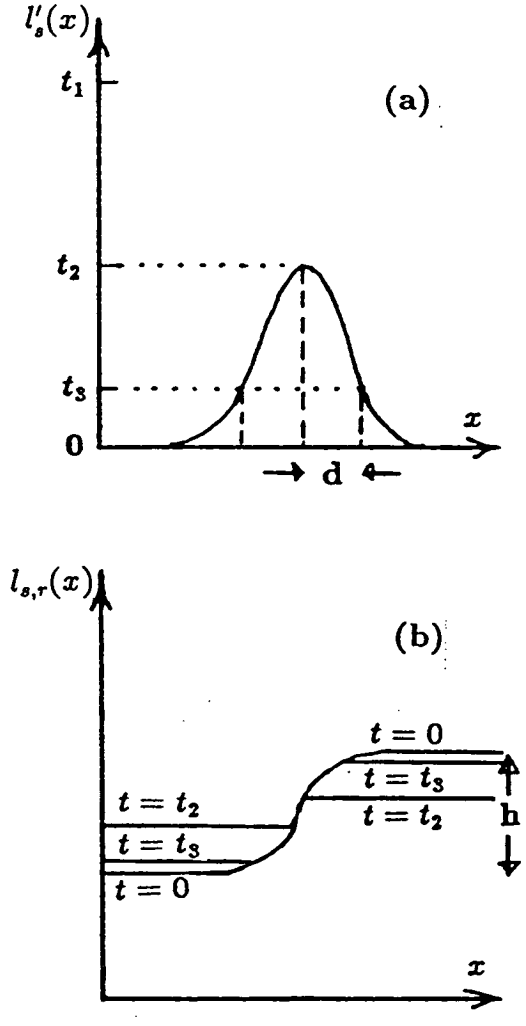


Figure C–1: *(a) A smoothed signal l_s creates a wide pulse of smaller gradients. (b) In this case, the height of the reconstructed step varies with value of t.*

C.2 Choosing a threshold for a smoothed image

C.2.1 The noiseless case—no smoothing

In this case the threshold operation is made clearer by considering the logarithmic signal: $l = \log g$. This simplifies the expression (4.3) to a simple difference threshold: i.e. the threshold is exceeded if

$$|l(x_1) - l(x_2)| \geq h, \tag{C.1}$$

where h corresponds to a 6% local contrast ($h = |\log 1.06|$). This is equivalent to thresholding the gradient l' of the signal l: i.e. the threshold is exceeded if

$$|l'| \geq t_1, \tag{C.2}$$

where t_1 corresponds to a gradient produced by the step of height h—see figure 4–3 (a) and (c). Using this threshold the retinex will preserve all steps in the signal exceeding the 6% local contrast—see figure C.1 (b). It will remove illumination variations that produce smaller contrasts.

C.2.2 The noisy case—smoothed images

As figure 4–3 (b) and (d) show smoothing widens the pulse of gradients corresponding to the step of height h. As a consequence, the pulse peaks at a lower gradient, t_2—so the step would be missed if a threshold t_1 were used. For a smoothed signal l_s the threshold value must be reduced to $t \leq t_2$ before the retinex starts to preserve the step. In order to avoid unduly contracting the dynamic range of the signal, the retinex ought to preserve as much of the step as possible. In this way the dynamic range test (equation (4.2)) maintains its potential for detecting specularities. Figure C.1 shows how much of the step height is preserved for various values of t:

- if $t = t_2$ only a tiny step height is preserved—so dynamic range is severely contracted.

- if $t = 0$ the entire step height h is preserved. However in this case the retinex will not fulfill its task of removing the illumination variations.

A compromise can be made by choosing a value of $t = t_3 : 0 < t_3 < t_2$—see figure C.1. In this way a large part of the step height can be preserved while the illumination variations are eliminated. An appropriate value of t_3 can be chosen empirically to suit the circumstances. The appropriate value will depend on the amount of smoothing applied. A convenient way to specify t_3 for an arbitrary amount of smoothing is given below:

1. Smooth by convolving the image signal l with a 2-D gaussian G_σ ($l_s = G_\sigma \otimes l$). The width of the gaussian σ determines the amount of smoothing: the larger σ the greater the smoothing. A value of σ just large enough to counteract the effects of the noise should be chosen. Using a width any larger unduly reduces the accuracy with which large gradients [24] can be located.

2. Select a threshold t_3 corresponding to the gradient, l', at distance d from the peak due to the step—as shown in figure C.1 (a). By specifying d in units of σ an appropriate value of t_3 can be selected when an arbitrary amount of smoothing is applied.

The results in this thesis were obtained using parameters $\sigma = 1$ pixel and a value of t_3 at distance $d = 1\sigma$.

Appendix D

Extracting profiles for the local tests

D.1 Introduction

This appendix describes the method used to extract, profiles along which the local tests can be applied. These are the profiles that were discussed in section 4.5.3. Figure 4–5 (page 93) depicts an example of one. Each such profile is specified completely by

1. The image location of the local maximum that it intersects.

2. Its direction in the image. As indicated in figure 4–5 the appropriate direction is perpendicular to a spine of maxima contained by an elongated blob, or radially through the maximum of a compact blob.

From this specification each profile can be extracted, by starting at the maximum and tracking in a straight line through the image in the two opposite directions. This is shown in figure D–1. So the extraction of profiles can be achieved in three stages:

1. Locating each local maximum.

2. Calculating the appropriate directions.

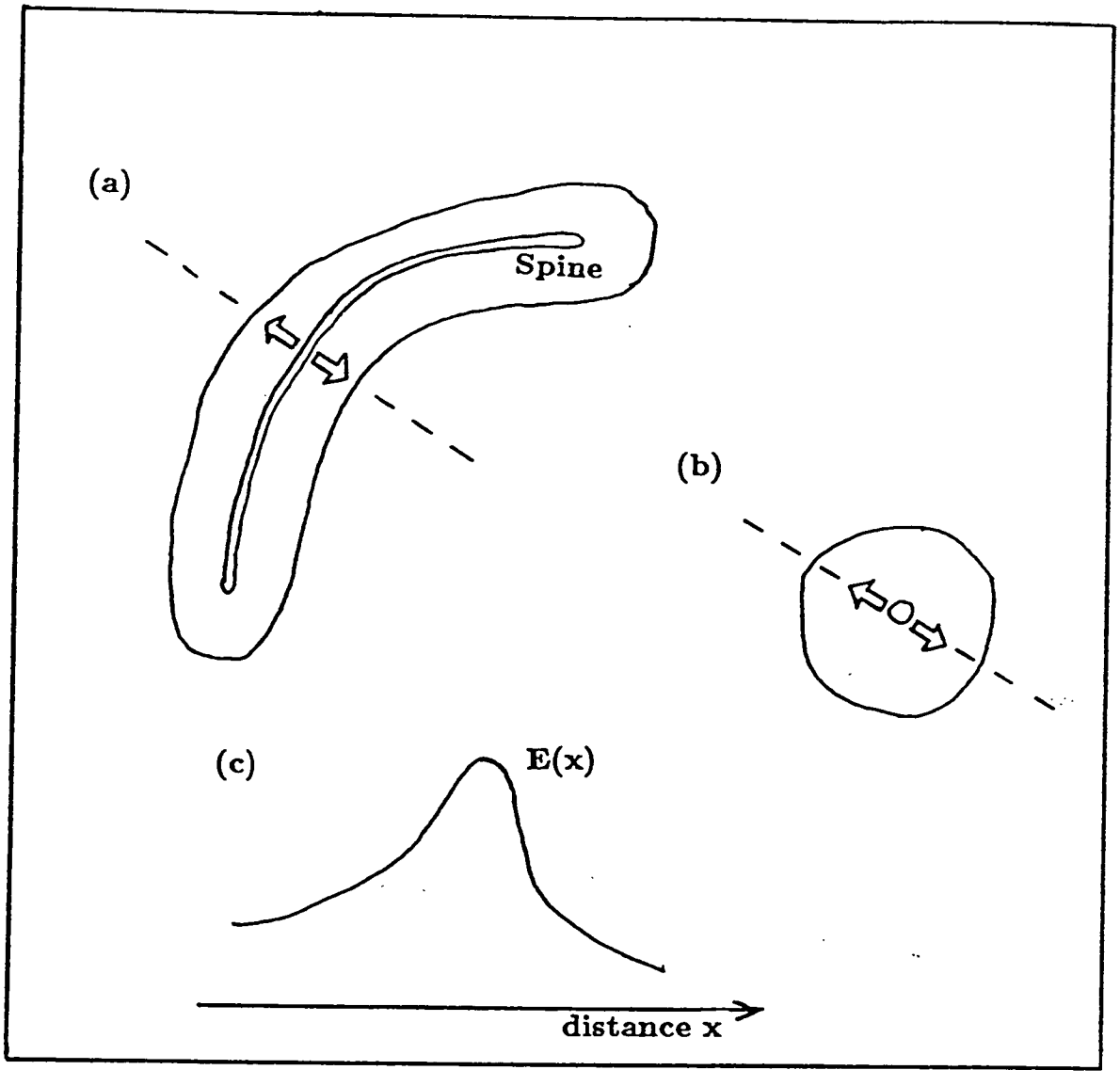


Figure D–1: *Profiles are extracted by tracking in the appropriate opposite directions, starting at the local maximum. (a) For a blob containing a spine the appropriate directions are along its perpendicular. (b) For a blob containing a single maximum, radial directions are tracked. (c) A typical profile extracted.*

3. Tracking through the image along those directions to extract the profile.

Only profiles that produce significant edges will be considered here[1]. These edges correspond to the borders of the blobs depicted in figure 4–5. An edge detector, the Canny operator [24] is used to mark the locations of edges in the image. This operator also provides an estimate of the direction of greatest ascent[2] at each edge location. Both the location and direction of the edges is used to initiate the search for appropriate profiles along which to apply the local tests. This is done as follows. From each edge location a straight path is followed through the image in the direction of greatest ascent—see figure 4–6 (a) (page 95). If the next edge location encountered along this path has a direction (of greatest ascent) approximately opposite to the path's direction (as in figure 4–6 (b)) then the profile along the path is suitable for the local tests. The profile's specification consists of the path's direction and the location of the maximum signal along the path.

By itself, this method often fails to extract profiles through many of the local maxima with spines. This is likely to occur when image noise corrupts the edge data. When this happens some of the edges surrounding a blob may be missing or the edge direction information may be poor. The method can be made noise-resistant by a slight modification: the single path initiated at each edge location is replaced by a narrow beam. Each beam, contains a set of paths centred about the original path direction—as shown in figure 4–7 (page 96). For the images shown in this thesis, it was found that a beam of width 10 degrees provided adequate noise resistance. In noiser circumstances a wider beam may be necessary. In general the narrowest beam that the noise-level permits should be used. In this way the number of paths that need to be tracked is kept to a minimum, so that the method is not unduly slow.

---

[1] See section 4.5.3 for reasons why.

[2] Greatest ascent in the image irradiance signal.

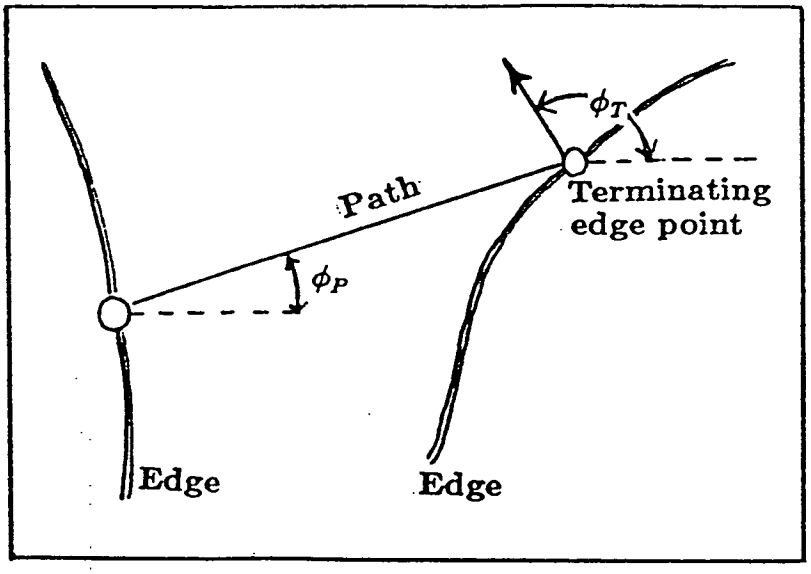


Figure D–2: *A path with direction ϕ_P is terminated by an edge that has a direction of greatest ascent ϕ_T.*

D.2 Selecting profiles from paths

Often more than one path intercepts a given local maximum. This is especially true when beams are used—as described above. Rather than extracting every profile along these paths, the most appropriate can be selected. This is done by considering the direction ϕ_P of each path that intercepts the local maximum. The path that is terminated by an edge in the 'most opposite' direction is chosen. The profile is extracted along this path. More precisely the path selected is that which minimises

$$\pi - |\phi_P - \phi_T|, \tag{D.1}$$

where ϕ_T is the direction of greatest ascent of the terminating edge—as shown in figure D–2.

Another consideration should be made choosing the path: any profile containing more than one local shading maxima is unsuitable for applying the local tests—and should be discounted. However, because noise often corrupts the image signal and creates spurious local maxima, additional care is needed.

Genuine shading maxima can be identified if the signal is smoothed before extracting the profile. This smoothing can be achieved by convolving the image with an appropriate gaussian—see appendix C. But smoothing, distorts the characteristic specular features that the local tests are looking for—i.e. local contrasts are reduced and sharp peaks are flattened. For this reason it is best to use the smooth signal for locating the local maxima, and the unsmoothed signal for extracting the profiles to which the test are applied.

It should be noted that other methods for identifying local maxima in images exist—notably Haralick's [51]. There is scope for improving the method described above by incorporating the more formal ideas of scale and smoothing that these other methods provide.

Appendix E

Derivation of the upper bound for the cylinder test

E.1 Introduction

In this appendix the upper bound U, on the second derivative of image irradiance at the peak in a profile is derived. The expression obtained for U is quoted in the discussion of the cylinder test in section 4.4.5. The derivation is in two parts:

1. deriving the upper bound corresponding to a profile through a whole cylinder.

