Specularity Removal for Shape-from-Shading

by

Frank C.H. Tong

B.Sc., Chinese University of Hong Kong, 1983

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE
in the School
of
Computing Science

SIMON FRASER UNIVERSITY
December 1987

Approval

Name: Frank C.H. Tong

Degree: Master of Science

Title of Thesis: Specularity Removal for Shape-from-Shading

Binay Bhattacharya
Chairman

Brian V. Funt
Senior Supervisor

Thomas W. Calvert

James P. Delgrande

Diane V. Ingraham
External Examiner

October 22, 1987
Date Approved

PARTIAL COPYRIGHT LICENSE

I hereby grant to Simon Fraser University the right to lend my thesis, project or extended essay (the title of which is shown below) to users of the Simon Fraser University Library, and to make partial or single copies only for such users or in response to a request from the library of any other university, or other educational institution, on its own behalf or for one of its users. I further agree that permission for multiple copying of this work for scholarly purposes may be granted by me or the Dean of Graduate Studies. It is understood that copying or publication of this work for financial gain shall not be allowed without my written permission.

Title of Thesis/~~Project~~/~~Extended Essay~~

Specularity Removal for Shape-from-Shading

Author:

(signature)

Frank Tong

(name)

Dec. 14, 87

(date)

Abstract

Specularity reflecting surfaces confuse traditional shape-from-shading algorithms because the variation in image intensity within a specularity does not directly relate to the cosine of the incident angle, as it would for a simple Lambertian reflector. To overcome this problem, color is introduced and a method of removing the specular component of the intensity variation is proposed based on a dichromatic model of surface reflection. Unlike Shafer's method for specularity removal, which is restricted to uniformly colored surface patches, our algorithm uses information from several differently colored regions. The problem of segmenting an image into color regions is successfully avoided as the specular component is calculated and removed using local operations only. The image resulting from specularity removal preserves the relative intensity of the diffuse component so it can then be input to a shape-from-shading algorithm. Our shape-from-shading algorithm is based on variational calculus. Without assuming the location of the scene illuminant, and allowing background illumination, the algorithm computes the shape from the diffuse component image in a more general setting than the existing algorithms do. In the thesis, the algorithm is formulated into a local relaxation scheme which allows a parallel network implementation.

To my parents.

Acknowledgements

I would like to express my deepest gratitude to my supervisor, Dr. Brian V. Funt for his valuable guidance and insightful suggestions. His patience in reading and correcting my thesis is appreciated very much. I should also be thankful to the members of my supervisory committee, Dr. Thomas W. Calvert and Dr. James P. Delgrande, whose constructive comments are indispensible to the completion of this thesis. My gratitude is also due to my external examinar, Dr. Diane V. Ingraham. Her encouragement and helpful suggestions were very incisive.

My gratefulness is also given to all my nice friends in the department who helped contributing to the stimulating environment for carrying out this research. I have to thank my very best friend, Mimi Kao, for her moral support which I most needed. I am grateful to my parents for their upbringing. I would like to dedicate to them whatever I have achieved in this research.

Finally, my grateful acknowledgement go to NSERC and Simon Fraser University for the financial support.

Table of Contents

Lists of Figures

Chapter 1
Introduction

1.1. Motivation

When light strikes a surface, it is reflected both diffusely and specularly. Surfaces facing the light source reflect the most diffuse light while surfaces oriented half-way between the source and the viewer exhibit the strongest specular reflection. Shading is the phenomenon of intensity variation in the image due to variation in surface orientation in the scene. Comprised of both diffuse and specular components, the intensity variation carries a lot of information about surface orientation.

Shading therefore provides rich cues about surface orientation. But shape-from-shading is a hard problem. The difficulty lies in the fact that the intrinsic characteristics -- for instance, the orientation, illumination and reflectance -- are all encoded in a single intensity value. As Barrow and Tenenbaum [BaTe78] have commented, while the encoding process gives unique value, the decoding is ambiguous. A single intensity value may result from an infinite number of combinations of illumination, orientation and reflectance.

Horn and Brooks have derived a numerical scheme [HoBr86,BrHo85] for computing local surface orientation from the image intensity. But to render the problem solvable, some

assumptions about the world were adopted. Smooth surfaces[1] of uniform reflectance[2], illuminated by a single point source at great distance were used to constrain the problem. Horn looked at the world of perfect diffuse reflectors[3], devoid of highlights, so that the reflectivity of the surfaces is kept simple enough to be modeled by a very simple function. In addition, color is excluded from consideration so that reflectance variations can be avoided. The location of the distant point source is usually assumed to be known. In his work [BrHo85], Horn relaxed this assumption to an unknown distant point source and derived a scheme which computes both the local surface orientation and the light source. However, evaluation of the light source requires global scans of the image data.

While Horn's method is sufficient to solve problems involving monochromatic Lambertian reflectors, it will not cope with the real world. In our daily life, we see color everywhere, also, it is common to find highlights reflected off smooth surfaces. In this thesis, a wider scope of shape-from-shading problems is investigated.

Extending beyond Horn's world of Lambertian reflectors, we work on the more realistic situation where, besides the diffuse component, a specular and an ambient component are reflected off the surfaces