2. evaluating the bound for an arbitrary portion of a cylinder.

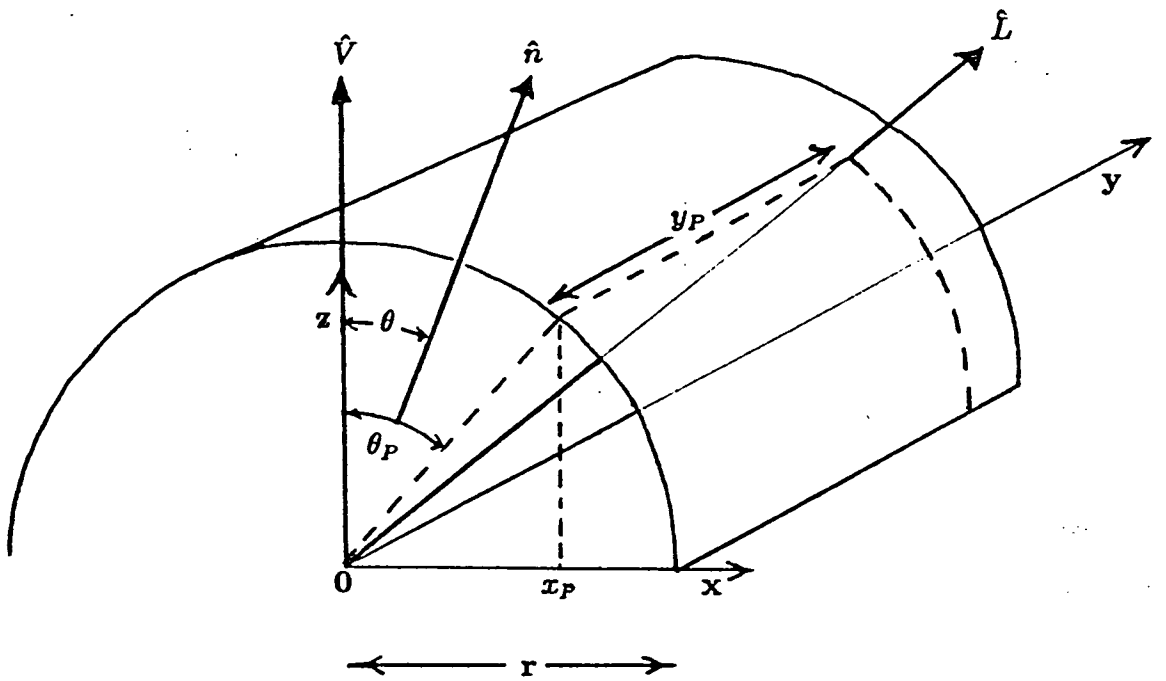


Figure E–1: *A cross-section through a cylindrical surface, with radius r. Viewing is along the vector* $\hat{\mathbf{V}}$. *The surface is lit along the vector* $\hat{\mathbf{L}}$. *The local surface normal at an arbitrary point along the cross-section is* $\hat{\mathbf{n}}$.

E.2 Derivation for a whole cylinder

Here, an image irradiance profile corresponding to a cross-section through a Lambertian cylinder is considered. As shown in figure E–1, the cross-section is taken in a direction perpendicular to the cylinder's axis—so that the profile captures the most rapid shading variations—in accordance with the principle discussed in section 4.5.3. x and y are image coordinates under parallel projection while z is depth. The view vector $\hat{\mathbf{V}}$ is as shown. In general V can lie out of the plane of the profile (the x-z plane). However as Lambertian shading is the same for all V—it is convenient to choose it as shown. The local surface normal $\hat{\mathbf{n}}$, at an arbitrary point along the profile has components:

$$\hat{\mathbf{n}} = (\sin\theta, 0, \cos\theta), \tag{E.1}$$

where θ is the angle shown. The direction of the illumination[1] is specified by vector $\hat{\mathbf{L}}$. In general $\hat{\mathbf{L}}$ can lie out of the x-z plane, so it is represented as:

$$\hat{\mathbf{L}} = (y_P^2 + r^2)^{-\frac{1}{2}}(r\sin\theta_P, y_P, r\cos\theta_P), \tag{E.2}$$

where r is the radius of the cylinder and length y_P and angle θ_P are as shown in figure E–1.

The collimated-plus-ambient illumination model assumed in section 4.5.5 results in an image irradiance equation along the Lambertian profile described by:

$$E(x) = E_{max}[A + (1 - A)\max(0, \hat{\mathbf{n}} \cdot \hat{\mathbf{L}})]. \tag{E.3}$$

This is obtained by combining the expressions (4.1) and (2.8). Again, E is image irradiance, A is the constant ambient term ($A:\ 0 \leq A \leq 1$, typically $A = 0.3$). E_{max} is the maximum value of E produced by the homogeneous Lambertian surface. For the purposes of deriving the upper bound on the second derivative

---

[1]The assumption of a directional illumination is discussed in section 4.5.5.

at the peak, only the image irradiance signal near the peak need be considered. Thus (E.3) simplifies to

$$E(x) = E_{max}[A + (1 - A)\hat{\mathbf{n}} \cdot \hat{\mathbf{L}}]. \tag{E.4}$$

The dot product:

$$\hat{\mathbf{n}} \cdot \hat{\mathbf{L}} = (y_P^2 + r^2)^{-\frac{1}{2}} r \cos(\theta - \theta_P), \tag{E.5}$$

derived from (E.1) and (E.2), is substituted into (E.4) to obtain:

$$E(x) = E_{max}[A + (1 - A)(y_P^2 + r^2)^{-\frac{1}{2}} r \cos(\theta - \theta_P)]. \tag{E.6}$$

The second derivative of E at an arbitrary point x along the profile is obtained using the chain rule:

$$\frac{d^2E}{dx^2} = \frac{d^2E}{d\theta^2}\left(\frac{d\theta}{dx}\right)^2 + \frac{dE}{d\theta}\frac{d^2\theta}{dx^2} \tag{E.7}$$

and the relationship between x and θ:

$$x = r \sin\theta. \tag{E.8}$$

The various terms in (E.7) are easily shown to be:

$$\frac{d\theta}{dx} = \frac{1}{r\cos\theta}, \tag{E.9}$$

$$\frac{d^2\theta}{dx^2} = \frac{\tan\theta}{r^2\cos^2\theta}, \tag{E.10}$$

$$\frac{dE}{d\theta} = -E_{max}(1 - A)(y_P^2 + r^2)^{-\frac{1}{2}} r \sin(\theta - \theta_P), \tag{E.11}$$

$$\frac{d^2E}{d\theta^2} = -E_{max}(1 - A)(y_P^2 + r^2)^{-\frac{1}{2}} r \cos(\theta - \theta_P). \tag{E.12}$$

So that:

$$\frac{d^2E}{dx^2} = -E_{max}(1 - A) r (y_P^2 + r^2)^{-\frac{1}{2}} \left(\frac{\cos(\theta - \theta_P) + \sin(\theta - \theta_P)\tan\theta}{r^2\cos^2\theta}\right). \tag{E.13}$$

At the profile's peak, $\theta = \theta_P$ and $x = x_P$, so:

$$\begin{aligned} |d^2E/dx^2|_{x=x_P} &= E_{max}(1 - A) r (y_P^2 + r^2)^{-\frac{1}{2}} / (r\cos\theta_P)^2, \\ &= E_{max}(1 - A) r (y_P^2 + r^2)^{-\frac{1}{2}} / (r^2 - x_P^2). \end{aligned} \tag{E.14}$$

This expression is maximum when $y_P = 0$ and $A = 0$. So the upper bound on $|d^2E/dx^2|$ at the peak x_P in a profile through a Lambertian cylinder of radius r is:

$$\frac{E(x_P)}{r^2 - x_P^2}. \tag{E.15}$$

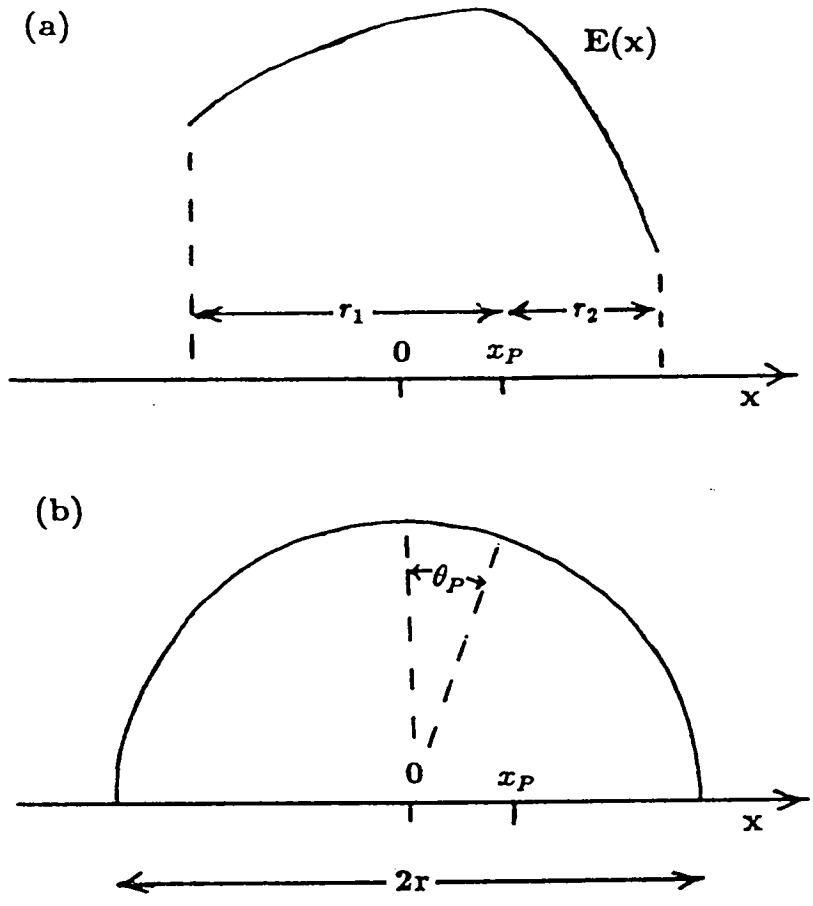


Figure E–2: *(a) An example of an image irradiance profile corresponding to a portion of the cross-section through a Lambertian cylinder (b).*

E.3 Evaluating the upper bound for an arbitrary portion of a cylinder

The cylinder test applies not only to a whole cylinder but also to smaller portions. Consequently, the upper bound (E.15) is not directly applicable, because the values of r and x_P cannot be measured from a profile. However the upper bound can be extended so that it is specified using parameters that can be measured from a profile. These parameters are the distances from the peak to the two flanking edges, r_1 and r_2 - as shown in figure E–2. The upper bound is extended by considering the largest value it can assume with respect to x_P

and r. From figure E–2 it can be seen that $r_1 + r_2 \leq 2r$ so that x_P is thus constrained:

$$|r - |x_P|| \geq \min(r_1, r_2). \tag{E.16}$$

So the upper bound (E.15) is largest when

$$x_P = r - \min(r_1, r_2). \tag{E.17}$$

In this case the upper bound can be expressed as

$$\frac{E(x_P)}{\min(r_1, r_2)[2r - \min(r_1, r_2)]}. \tag{E.18}$$

Again, from figure E–2 it can be seen that this upper bound is a maximum value, U, when $2r = r_1 + r_2$, so that

$$\begin{aligned} U &= E(x_P)/(\min(r_1, r_2)[r_1 + r_2 - \min(r_1, r_2)]), \\ &= E(x_P)/(\min(r_1, r_2)\max(r_1, r_2)), \\ &= E(x_P)/(r_1 r_2). \end{aligned} \tag{E.19}$$

It is this latter expression that is quoted as equation (4.5) in the cylinder test.

Appendix F

Obtaining a compact description of specularities

F.1 Introduction

This appendix shows how the compact description of specular blobs—defined in section 4.6.3—is extracted directly from the specularity map. A blob's description contains the parameters shown in figure 4–11 (page 107): the image coordinates of its centroid, (x_c, y_c), its approximate dimensions: height h and width w and the direction of its major axis, θ. In addition there is a boolean parameter e, that is set if the blob's shape is approximately elliptical.

Although the above description often provides enough information for shape to be successfully inferred, some specularities are not well represented by it. For example the curving nature of the crescent shaped specularities in figure 4–13 (page 110) is not represented. Scope exists for extending this scheme to provide more accurate compact descriptions for a greater variety of specularities.

Each parameter listed in the description above is obtained by a statistical measure involving the image location of each pixel within the blob[1]. Here, each

[1]The statistical measures are based purely on position. An improved description

location is denoted (x_i, y_i), where the index i is an integer in the range 1 to n—the total number of pixels in the blob. The centroid of each blob is located at its centre of gravity:

$$(x_c, y_c) = \left((1/n) \sum_{i=1}^{n} x_i,\ (1/n) \sum_{i=1}^{n} y_i \right). \tag{F.1}$$

The parameters h,w and θ are estimated using the principal axes method (see next section) while e is obtained by a least-squares fit of an ellipse to the blob's perimeter (see section F.3)

F.2 Principal axes method

The parameters h, w and θ are estimated using the principal axes method. A detailed description of the method is given in [9] page 486: in this section only the basic steps are outlined. These steps are

1. Construct the scatter matrix M_s.
2. Compute the eigenvalues and eigenvectors of M_s.
3. Use the eigenvectors to obtain the direction of the major axes of the blob θ.
4. Use the eigenvalues to obtain the blob dimensions h and w.

M_s is constructed as follows:

$$\mathsf{M}_s = \begin{pmatrix} a_s & b_s \\ b_s & c_s \end{pmatrix} \tag{F.2}$$

where

$$a_s = (1/n) \sum_{i=1}^{n} (x_i - x_c)^2, \tag{F.3}$$

---

might be obtained if the irradiance signal within the blob were also exploited—perhaps by fitting a patch to the irradiance function at the blob.

$$b_s = (1/n)\sum_{i=1}^{n}(x_i - x_c)(y_i - y_c), \tag{F.4}$$

and

$$c_s = (1/n)\sum_{i=1}^{n}(y_i - y_c)^2. \tag{F.5}$$

The eigenvalues of M_s can be shown to be:

$$\lambda_1 = (a_s + c_s + r_s)/2,\ \lambda_2 = (a_s + c_s - r_s)/2, \tag{F.6}$$

where

$$r_s = \sqrt{(a_s - c_s)^2 + 4b_s^2}. \tag{F.7}$$

The corresponding (unnormalised) eigenvectors are then:

$$\mathbf{X_1} = \begin{pmatrix} -2b_s \\ a_s - c_s - r_s \end{pmatrix},\ \mathbf{X_2} = \begin{pmatrix} -2b_s \\ a_s - c_s + r_s \end{pmatrix}. \tag{F.8}$$

The x and y components of $\mathbf{X_1}$ are used to obtain the direction θ:

$$\theta = \tan^{-1}\left(\frac{\mathbf{X}_{1y}}{\mathbf{X}_{1x}}\right) = \tan^{-1}\left(\frac{a_s - c_s - r_s}{-2b_s}\right). \tag{F.9}$$

The estimates of h and w can be shown to be:

$$h = 1 + 2\sqrt{\lambda_1} = 1 + \sqrt{2(a_s + c_s + r_s)}\ \text{pixels}, \tag{F.10}$$

$$w = 1 + 2\sqrt{\lambda_2} = 1 + \sqrt{2(a_s + c_s - r_s)}\ \text{pixels}. \tag{F.11}$$

For the purposes of error analysis the values of h and w may be regarded as accurate up to ± 1 pixel.

F.3 Ellipse fitting

The boolean value e is computed for each blob by measuring how well an ellipse fits the blob's perimeter. The dimensions and orientation of the ellipse to be fitted are determined using the parameters h, w and θ, provided by the principal axes method, above. The length of its semi-major and semi-minor axes are $a_e = h/2$ and $b_e = w/2$ respectively. The major axis makes an angle θ with the direction of the x-axis in the image—as shown in figure 4–12 (page 108). The ellipse is fitted so that its origin coincides with the blob's centroid (x_c, y_c). The fitting process involves computing an error measure ϵ that represents the deviation of the blob's perimeter from the ellipse being fitted. The equation of an ellipse:

$$\frac{x_e^2}{a_e^2} + \frac{y_e^2}{b_e^2} = 1 \tag{F.12}$$

is used to derive the expression for ϵ. Equation (F.12) is expressed in the ellipse's local coordinate system: x_e and y_e are measured along the semi-major and semi-minor axes respectively—see figure 4–12 (page 108). The square deviation for each pixel located on the blob's perimeter[2], $(x_{e,i}, y_{e,i})$ is

$$d_i^2 = \left(\frac{x_{e,i}^2}{a_e^2} + \frac{y_{e,i}^2}{b_e^2} - 1 \right)^2. \tag{F.13}$$

The mean square deviation for all m pixels on the perimeter is

$$\epsilon^2 = (1/m) \sum_{i=1}^{m} d_i^2. \tag{F.14}$$

in order to compute ϵ the local coordinates $(x_{e,i}, y_{e,i})$ must be derived from the global image coordinates (x_i, y_i). This is done by a coordinate frame transformation T_e:

$$(x_{e,i}, y_{e,i}) = \mathsf{T}_e \begin{pmatrix} x_i - x_c \\ y_i - y_c \end{pmatrix} \tag{F.15}$$

---

[2]For practical purposes a pixel lies on the perimeter if some, but not all of its 4 nearest neighbours lie within the blob.

where

$$\mathsf{T}_e = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}. \tag{F.16}$$

Once computed ϵ is thresholded against a chosen value, t_e in order to determine e for the blob:

$$e = \begin{array}{ll} 1 & \text{if } \epsilon \leq t_e, \quad \text{i.e. ellipse description accepted,} \\ 0 & \text{otherwise.} \end{array} \tag{F.17}$$

For the specularities described in this thesis a value of $t_e = 1.6$ pixels was found to usefully indicate elliptical blobs.

Appendix G

Inferring surface shape: mathematical appendix

G.1 Derivation of constraint in (r_1, r_2) space

Restating equation (6.14) and substituting for $r_1 = 1/k_1$, $r_2 = 1/k_2$:

$$\begin{aligned} H_{xx} &= k_1 \cos^2\varphi + k_2 \sin^2\varphi &&= (r_2 \cos^2\varphi + r_1 \sin^2\varphi)/(r_1 r_2), \\ H_{xy} &= (k_2 - k_1)\cos\varphi \sin\varphi &&= (r_1 - r_2)\cos\varphi \sin\varphi/(r_1 r_2), \end{aligned} \tag{G.1}$$

gives:

$$H_{xx}{}^2 + H_{xy}{}^2 = (r_2^2 \cos^2\varphi + r_1^2 \sin^2\varphi)/(r_1 r_2)^2. \tag{G.2}$$

Defining

$$A = \frac{H_{xx}}{H_{xx}{}^2 + H_{xy}{}^2} = \frac{(r_2 \cos^2\varphi + r_1 \sin^2\varphi) r_1 r_2}{r_2^2 \cos^2\varphi + r_1^2 \sin^2\varphi}, \tag{G.3}$$

$$B = \frac{H_{xy}}{H_{xx}{}^2 + H_{xy}{}^2} = \frac{(r_1 - r_2) r_1 r_2 \cos\varphi \sin\varphi}{r_2^2 \cos^2\varphi + r_1^2 \sin^2\varphi}, \tag{G.4}$$

gives

$$(r_1 - A) = \frac{r_1^2 \sin^2\varphi (r_1 - r_2)}{r_2^2 \cos^2\varphi + r_1^2 \sin^2\varphi}, \tag{G.5}$$

$$(r_2 - A) = \frac{r_2^2 \cos^2\varphi (r_2 - r_1)}{r_2^2 \cos^2\varphi + r_1^2 \sin^2\varphi} \tag{G.6}$$

and so the equation of a right-rectangular hyperbola (6.16) is obtained:

$$(r_1 - A)(r_2 - A) = -\frac{(r_1 - r_2)^2 r_1^2 r_2^2 \cos^2\varphi \sin^2\varphi}{(r_2^2 \cos^2\varphi + r_1^2 \sin^2\varphi)^2} = -B^2. \tag{G.7}$$

G.2 Locally hyperbolic surface: derivations

When $k_1 = k_h$ and $k_2 = -k_h$ the hessian is given by equation (6.17) which is restated below:

$$\mathsf{H} = \mathsf{R}(\theta_0) k_h \begin{pmatrix} 1 & . \\ . & -1 \end{pmatrix} \mathsf{R}(-\theta_0). \tag{G.8}$$

The rotation R is provided by equation (5.4) so that

$$\begin{aligned} \mathsf{H} &= \mathsf{R}(\theta_0) k_h \begin{pmatrix} 1 & . \\ . & -1 \end{pmatrix} \begin{pmatrix} \cos\theta_0 & -\sin\theta_0 \\ \sin\theta_0 & \cos\theta_0 \end{pmatrix}, \\ &= \mathsf{R}(\theta_0) k_h \begin{pmatrix} \cos\theta_0 & -\sin\theta_0 \\ -\sin\theta_0 & -\cos\theta_0 \end{pmatrix}, \\ &= k_h \begin{pmatrix} \cos\theta_0 & \sin\theta_0 \\ -\sin\theta_0 & \cos\theta_0 \end{pmatrix} \begin{pmatrix} \cos\theta_0 & -\sin\theta_0 \\ -\sin\theta_0 & -\cos\theta_0 \end{pmatrix}. \end{aligned} \tag{G.9}$$

From this expression equation (6.18) becomes apparent:

$$\mathsf{H} = k_h \begin{pmatrix} \cos^2\theta_0 - \sin^2\theta_0 & -2\cos\theta_0\sin\theta_0 \\ -2\cos\theta_0\sin\theta_0 & \sin^2\theta_0 - \cos^2\theta_0 \end{pmatrix} = k_h \begin{pmatrix} \cos 2\theta_0 & -\sin 2\theta_0 \\ -\sin 2\theta_0 & -\cos 2\theta_0 \end{pmatrix}. \tag{G.10}$$

The expressions for orientation (equation (6.19)) and magnitude (equations (6.20) and (6.21)) derive from the expression above and the constraint equation: $\mathsf{H}\,\mathbf{x} = -(\delta n_1, \delta n_2)^T$

$$k_h \begin{pmatrix} \cos 2\theta_0 & -\sin 2\theta_0 \\ -\sin 2\theta_0 & -\cos 2\theta_0 \end{pmatrix} \begin{pmatrix} R_1 \\ R_2 \end{pmatrix} = - \begin{pmatrix} \delta n_1 \\ \delta n_2 \end{pmatrix}. \tag{G.11}$$

This provides two equations:

$$k_h(R_1 \cos 2\theta_0 - R_2 \sin 2\theta_0) = -\delta n_1, \tag{G.12}$$

$$k_h(R_1 \sin 2\theta_0 + R_2 \cos 2\theta_0) = \delta n_2. \tag{G.13}$$

These rearrange to provide equations (6.20) and (6.21), while orientation is obtained by dividing the first by the second:

$$(R_2 \sin 2\theta_0 - R_1 \cos 2\theta_0)/(R_1 \sin 2\theta_0 + R_2 \cos 2\theta_0) = (\delta n_1/\delta n_2), \tag{G.14}$$

$$R_2 \tan 2\theta_0 - R_1 = (\delta n_1/\delta n_2)(R_1 \tan 2\theta_0 + R_2), \tag{G.15}$$

$$\tan 2\theta_0 = \frac{R_1 + (\delta n_1/\delta n_2)R_2}{R_2 - (\delta n_1/\delta n_2)R_1} = \frac{\delta n_2 R_1 + \delta n_1 R_2}{\delta n_2 R_2 - \delta n_1 R_1}. \tag{G.16}$$

Inverting this expression gives equation (6.19).

Appendix H

Monocular analysis: simple formulation

As for the stereo analysis, a linearised approximation, in the form of equation (5.15), page 142, is again used to represent a smooth surface. For the monocular analysis the following information is assumed to be available:

- $\mathbf{V}$—the view vector. Binocular stereo cannot be used to estimate this in a purely monocular system, but in a mixed system employing both monocular and stereoscopic analysis, $\mathbf{V}$ is available as before (section 6.3.2).

- $\mathbf{S}$—the position of the origin of the particular light source that created the specularity. Some independent knowledge source is assumed to supply this.

- $\mathbf{s}$—a light source "base-line" vector. This describes the displacement between a general point on the light source, relative to $\mathbf{S}$.

- $\widehat{\mathbf{L_B}}$—the direction of $\mathbf{L_B}$: the vector shown on figure H–1.

- $\widehat{\mathbf{V_B}}$—the direction of $\mathbf{V_B}$: the vector shown on figure H–1.

The vectors $\mathbf{V}$, $\mathbf{s}$ and $\mathbf{S}$, are also shown in the figure. A point on the distributed light source is specularly reflected at surface point A so as to traverse view

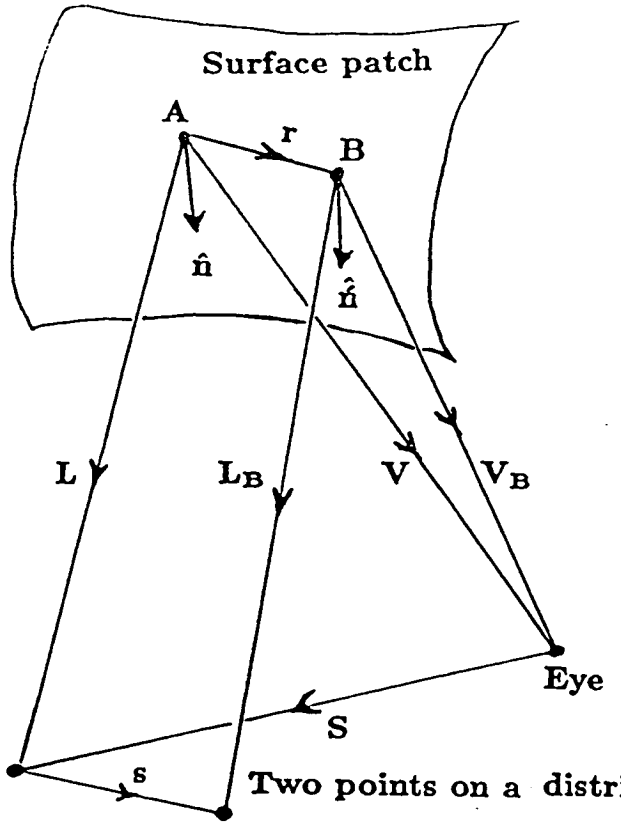


Figure H–1: *Geometry for the simple formulation of the monocular analysis (c.f. figure 5–6, page 145). A ray from a point on the distributed light source is specularly reflected at surface point A, with reflected ray along* **V**. *The ray from a point displaced from the first by the base-line vector* **s** *is similarly reflected at surface point B along vector* $\mathbf{V_B}$. *The vectors marked are used in the accompanying analysis.*

vector **V**. A point displaced from the first by the base-line vector **s** is similarly reflected at surface point B along view vector $\mathbf{V_B}$.

The simple formulation is analogous to that given for the stereo analysis. Again the vector $(\delta n_1, \delta n_2)^T$ is provided by considering the equations of mirror-like reflection at A and B and **x** from a simple vector equality, so that:

$$\begin{aligned} \mathsf{H}\mathbf{x} = - \;\; & (\delta n_1, \delta n_2)^T, \\ \text{where} \;\; & (\delta n_1, \delta n_2)^T = \left((\hat{\mathbf{n}}'-\hat{\mathbf{n}})\cdot\hat{\mathbf{i}},\, (\hat{\mathbf{n}}'-\hat{\mathbf{n}})\cdot\hat{\mathbf{j}}\right)^T \\ \text{and} \;\; & \mathbf{x} = \left(\mathbf{r}\cdot\hat{\mathbf{i}},\, \mathbf{r}\cdot\hat{\mathbf{j}}\right)^T, \end{aligned} \qquad \text{(H.1)}$$

where $(\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}})$ are again the unit vectors along the coordinate axes of the local frame, defined by equation (6.7).

$$\hat{\mathbf{n}} - \hat{\mathbf{n}} = (\hat{\mathbf{V}} + \hat{\mathbf{L}})/|\hat{\mathbf{V}} + \hat{\mathbf{L}}| - \left(\hat{\mathbf{V}_B} + \hat{\mathbf{L}_B}\right) / \left|\hat{\mathbf{V}_B} + \hat{\mathbf{L}_B}\right|, \tag{H.2}$$

where $\mathbf{L} = \mathbf{V} + \mathbf{S}$, and $\mathbf{r} = \mathbf{L} + \mathbf{s} - \mathbf{L_B}$. The only complication is in obtaining the magnitude $|\mathbf{L_B}|$ which is easily shown to be:

$$|\mathbf{L_B}| = \frac{(\mathbf{L}+\mathbf{s})\cdot\hat{\mathbf{n}}}{\hat{\mathbf{L_B}}\cdot\hat{\mathbf{n}}}. \tag{H.3}$$

The advantage of this formulation is that the short base-line assumption (i.e. $|\mathbf{s}| \ll |\mathbf{L}|\cos\sigma$—analogous to $|\mathbf{d}| \ll |\mathbf{V}|\cos\sigma$ for the stereo analysis) is no longer required. The angular variables corresponding to δ_m and α in the transformation equation (5.26) are $\hat{\mathbf{V}_B} - \hat{\mathbf{V}}$ and $\hat{\mathbf{L}_B} - \hat{\mathbf{L}}$, respectively. In this form, the transformation is no longer linear and is much harder to apply. For the applications discussed below the earlier linear transformation is considered.

Appendix I

Implementation details

I.1 Camera geometry

The camera geometry used in this thesis is as shown in figure 6–3: the cameras are oriented so that their optical axes verge at a point equidistant from their focal points while their vertical axes are perpendicular to the plane containing the optical axes. In this case vector $\mathbf{d}$ is determined by the vergence angle θ_v and the base-line separation b, i.e. in the left-eye frame:

$$\mathbf{d} = (b\cos\theta_v/2,\ 0, -b\sin\theta_v/2)^T. \tag{I.1}$$

In practice b and θ_v are usually fixed so that they can be accurately measured beforehand.

For this camera geometry the transformation matrix $\mathsf{T}_{r\to l}$—with which the angular position of the specularity in the right image is transformed into the left-eye is

$$\mathsf{T}_{r\to l} = \begin{pmatrix} \cos\theta_v & . & \sin\theta_v \\ . & 1 & . \\ -\sin\theta_v & . & \cos\theta_v \end{pmatrix}. \tag{I.2}$$

I.2 Combined inference

For a small field of view ($\sim 10°$) measurements made in the image plane provide adequate approximations to those made in polar projection. This is the case for the images processed by the system. In such cases, the descriptions of specularities (section 4.6.3) —in terms of height h, width w, orientation θ and boolean variable e—produced by the detector, are interpreted directly for the monocular analysis:

- If the source is circular and slant less than 5 degrees and the specularity is an ellipse (i.e. $e=1$) then the ratio $|k_1/k_2|$ is estimated as $\sqrt{a}$.
- If $a>2$ and the source is compact then the direction of least curvature is obtained as follows.

Four steps are required:

1. Angle θ in the specularity's description provides the unit vector of the projection of θ_{le} in the image: $\hat{\mathbf{X}}_e=(\cos\theta,\sin\theta)$.
2. Using $\hat{\mathbf{X}}_e$, the direction in the local surface frame is estimated: $\theta_{min}=\tan^{-1}\left(\hat{\mathbf{X}}_e\cdot\hat{\mathbf{j}}/\hat{\mathbf{X}}_e\cdot\hat{\mathbf{i}}\right)$. Again the axes of the local coordinate frame lie along the unit vectors $\left(\hat{\mathbf{i}},\hat{\mathbf{j}},\hat{\mathbf{k}}\right)$ - see page 154.
3. Next equation (5.34) is used. The local coordinate frame is rotated through an angle $\theta_{min}+\pi/2$. Consequently the stereo measurements, expressed in the new frame: $(\acute{\delta}n_1,\acute{\delta}n_2)$ and $(\acute{R}_1,\acute{R}_2)$, supply estimates of the absolute values of the principal curvatures: k_{max} and k_{min} ($|k_{max}|>|k_{min}|$).
$$(k_{max},k_{min})=-(\acute{\delta}n_1/\acute{R}_1,\acute{\delta}n_2/\acute{R}_2). \tag{I.3}$$
4. Finally the principal curvatures are obtained:
$$\begin{aligned}\textbf{if } k_{min}<k_{max} \textbf{ then}\quad & (k_1,k_2,\theta_0)=(k_{max},k_{min},\theta_{min}+\pi/2)\\ \textbf{else}\quad & (k_1,k_2,\theta_0)=(k_{min},k_{max},\theta_{min}).\end{aligned} \tag{I.4}$$

Appendix J

Software system

J.1 Introduction

The specularity detection and shape-from-specularity schemes are implemented as executable programs on a SUN-3 mini-computer under Unix 4.2. The programs are indicated in the figure as the straight-edged boxes in figure 7–1 on page 183. Some of these boxes represent more than one program: e.g. specularity detection includes both the retinex preprocessing and the tests for specularity. Below, details of the programs associated with each box are given. In addition the format of the input and output of each program is specified. All the programs reside in the directory `visionbin`.

J.2 Image file format

Grey-level images consist of a byte (8 bits) per pixel. The IFF image file format is used to represent them. In this format an image consists of:

1. a 512 byte-long header that contains the *height* and *width* of the image in pixels and specifies the viewing geometry.

2. A sequence of *height*×*width* bytes representing the grey-levels at each pixel, in raster-scan order: i.e. from left to right then from top to bottom.

An IFF image, *file.iff* is displayed on the screen using

```
% displaytool file.iff
```

One way to produce an IFF stereo image-pair: *filel.iff* and *filer.iff*, is to specify a stereo view using a WINSOM model file [95]: *file.model* and then run:

```
% winsom -r file
```

J.3 Edge detection

Edge detection is achieved by convolving the derivative of a gaussian (width *sigma*—default 1 pixel) with the logarithmic (base 2) image signal:

```
% canlog [-ssigma] file.iff file.canlog
```

followed by hysteresis thresholding, to remove all but the prominent edge segments:

```
% hyster [-hhi] [-llo] file.canlog file.hyster
```

Both *file.canlog* and *file.hyster* are in the format described below. The lower and higher thresholds may be set using the optional parameters *lo* and *hi*. Their default settings correspond to a 2% and 20% contrast between neighbouring pixels, respectively.

J.4 Edge file format

Both the `canlog` and `hyster` programs output edge segments in the following format[1]:

1. An IFF header of type EDGEL, again specifying the image dimensions and viewing geometry.

2. A 16 bit integer specifying the number of edge segments: n.

3. A set of n edge segments: each represented by a quadruple of floating point numbers (4 bytes each). A quadruple describes (1) the row pixel coordinate (from the top), (2) the column pixel coordinate (from the left), (3) the edge contrast and (4) the edge orientation (-180 to $180°$).

The values of the quadruples of all the edge segments are printed out using:

```
% predges file.edgel
```

The edge segments are displayed on the screen using:

```
% displaytool [-D] file.edgel
```

The `-D` option renders segments of opposite orientation in "opposite" colours.

[1]The AIVRU format with added header [1]

J.5 Specularity detection

J.5.1 The tests

The three tests for specularity are applied using:

`% dotests` *[-rl] [-ww] [-l file.retinex] file.hyster file.iff file.tests*

Unless specified, the maximum path length, l is 20 pixels and the beam width, w is 10°. The retinex-based test is only applied if the `-l` option is specified and *file.retinex* has been created using the `retinex` program—see below. The input *file.hyster* and *file.iff* are in the formats described above. The output in ASCII, in the Backus Naur form[2] is described below

file.tests	=	[# *An_optional_one_line_comment*]
		{ *column row evidence code a b* }
column, row	=	integer
evidence	=	"Good"|"Some"|"No"
code	=	{"r"|"c"|"l"|"n"}
a, b	=	integer

Thus at each (spine) point that the tests are applied, its image location (*column, row*) in pixels is written out, along with a *code* indicating the tests, if any, it has passed. The *code*[3] contains "r" if the retinex-based test is passed, "c" if the cylinder test is passed, "l" if the local contrast test is passed and "n" if none is passed. The *evidence* and the values of a and b are used by the `dodescribe` program.

---

[2]Enclosure within [···] is a simple optional clause. Enclosure within {···} indicates a repeatable clause. Alternatives in a clause are separated by a | sign.

[3]The *code* is case independent.

The format above can be converted into an image using `showtests` then displayed:

```
% showtests [-r] [-c] [-l] [-n] file.hyster file.test file.showtests
% displaytool file.showtests
```

where options specify initials of which tests are to be displayed. The points where any of the specified tests are passed are marked in white upon a grey background. Spatial reference is provided by marking the edge segments in black.

J.5.2 Retinex pre-processing

Retinex processing is applied to the original grey-level image, *file.iff*. The output is in the form of another IFF image, *file.retinex*:

```
% retinex thresh_op iter_ops report_ops file.iff file.retinex
```

The threshold stage is controlled through:

thresh_op = `[-t`*t*`]` `[-C`*file.retcan*`]`

Either a percentage local contrast threshold, t (default 6%) is specified or noise-resistant thresholding is achieved using *file.retcan*, the output from the program:

```
% retcan file.iff file.retcan
```

The iterative scheme that follows thresholding is specified through

iter_ops = `[-u`*u*`]` `[-i`*i*`]` `[-e`*e*`]` `[-j`*j*`]`

The scheme can be made to start with a retinex at a uniform grey-level, u vis the `-u` option, otherwise the logarithmic input signal is used. If `-i` is used the scheme stops after i iterations. If `-e` is used the scheme stops when the error-norm falls below 2^{-e}. The acceleration parameter is specified using `-j` to

be $1+j/16$—by default $j=15$. The progress of the scheme can be watched by using the *report_ops* = *[-Ff] [-D]* A change in the error-norm is reported after every *f* iterations, if `-F` is specified. The retinex image-so-far is dumped out to *file.retinex.dump* every time the error-norm halves, if `-D` is used. This last option allows the user to abort the program without losing the output-so-far.

J.6 Describing Specularities

The program `dodescribe` takes the output of `dotests` (*file.tests*) and creates a description of the specular-blobs:

% `dodescribe` *file.hyster file.iff file.tests file.desc*

The output (*file.desc*), in ASCII is described below in the Backus Naur form:

file.desc	=	size *image_width image_height*
		{*blob_structure*}
blob_structure	=	blob *id evidence n*
		peak *p_col p_row grey_level*
		centroid *c_col c_row*
		angle height width
		{*column row class*}
image_width, image_height	=	integer
evidence	=	"Good"|"Some"|"No"
id, n, column, row	=	integer
p_col, p_row grey-level	=	integer
c_col, c_row, angle, width, height	=	float
class	=	*evidence* |"0"

The first line specifying the image dimensions is followed by a sequence of *blob_structures*. Each of these consists of a unique *id* number, the *evidence* of specularity associated with it, the number of pixels within it: *n*, the image

location of its peak[4] (*p_col,p_row*), its peak *grey_level*, its centroid (*c_col,c_row*), its dimensions *angle*, *column* and *height*. Finally its constituent pixels are listed, giving their image location (*column,row*) and the evidence *class*. The latter is "0" if the pixel is not a spine point, otherwise it is the evidence obtained at it by `dotests`. Thus no information is discarded.

J.7 Edge pruning

The edge segments associated with specularities are removed before they are used for binocular stereo.

```
% prune file.hyster file.desc file.pruned
```

The inputs are *file.desc* created by `dodescribe` and the edge-map *file.hyster*. The output is again in edge file format.

J.8 Binocular stereo

The left and right pruned edge files are rectified prior to matching:

```
% rectify -L filel.pruned filel.rectify
% rectify -R filer.pruned filer.rectify
```

In addition a reference file is created so that the de-rectification can be applied later:

```
% rectify -r -L filel.pruned file.ref
```

---

[4]In case of dispute, the peak nearest the centroid is used.

The PMF edge matcher[5] reads as input the rectified edge files, *filel.rectify*, *filer.rectify* and outputs a description of the matches in ASCII, *file.pmf*.

```
% pmf filel.rectify filer.rectify > file.pmf
```

in ASCII in the Backus Naur form described below:

file.pmf	=	{ *m row column d contrast orientation s* }
m	=	"M"|"U"
row, column, d	=	float
contrast, orientation, s	=	float

Each line in the file corresponds to an edge segment in the left-rectified edge file, at location (*column*, *row*) with its associated *contrast* and *orientation*. Only edge segments with *m*="M" match with ones in the right edge file. The horizontal disparity: *d* is in pixel units. The strength of the match *s* is also provided.

Depth estimates are obtained from the output of `pmf` (*file.pmf*) by triangulation and de-rectification. This is achieved by a single program:

```
% threeD file.ref file.pmf file.3d
```

Where the input file, *file.ref* was created by `rectify`. The output, *file.3d* is in edge-file format. Again the first two fields in the quadruple of each edge segment are the *row* and *column*—in the unrectified left image. However the last two fields are now the *depth* and the disparity *d*. The nearest non-horizontal edge segment to each unmatched specularity in the left image is extracted from *file.3d* as described below.

---

[5]Code supplied by AIVRU [1].

J.9 Matching Specularities

Specular blobs are matched using their descriptions in the left and right images–obtained using `dodescribe`—*filel.desc* and *filer.desc*.

% `matcher` *[weights] [-3 file.3d] [-rr] filel.desc filer.desc file.match*

Where *weights* = *[-Dd] [-Aa] [-Pp] [-Oo] [-Uu]*. Each of the default weights—given in section 6.3.2—can be multiplied by a factor: d for ω_D, a for ω_A, p for ω_P, o for ω_O and u for ω_U. The output, written in ASCII, to *file.match* is simply a set of pairs of numbers. Each pair consists of the id of the specular blob in the left image *left_id*, followed by that of the matching blob in the right *right_id*. If `-3` is specified, a depth estimate *depth* is extracted from *file.3d* for each pair and output with them. The complete Backus Naur form of output—as read by the `tool`—allows three optional fields accompanying each pair:

file.match	=	[# *An_optional_one_line_comment*]
		{ *left_id right_id [d_field [l_field [a_field]]]* }
d_field	=	depth *column row depth*
l_field	=	light *light_id*
a_field	=	answer r_1, r_2
column, row, depth	=	**float**
left_id, right_id, light_id	=	**integer**
r_1, r_2	=	"infinite"|**float**

The *d_field* describes the image position (*column, row*) and *depth* of the reference point. The *l_field* allows a known light source to be assigned to the specularity. The *light_id* is an index to a data-base of light sources, stored in *file.light*—in the format described below. When the principal radii of curvature r_1 and r_2 are known, they can be supplied to the `tool` through the *a_field*. The `tool` makes no use of these values other than to mark the white crosses or boxes on the (r_1, r_2) graph, so that the result can be compared with the known answer.

J.10 Light source data-base

The light source data-base is specified in an ASCII file (*file.light*). The format of this file is:

$$\begin{aligned} \textit{file.light} &= \{\ x\ y\ z\ [\textit{shape}]\ \} \\ x,\ y,\ z &= \text{float} \\ \textit{shape} &= \text{“c” | “o”} \end{aligned}$$

If the light source is compact then *shape*=“c” and if it is circular then *shape*=“o”. The position of the light source in the left camera frame is (x, y, z). Each light source can be accessed from the data-base using its *light_id*. The i^{th} light source in *file.light* has a *light_id*=i. Any *file.light* can be extracted directly from stereo winsom model file, *file.model* using:

```
% modelig < file.model > file.light
```

For a real scene, the file has to be created by hand from known measurements.

J.11 The `tool`

J.11.1 Start up

The interactive `tool` that infers shape-from-specularity runs as follows:

```
% tool [-n] [-m] {dir}
```

Details of its use are given in chapter 7. The `-n` option runs the `tool` without an accompanying `displaytool`—which has its advantages if space is at a premium. The `-m` option starts up the tool in *monocular-mode*—useful when a full stereo pair is not available. Every *dir* argument specifies a Unix directory containing the files required to apply the scheme to a given scene. Each *dir* is put into the `tool`'s library list—for rapid access. The `tool` is started up

with the first in the list. Each directory accessed should contain the following files: `file.light`, `file.match`, `filel.desc`, `filer.desc`, `filel.iff`, `filer.iff`, `filel.pruned`, `filer.pruned`, `filel.hyster`, `filer.hyster`, `filel.edgedisplay`, `filer.edgedisplay`, `filel.blob`, `filer.blob`.
The last four files are used by the `tool` purely for cosmetic reasons: a file: *file.edgedisplay* is generated using:

```
% edgedisplay file.hyster file.desc file.edgedisplay
```

and a *file.blob* by:

```
% showblobs file.iff file.desc file.blob
```

When `-n` is specified it is possible to run the `tool` with a much reduced file set: `file.light`, `file.match`, `filel.desc`, `filer.desc`, `filel.edgedisplay`, `filer.edgedisplay`.
In *monocular-mode* only the following files are needed: `file.light`, `filel.desc`, `filel.iff`, `filel.hyster`, `filel.edgedisplay`, `filel.blob`.

In this case, for stereo-mono compatibility, the `tool` does not distinguish between the root-name `file` or `filel`.

J.11.2 Linkable code

The key computational modules used by the `tool` are available as "C" routine calls. These routines can be called from any other "C" program provided the source files specified below are loaded and included. All these source files are to be found in the directory `toolsrc`. All the routines require the include statements

```
#include <math.h>
#include "vecs.h"
#include "XYZ.h"
```

The maths library must be linked using `-lm` during loading. Also the suite of vector manipulation routines must be linked from the source file `vecs.c`. Each 3-D vector is represented by three consecutive floating point numbers, e.g:

```
float vector[3];
```

For vectors in the local surface frame (l.s.f.), the numbers are, in order, the i^{th}, j^{th} and k^{th} components. For vectors in the left view frame (l.v.f.) they are the the x, y and z components. All vectors are in the l.v.f. unless otherwise stated.

Iterative refinement of the depth data is carried out by the following routine, linkable from file `refine.c`:

```
int iterative_refinement(F, unit_V, S,  V, unit_n, L)
float F[3], unit_V[3], S[3];     /* input  vectors */
float V[3], unit_n[3], L[3];     /* output vectors */
```

As input, `F` is **F**, `unit_V` is $\hat{\mathbf{V}}$ and `S` is **S**. As output, `V` is **V**, `unit_n` is $\hat{\mathbf{n}}$ and `L` is **L**. The value 1 is returned, unless $|\mathbf{V}| < 0.1|\mathbf{F}|$ when 0 is returned, or $\hat{\mathbf{V}} \cdot \hat{\mathbf{n}} \leq 0$ when -1 is returned.

When **V** is known the orientation can be obtained using the following routine, linkable from file `doorient.c`:

```
#include "specstereo.h"

int orient(V, S,  unit_n, L,  slant, tilt)
float V[3], S[3];       /* input  vectors */
float unit_n[3], L[3]; /* output vectors */
float *slant,*tilt;      /* output angles (in degrees) */
```

As input, `V` is **V** and `S` is **S**. As output `unit_n` is $\hat{\mathbf{n}}$ and `L` is **L**. The `slant`, σ and `tilt`, τ (in degrees) are also updated. The value 1 is returned, unless `orient` fails. The return codes defined in `specstereo.h` are `FAILS_V_BEHIND_SURFACE` if the surface is behind the camera, `FAILS_V_TOO_CLOSE` if $|\mathbf{V}| = 0$, `FAILS_L_TOO_CLOSE` if $|\mathbf{L}| = 0$ and `FAILS_NORMALISE` if $|\mathbf{n}| = 0$.

Transformation of unit vectors from the l.v.f. into the l.f.s. is achieved by the routine `Transform`. Before it can be used a `FRAME` structure must be declared and assigned. The structure is defined in file `frame.h` as:

```
typedef struct frame_struct{
  float i[3],j[3],k[3];
}FRAME, *FRAMEP;
```

This represents the unit vectors $\left(\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}\right)$ defined in equation (6.7). The values of these vectors are assigned using the following routine, linkable from file `frame.c`:

```
#include "specstereo.h"
#include "frame.h"

mk_frame(unit_n,unit_V,frame)
float unit_n[3], unit_V[3]; /* input vectors */
FRAME *frame;               /* output frame  */
```

Required as input are $\hat{\mathbf{n}}$, (`unit_n`) and $\hat{\mathbf{V}}$ (`unit_V`). The contents of the `frame` structure are updated. Now this `frame` is used to transform a unit vector (`u`) in the l.v.f. into a unit vector (`v`) in the the l.s.f. via a call to `Transform`—also linkable from `frame.c`:

```
#include "specstereo.h"
#include "frame.h"

Transform(u,v,frame)
float u[3];   /* input  unit vector */
float v[3];   /* output unit vector */
FRAME *frame; /* (input) frame      */
```

The following structure is defined in `monocular.h` to represent a description of a specular blob.

```
typedef struct desc_struct {
    float x,y,theta,h,w,aspect_ratio;
    char shape, state;
    short match_index;
} DESC, *DESCP;
```

The fields `x`, `y`, `theta`, `h`, `w` directly correspond to the parameter of the blob-description defined in section 4.6.3: x_c, y_c, θ, h and w. Four different shapes

of blob are supported. In `monocular.h`, `BLOB_ELONGATED`, `BLOB_CIRCULAR`, `BLOB_NOT_ELONGATED_ELLIPTICAL` and `BLOB_NOT_ELONGATED_IRREGULAR` define valid values of `shape`. If *good* evidence for specularity exists at the blob then `state`='G'. If only *some* evidence exists then `state`='S'. The `match_index` permits matched blobs to be linked.

Monocular analysis of a blob can be achieved by calling the following routine, linkable from files `monocular.c` and `frame.c`:

```
#include "monocular.h"
#include "frame.h"

int monocular_analysis(desc, frame,   angle, err_angle,
          ratio, centre_x,centre_y,focal_length_in_pixels)
DESC *desc;                    /* input blob description */
FRAME *frame;                  /* local surface frame    */
float *angle,*err_angle;       /* output angle (degs)    */
float *ratio;                  /* output  ratio          */
short centre_x,centre_y;       /* input image coords     */
float focal_length_in_pixels; /* input */
```

As well as the description `desc`, some information about the (left) image must be supplied as input: the image coordinate at the optic axis (`centre_x`, `centre_y`) and the `focal_length_in_pixels`. As output, are `ratio`: a rough estimate of $|r_2/r_1|$ and `angle`: an estimate of θ_l in the l.s.f. (in degrees). A rough estimate of the uncertainty in θ_l is provided by `err_angle`. The routine returns one of five values, each defined in `monocular.h`. If `SUCCEED_CIRCULAR` is returned then `ratio`≈ 1 and `angle` is undefined. If `SUCCEED_ANGLE_GOOD` is returned then both `ratio` and `angle` are defined. If `SUCCEED_ANGLE_BAD` is returned then `ratio` is unreliable and `angle` is undefined. If `FAIL_BLOB_SHAPE` is returned then `monocular_analysis` failed because the blob shape was irregular. If `FAILS_V1` is returned then `monocular_analysis` failed due to unfavourable circumstances.

To use the routine `stereo_analysis` the following files must be linked—in addition to those stipulated above: `frame.c`, `vecerr.c` and `specstereo.c`.

```
#include "specstereo.h"
#include "frame.h"

int stereo_analysis(r, err_r,  dN, err_dN, V, err_V,
                B_R, err_B_R,  S, err_S,  b, theta_v,  frame)
float r[3], err_r[3],  dN[3], err_dN[3]; /* output params */
float V[3], err_V[3],   S[3],  err_S[3]; /* input  params */
float unit_W_R[3],      err_unit_W_R[3]; /* input  params */
float b,theta_v;                         /* view   params */
FRAME *frame;                            /* frame (input) */
```

Three inputs: `V`, `S` and `unit_W_R` represent respectively the vectors $\mathbf{V}$, $\mathbf{S}$ and $\widehat{\mathbf{W}_\mathbf{r}}$. Two outputs: `r` and `dN` represent respectively the vectors $\mathbf{r}$ and $\delta\boldsymbol{n}$. The uncertainty associated with each vector is accessed through the parameters with the prefix `err_`. Also required as input are `b`: the base-line separation in metres, `theta_v`: the vergence angle θ_v and `frame`: the local coordinate frame—as created by `mk_frame`. When successful, `stereo_analysis` returns 1. If it fails the return codes (defined in `specstereo.h`) provide the reason why: `FAILS_N` indicates a bad surface normal, `FAILS_WR` indicates a bad $\widehat{\mathbf{W}_\mathbf{R}}$, `FAILS_WZ` indicates that $|\mathbf{W}|=0$ and `FAILS_N1` indicates a bad value of $\hat{\mathbf{n}}$ occurred.

Once obtained the values of vectors $\mathbf{r}$ and $\delta\boldsymbol{n}$ can be used to apply the six applications—listed in table 6–2. The routines for these applications, below, all take as input: `r` and `dN`. Again the prefix `err_` denotes a corresponding uncertainty. For all these routines the file `vecerr.c` must be linked.

```
#include "specstereo.h"

int do_cylindrical_case(r,err_r, dN,err_dN,
                                      r_c,err_r_c, theta_c)
float r[3],err_r[3], dN[3],err_dN[3]; /* input vectors    */
float *r_c, *err_r_c;                 /* output radius(m) */
float *theta_c;                       /* output dir(degs) */
```

This routine resides in file `cyl.c` and outputs the cylindrical parameters: r_c (`r_c`) and θ_c (`theta_c`). A value of 1 is returned unless the $r_c \approx \infty$ when `SPECIAL_CASE_INFINITE`, defined in `specstereo.h`, is returned.

```
#include "specstereo.h"

do_hyp_case(r,err_r, dN,err_dN, r_h,err_r_h, theta_h)
float r[3],err_r[3], dN[3],err_dN[3]; /* input vectors    */
float *r_h, *err_r_h;                 /* output radius(m) */
float *theta_h;                       /* output dir(degs) */
```

This routine resides in file hyp.c and outputs the hyperbolic ($r_1 = -r_2$) parameters: r_h (r_h) and θ_h (theta_h). A value of 1 is returned unless $r_h \approx \infty$ when the value is SPECIAL_CASE_INFINITE.

```
int get_Hxx_Hxy(r,err_r,dN,err_dN, Hxx,err_Hxx, Hxy,err_Hxy)
float r,err_r,   dN,err_dN;              /* input  params */
float *Hxx,*err_Hxx, *Hxy,*err_Hxy;      /* output params */
```

This routine resides in file geth.c and estimates the values and uncertainities of H_{xx} (Hxx) and H_{xy} (Hxy). These values can then be used to determine the constraint graph. The value 1 is returned unless division by zero has occurred when the value is 0.

```
#include "specstereo.h"

get_radii(code_rg,rg,err_rg, code_rl,rl,err_rl,
                        r,err_r, dN,err_dN, angle,err_angle)
int   *code_rg, *code_rl;              /* output codes    */
float *rg,*err_rg,  *rl,*err_rl;       /* output params   */
float r[3],err_r[3],  dN[3],err_dN[3]; /* input  params   */
float theta_l,  err_theta_l;           /* input dir (degs)*/
```

This routine resides in file getradii.c and estimates the values and uncertainties of the greatest and least radii of curvature: r_{gr} (rg) and r_{le} (rl). The direction of least curvature in the l.s.f. θ_{le} (theta_l) must be supplied. The output codes code_rg and code_rl indicate whether the estimates rg and rl, respectively, are good. The following values, defined in specstereo.h are possible: GOOD_RADIUS if the estimate is good, NEG_INFINITE_RADIUS if the radius$=-\infty$, POS_INFINITE_RADIUS if the radius$=\infty$ and UNDETERMINED_RADIUS if the radius could not be determined.

```
#include "specstereo.h"

int non_spherical(r,err_r, dN,err_dN,  R_i_j_zero)
float r[3], err_r[3],  dN[3], err_dN[3]; /* input params */
char R_i_j_zero;                         /* input param  */
```

This routine resides in file spheretest.c and applies the non-spherical test. The additional input parameter R_1_2_zero ¡ should only be non-zero if $R_1 \approx 0$ or $R_2 \approx 0$. The return values defined in specstereo.h indicate the result: NOT_A_SPHERE, COULD_BE_SPHERE and CANT_APPLY_SPHERE_TEST. If the value is not NOT_A_SPHERE then the spherical interpretation, below can be made.

```
#include "specstereo.h"

int do_spherical_case(r,err_r, dN,err_dN,  r_s,err_r_s)
float r[3],err_r[3], dN[3],err_dN[3]; /* input vectors    */
float *r_s, *err_r_s;                 /* output radius(m) */
```

This routine resides in file sph.c and outputs the spherical parameter: r_s (r_s). A value of 1 is returned unless the $r_s \approx \infty$ when the value is SPECIAL_CASE_INFINITE.

Appendix K

Published paper: Computing lightness

This appendix consists of paper [22] published in Pattern Recognition Letters 5, in February 1987. It gives details of the implementation of the lightness computation used by the retinex based test in chapter 4.

Pattern Recognition Letters 5 (1987) 129–138
North-Holland

February 1987

Computing lightness

Gavin BRELSTAFF and Andrew BLAKE
Department of Computer Science, University of Edinburgh, Scotland

Received October 1985
Revised 3 April 1986

Abstract: Retinex theory predicts perceived colour. Each of its three retinex systems determines a field of lightness that directly corresponds to surface reflectance, for the class of images called Mondrians. We describe in detail and demonstrate a complete, noise-resistant computational scheme for the retinex. A local, parallel implementation is proposed. We investigate how the scheme might provide an image segmentation based on surface materials in the real world.

Key words: Colour perception, retinex theory.

1. Introduction

The phenomenon of colour constancy demonstrates that perceived colour of a point in a scene is not entirely determined by the spectral content of the light received locally on the retina. There is an additional global effect.

Land's retinex theory (Land and McCann, 1971; Land, 1983) impressively predicts perceived colour. The retinex system that it employs, effectively converts a field of monochromatic intensities into a field of lightnesses. The monochromatic scheme may be applied in each of the 3 colour channels to predict perceived colour. A single channel retinex is useful for estimating monochromatic surface reflectance.

The first computational scheme of the retinex (Land and McCann, 1971) was developed to work on images called Mondrians where lightness corresponds to relative reflectance of the surface being imaged.

Horn (1974) recognised that this method of recovering reflectance, might produce useful surface descriptions. He developed a local, parallel scheme to compute lightness, that was considered biologically feasible. However the scheme was incomplete, and Blake (1985a, b) modified it, by imposing formal boundary conditions at the image perimeter, to produce a complete, robust, local, parallel retinex scheme. This paper adapts this scheme for conventional images and demonstrates using a serial implementation that it can recover reflectance of Mondrians. For this we use an iterative method (successive over-relaxation).

A study of the effect of noise reveals that the initial stage of the scheme is in fact a very primitive edge detector. We describe how the Canny operator – a sophisticated edge detector – is integrated into a noise-resistant scheme.

We find that lightness[1] does not generally correspond to reflectance in real images. However it remains useful for predicting perceived colour and can sometimes provide an image segmentation based on surface material. This is illustrated using

We are grateful to IBM and SERC for a CASE studentship for G. Brelstaff, to SERC for support under grant GR/D 1439.6 and to the Royal Society for the IBM Research Fellowship for A. Blake. Thanks are due also to A. Zisserman and J. Dow for invaluable assistance.

[1] This is lightness as specified by Land's designators (Land, 1983), not as in hue, saturation and lightness colour mixing systems.

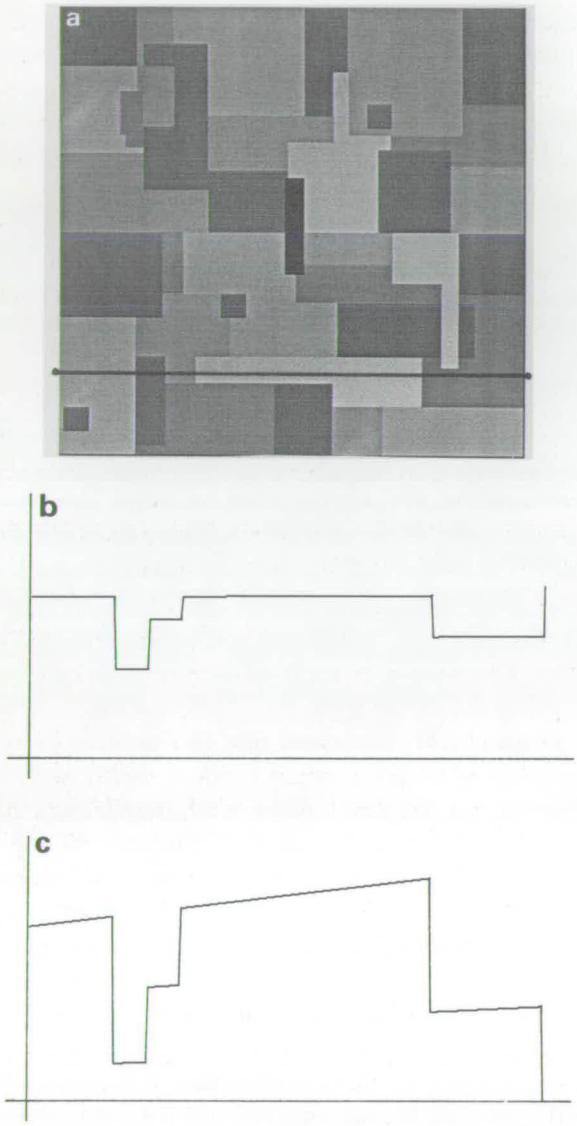


Figure 1. A Mondrian image. (a) has intensity profile (c) along the line marked on it, (b) is the computed lightness.

the lightness fields computed from some realistic images. We briefly outline a possible 'real-time' parallel implementation.

2. Lightness computation

Blake's reformulated lightness computation (Blake, 1985a, b) imposes formal boundary conditions, so it can be applied to any Mondrian. He showed it to be a robust, locally parallel equivalent to Land's retinex. The 2D gradient operator ∇ is applied to log-intensity $p(x)$ to detect edges. A threshold operator T is defined for vector fields:

$$T(A) = t(|A|^2 - \lambda^2)A \tag{1}$$

where t is a smooth approximation to a step function, as in Figure 2, and λ is a threshold. The formal statement of Land and McCann's scheme is that lightness is

$$l(x, y) = \int_P \boldsymbol{E} \cdot \mathrm{d}\boldsymbol{r},$$

where $\boldsymbol{E} = T(\nabla \mathbf{p})$, arbitarily assigning l at some point (x_0, y_0) and then integrating directly over paths P from (x_0, y_0) to each point (x, y). Blake has shown that this computation can be performed locally in parallel by minimising the functional

$$F = \int_{\text{image}} (\nabla l - \boldsymbol{E})^2, \tag{2}$$

which is equivalent to solving

$$\nabla^2 l = \nabla \cdot \boldsymbol{E}, \tag{3}$$

under the boundary condition that $\boldsymbol{n} \cdot \nabla l = \boldsymbol{n} \cdot \boldsymbol{E}$ on the contour C that surrounds the image. (Vector $\boldsymbol{n}$ is the unit normal to C.)

3. A modified scheme for real images

We have modified Blake's scheme for use on real images. The main considerations in doing this were:

- real images conventionally have a rectangular picture structure, not hexagonal.
- real images contain noise.
- real images often break the rules of the Mondrian world (see Section 3.3).

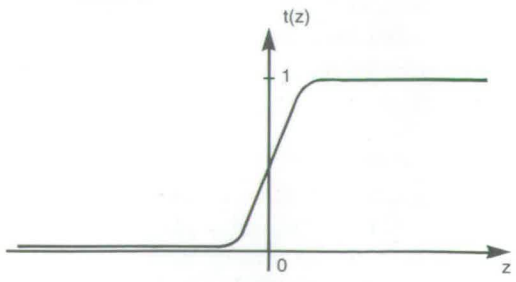


Figure 2. Blake's threshold function as in equation (1).

3.1. *Rectangular grid images*

Blake's formulation uses equilateral triangular elements, sampling intensities at the nodes of a hexagonal grid. However, conventional images produced by cameras, sample intensity at the nodes of a rectangular grid, having a standard aspect ratio A. We use right-angled triangular elements in a new tesselation Z – so that adjacent pairs of triangles cover a rectangular area – as shown in Figure 3. The set of vertices V of the triangles in Z is a rectangular grid of nodes, at which the intensities are sampled. This is compatible with conventional cameras and displays.

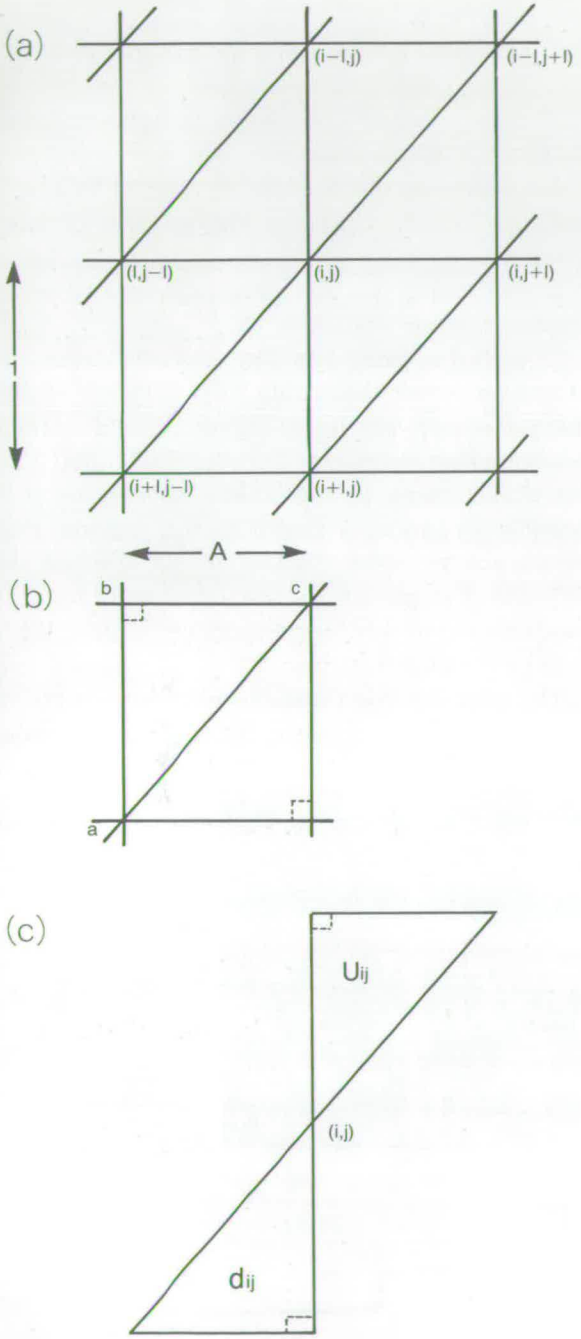


Figure 3. The image is discretely sampled at each node (i, j) of a rectangular grid. (a) Triangular linear elements are used to estimate intensity gradients. (b) Each element is labelled by a boolean flag (c).

Each vertex is indexed by coordinate (i, j) as in Figure 3 and has a set $N(i, j)$ of up to 4 (fewer on the boundary) nearest neighbouring vertices. At each vertex, p_{ij} denotes the log-intensity and $l_{i,j}$ the estimate of lightness.

The scalar fields are modelled using linear finite elements in the form of the right-angled triangles, so the gradient of a field $\boldsymbol{B}$ within an element having values a, b and c at its vertices, and aspect ratio A as shown in Figure 3, obeys:

$$|\nabla \boldsymbol{B}|^2 = g(a, b, c) = (a-b)^2 + (b-c)^2/A^2.$$

The function g is used to estimate $|\nabla p|$ within each triangle. So the edge detection stage consists of thresholding this value in the triangles above and below each vertex – i.e. by assigning the arrays of boolean flags $u_{i,j}$ and $d_{i,j}$ as below (see Figure 3).

$$\begin{aligned}&\text{For } (i,j) \in U = \{(i,j)\text{: the triangle } \textit{above}\ (i,j) \text{ is in } Z\},\ u_{i,j} = 1 \text{ if } g(p_{i,j}, p_{i-1,j}, p_{i-1,j+1}) > \lambda^2,\\ &\qquad\qquad 0 \text{ otherwise.}\\ &\text{For } (i,j) \in D = \{(i,j)\text{: the triangle } \textit{below}\ (i,j) \text{ is in } Z\},\ d_{i,j} = 1 \text{ if } g(p_{i,j}, p_{i+1,j}, p_{i+1,j-1}) > \lambda^2,\\ &\qquad\qquad 0 \text{ otherwise.}\end{aligned} \tag{4}$$

For convenience, we will define $u_{i,j} = 0$, for $(i,j) \notin U$, and $d_{i,j} = 0$, for $(i,j) \notin D$.

Using these flags the discrete version of the functional (F in equation (2)) can be expressed in terms of the function g.

$$\begin{aligned}F = \sum_{i,j \in U} & g(l_{i,j} - u_{i,j}p_{i,j}, l_{i-1,j} - u_{i,j}p_{i-1,j},\\ & l_{i-1,j+1} - u_{i,j}p_{i-1,j+1})\\ + \sum_{i,j \in D} & g(l_{i,j} - d_{i,j}p_{i,j}, l_{i+1,j} - d_{i,j}p_{i+1,j},\\ & l_{i+1,j-1} - d_{i,j}p_{i+1,j-1}),\end{aligned}$$

which can be minimised by differentiating with respect to each $l_{i,j}$ by solving the sparse set of

simultaneous equations $\partial F/\partial l_{i,j}=0$, which are explicitly written below.

$$\forall i,j \in V, \quad \sum_{k,m \in N(i,j)} c_{k,m} f_{i,j,k,m}(l_{i,j}-l_{k,m}) = r_{i,j}$$

where $f_{i,j,k,m}=0,1$ or 2 is the number of triangles in the tesselation on which both (i,j) and (k,m) are vertices,
where N, V are as defined above, and
where $c_{k,m}=1/A^2$ if $k=i$, 1 otherwise, and (5)

$$\begin{aligned} r_{i,j} = {} & (u_{i,j}+d_{i-1,j})(p_{i,j}-p_{i-1,j}) \\ & +(d_{i,j}+u_{i+1,j})(p_{i,j}-p_{i+1,j}) \\ & +((u_{i+1,j-1}+d_{i-1,j})(p_{i,j}-p_{i,j-1}) \\ & +(u_{i+1,j}+d_{i-1,j+1})(p_{i,j}-p_{i,j+1}))/A^2. \end{aligned}$$

Simultaneous over-relaxation (JOR) provides an iterative method to solve these equations, as shown below. Convergence is ensured by restricting the relaxation parameter ω so that $0<\omega<1$ (Blake, 1983).

$$\begin{aligned} l_{i,j}^{(n)} = {} & (1+\omega) l_{i,j}^{(n-1)} \\ & +\omega\left(\sum_{k,m \in N(i,j)} c_{k,m} f_{i,j,k,m} l_{k,m}^{(n-1)} + r_{i,j}\right) \Big/ M_{i,j}, \end{aligned}$$

where $M_{i,j}=\sum_{k,m \in N(i,j)} c_{k,m} f_{i,j,k,m}$. (6)

Here $M_{i,j}$ is a normalising constant for the contribution to the sum from each vertex, that takes into account how many of its surrounding triangles are in the tesselation Z. The log-intensities were used as the initial estimates of lightness: $l^{(0)}=p$.

3.2. *Combating noise using the Canny operator*

There are two distinct stages to lightness computation in Mondrians:

1. identifying the borders of patches by thresholding. This is really a primitive form of edge detection.

2. minimising the functional, constructed using the edges found (equation (2)).

In principle noise can disrupt both stages. In practice however the second stage is robust to the fluctuations in the initial lightness field – their random effects averaging out when integrated over the whole image. In fact at low noise levels it is possible to choose a threshold that ignores gradients due to noise, finding only those due to reflectance steps. So the minimisation produces patches of uniform lightness values almost identical to the noiseless case.

The edge detection stage is more susceptible to noise. The random intensity fluctuations can make

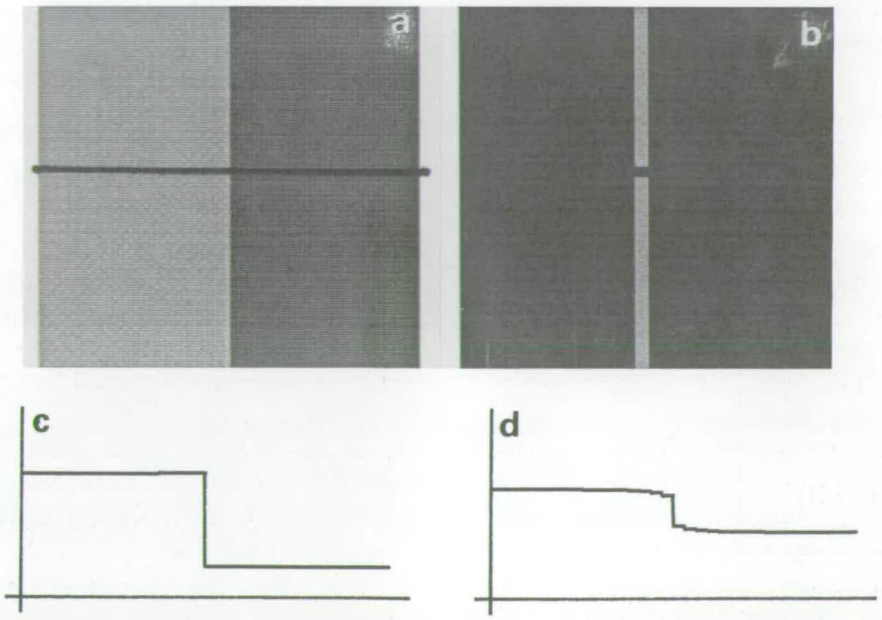


Figure 4. A gap cause a local distortions in the lightness field. (a) is the intensity image. (b) is the edge map showing the gap, (c) and (d) are profiles of intensity and computed lightness along the line marked in (a) passing through the gap.

impossible to choose one threshold that suppresses noise within a patch while retaining the atch border. A gap in an otherwise complete edge Figure 4), causes distortion in its vicinity. In the ext sections an improved edge detector is used to ombat the effects of noise that we have discussed ere.

.2.1. *Criteria for edge detection*

The discussion above suggests the following riteria for the edge detector used by the scheme. hey are in order of importance:

1. patch boundaries should be detected as closed ontours.
2. boundaries should be accurately located.
3. spurious edge elements should be avoided.

he primitive edge detector so far described decides hether each triangular element is an edge, using nly the three intensities sampled as its vertices. In ffect signal to noise ratio (SNR) is improved by sing an operator with a larger support as an estiate of $|\nabla p|$, for use in place of $g(\cdots)$ in equation 4). The increased support reduces the accuracy ith which each edge element is localised but tends o avoid gaps in edges especially when hysteresis is sed.

.2.2. *The Canny operator*

The Canny operator (Canny, 1983) is an edge letector designed to optimise the product of SNR nd edge localisation. It applies a derivative of a D gaussian: ∇G. Setting the width σ of the gaussian determines the SNR-localisation trade off. In n adaptive scheme the amount of noise in an mage might be estimated by its RMS value and sed to determine σ. In our discrete scheme the perator is applied by locally convolving the image ith masks that approximate two orthogonal components of ∇G. At each sample point an edge trength is delivered. Local maxima of strength epresent potential edge points. The operator also roduces an estimated normal direction to the otential contour, on which each edge point would ie. To ensure only one response per edge point, ny strength that is not a local maximum along its ormal direction is suppressed.

The Canny operator was incorporated, to improve edge detection by applying it to the (log-intensity) image beforehand – producing at each vertex a boolean flag $C_{i,j}$, that is set on, only if there is an edge in the rectangular element above and to the right of the vertex. Parameters including σ determine the number and distribution of these edges in this case the basic scheme uses the following assignments (using the same notation as equation (4)):

$$\begin{aligned} &\text{For } (i,j)\in U, \quad u_{i,j}=C_{i,j}. \\ &\text{For } (i,j)\in D, \quad d_{i,j}=C_{i+1,j-1}. \end{aligned} \tag{7}$$

An example lightness field produced by a noisy Mondrian by employing the Canny operator is shown in Section 4.2.

3.3. *Real scenes*

In real scenes, surfaces can be curved, creased, glossy, textured, lie at different depths, have non-uniform reflectance, and be self-luminous. Although radiated flux from illuminants usually forms a smooth, continuous field at any given place (hard-edged spot-lights are an exception), its interaction with some orientations of surface causes large discontinuities in illumination. So in general (unlike for Mondrians) the lightness field is not a reliable estimate of reflectance although it still predicts perceived colour. Below we discuss how useful is the description provided by a lightness field for a variety of real world features.

3.3.1. *Real world features*

Shading is caused by small orientation variations. Gentle shading results in small intensity gradients. These can be ignored by using a suitable threshold and so produce a uniform region of computed lightness values (as in Figure 5). This usefully represents a single patch of material. However steeper shading gradients can make it impossible to choose a threshold that distinguishes the now larger intensity gradients from material boundaries. The resulting lightness distribution is less useful – it may be non-uniform over a patch of uniform reflectance.

Surface orientation discontinuities (e.g. creases)

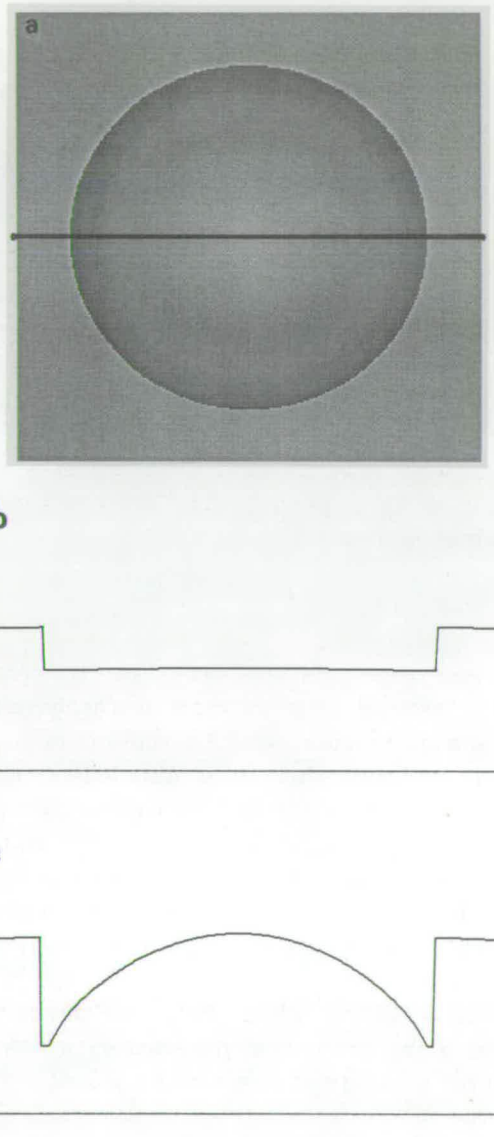


Figure 5. (a) is shaded matte sphere, (c) and (b) are profiles of intensity and computed lightness along the line marked in (a).

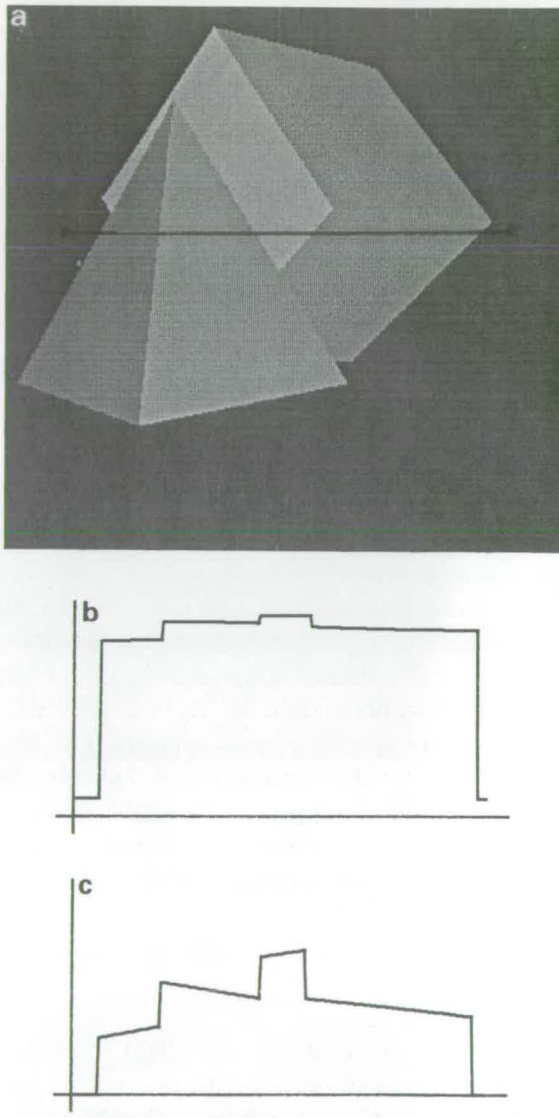


Figure 6. (a) is a pyramid in front of a cube, containing bot orientation and depth discontinuities, (c) and (b) are profiles intensity and computed lightness along the line marked in (a

cause large intensity gradients that thresholding cannot distinguish from material boundaries. The lightness field computed across the discontinuity will change – falsely indicating a material boundary. The edge between the faces of the pyramid in Figure 6 demonstrates this.

Depth discontinuities. Under normal viewing conditions depth variations produce only small intensity gradients. These can be successfully distinguished from material boundaries in the way gentle shading was. However, they are often accompanied by an orientation discontinuity – e.g the edge between the cube and pyramid in Figur 6. In this case the edge falsely indicates a materia boundary as discussed above, unless, as is ofte the case, it is a real one.

Shadows. At a sharp shadow border the sam analysis holds as for a crease – it is misclassified a a material boundary. At a very gentle borde (penumbra) the illumination gradient is small an can be ignored (as for gentle shading), to usefull indicate that it is not a material boundary. Figur

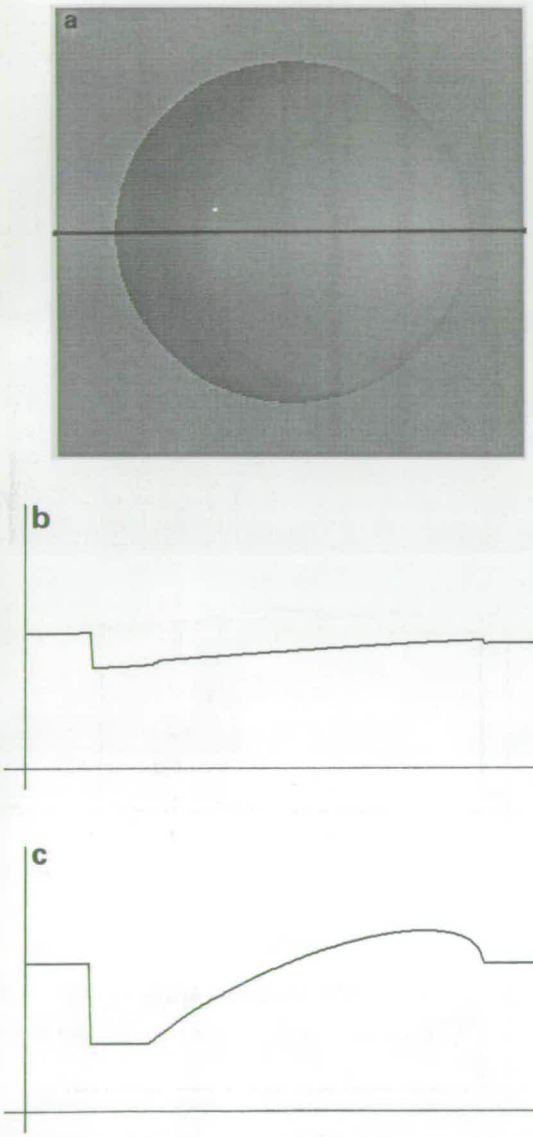


Figure 7. (a) is matte sphere with a shadow region, (c) and (b) are profiles of intensity and computed lightness along the line marked in (a).

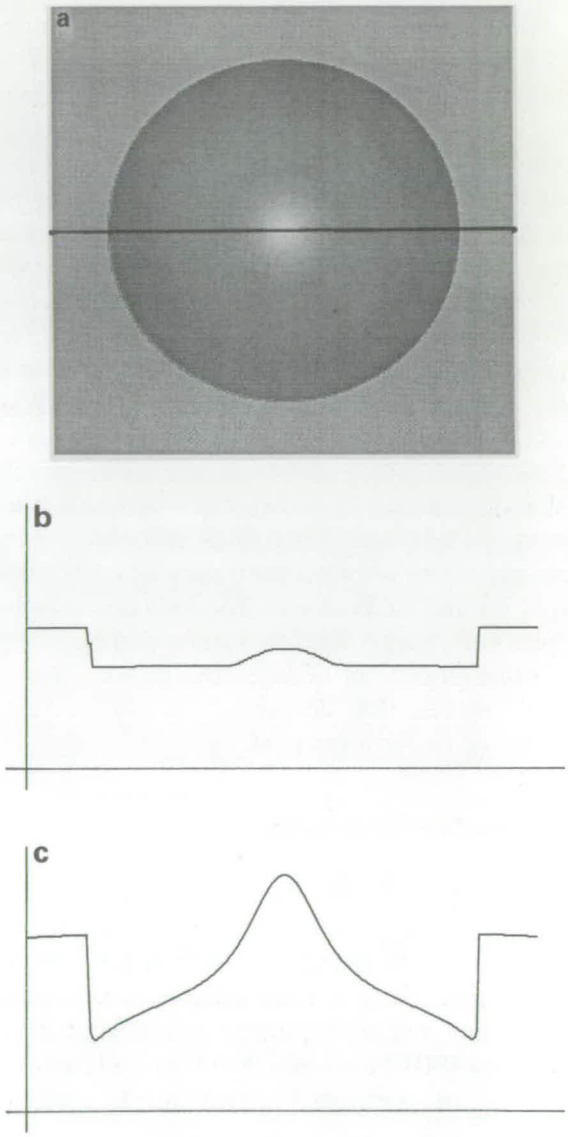


Figure 8. (a) is glossy sphere with a specularity, (c) and (b) are profiles of intensity and computed lightness along the line marked in (a).

7 represents a case in between these two extremes, where the threshold used detected several edges across the border. The resulting lightness field varies smoothly so it does not usefully indicate the underlying uniform material.

Light sources and fluorescent materials are self-luminous surfaces, having large intensity gradients at their borders, which are material boundaries. Within their image region a uniform lightness field is computed, producing a useful segmentation in terms of surface material.

Specularities are caused where a glossy surface is oriented to reflect light as a mirror does. The nature of the intensity gradients at their borders depend on the amount of gloss and the local surface orientation. They are not caused by material boundaries. A specularity with a sharp border is treated as if it were a light source and erroneously segmented from the material that surrounds it. Figure 8 shows a specularity with a gentler profile. Two regions of uniform lightness that wrongly represent different materials are apparent. (The

smooth variation in the lightness field at the border is for similar reasons to that for the shadow.)

Texture, is caused by local fluctuations of surface orientation or reflectance. The intensity gradients produced by it can be treated in the same way that noise was in Section 3.2. So in a region where these gradients are small enough to be distinguished from material boundaries, an average uniform lightness field can be computed, to correctly indicate a patch of a single material. For materials where texture is very prominent, the lightness field becomes distorted by the presence of extra edges and gaps – like the Mondrian in noise. Texture fluctuations persist in the lightness distribution resulting in a non-uniformity which no longer usefully describes a single material. An intermediate case is shown in Figure 9 which is a textured version of Figure 6. The lightness profiles across both images are very similar despite a considerable amount of texture. The regions of uniform lightness that do persist, usefully indicate regions of uniform material.

3.4. *Serial implementation*

Ideally, to fully exploit our scheme, it should be implemented in parallel using special hardware – one possibility is described in Section 3.4.2. However, a serial computer (an Edinburgh University APM) was used to implement a successive over-relaxation (SOR) method of solution – a serial version of JOR, mentioned in Section 3.1.

3.4.1. Basic scheme

The program executed the following steps:

1. Convert (1 byte) input intensities into a (4 byte) integer representing the log-intensity.

2. Assign the values of $M_{i,j}$, by examining the number of triangles surrounding each vertex (i, j) – as in equation (6).

3. Use the threshold operation to assign the boolean flags $u_{i,j}$ and $d_{i,j}$ – as in equation (4) or using the Canny operator (see (7)).

4. Use flags $u_{i,j}$ and $d_{i,j}$ with the log-intensities to compute $r_{i,j}$ as (4 byte) integers – as in equation (5).

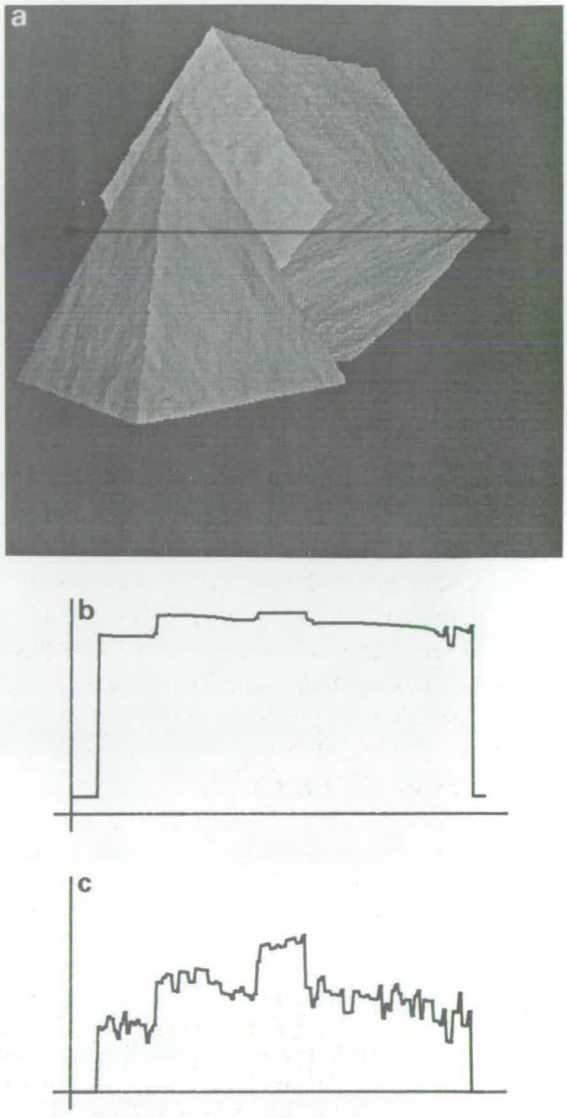


Figure 9. (a) is a texture version of the scene in Figure 6, (c) and (b) are profiles of intensity and computed lightness along the line marked in (a).

5. Use the log-intensities as the first guess at the lightnesses $l^{(n)}$, where $n=0$.

6. Use the formula given by equation (6) to compute the value of $l^{(n)}$ given $l^{(n-1)}$.

7. Repeat step 6 until the maximum individual change in lightness over the whole array is within a preset tolerance, or up to a maximum number of iterations.

Note: in step 6, there is no need for two lightness arrays, ($l^{(n-1)}$ and $l^{(n)}$), because the SOR scheme replaces $l_{i,j}^{(n-1)}$ with $l_{i,j}^{(n)}$ as soon as it becomes avail-

ible. The implementation required 68 bits per pixel.

3.4.2. A possible parallel implementation

Special hardware such as the CLIP4 image processing system (Duff, 1978) could be used for a 'real time', locally parallel implementation of our lightness computation scheme. A conservative estimate of the run time, calculated by considering the number and type of operation required to execute step 6 in the basic scheme (see Section 3.4.1), is 3 ms, using 16 bit precision. (Assuming CLIP specification as originally published and with additional storage (128 rather than 32 plane)). Such an implementation directly driven by a camera would provide an excellent test of the scheme on real images. It would be interesting to see whether the computation is in fact coordinate-frame-invariant as claimed in (Blake, 1985a, b), by rotating the image in real time and observing whether the lightness is altered. Note that speed increases could probably be achieved by multi-resolution techniques (Terzopoulos, 1984).

4. Results

4.1. Mondrians

Figure 1 shows the scheme has successfully computed lightness for a (256×256) Mondrian image. The illumination varies linearly from 100% at the top left of the image, down to 50% at the lower right, causing small intensity variations. These are not easily seen, but are evident in the intensity profile. Using a threshold equivalent to a 6% intensity ratio the computed lightness field consists entirely of uniform areas, corresponding to uniform reflectance of each Mondrian patch. The lightness profile demonstrates that there is a global correspondence. The two lightest regions in the profile that are disconnected and assigned an equal lightness, are in fact of equal reflectance. A satisfactory rate of convergence was achieved using $\omega \simeq 0.5$.

4.2. Noise

Figure 10 shows use of the Canny operator to combat the effect of noise on the lightness computation. The image is a Mondrian (having a minimum step between patches of 40 grey-levels) with added gaussian noise of RMS value 20 grey-levels.

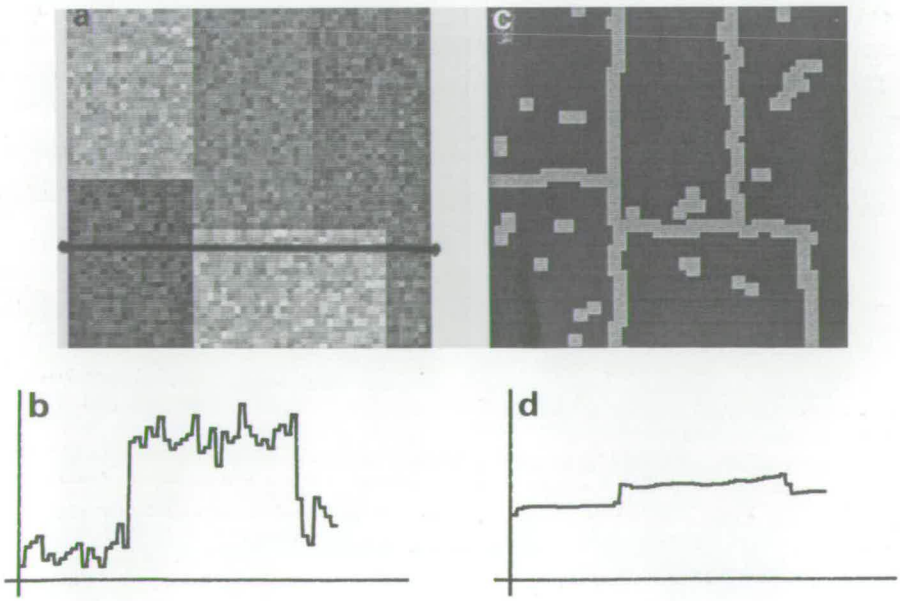


Figure 10. Using the Canny operator on the noisy Mondrian (a) produces edge map (b), (c) and (d) are profiles of intensity and computed lightness along the line marked in (a).

The edge map obtained as stated in Section 3.2.2, using a σ value of 2 pixels, is a significant improvement on that obtained by the simple threshold method. The resulting lightness field is remarkably uniform within patches. The profile through it, shows that the large intensity variations have been averaged out by the minimisation stage as mentioned in Section 3.2. This is an advantage when lightness is being interpreted as reflectance. However observed noise features should be retained if the goal is to predict perceived colour.

4.3. Features of real images

Section 3.3 discusses how the lightness computed in images containing real world features, not found in the Mondrian world, can be used to both predict perceived colour and to segment images in terms of materials. Figures 5 to 9 illustrate this discussion.

References

Blake, A. (1983). *Parallel Computation in Low-Level Visi* Ph.D. thesis, CST-27-83, Dept. of Computer Science, Ed burgh University.

Blake, A. (1985a). Boundary conditions for lightness compu tion in Mondrian world. *Comput. Vision Graphics Ima Processing* 32, 314–327.

Blake, A. (1985b). On lightness computation in the Mondri world. *Proceedings of the Wenner–Gren Conference on C tral & Peripheral Mechanisms in Colour Vision*. MacMill New York, 45–59.

Canny, J.F. (1983). *Finding Edges and Lines in Images*. Mast Thesis, AI-TR 720, A.I. Lab., M.I.T.

Duff, M.J.B. (1978). Review of the CLIP. *National Compu Conference* 1978, 1011–1060.

Horn, B.K.P. (1974). Determining lightness from an ima *Comput. Graphics Image Processing* 3, 277–299.

Land, E.H. and J.J. McCann (1971). Lightness and reti theory. *J. Opt. Soc. Am.* 61(1), 1–11.

Land, E.H. (1983). Recent advances in retinex theory and so implications for cortical computation: Color vision a natural images. *Proc. Natl. Acad. Sci. U.S.A.* 80, 5163–51

Terzopoulos, D. (1984). Efficient multiresolution algorith for computing lightness, shape-from-shading, and opti flow. *Proc. A.A.A.I.* 1984, 314–317.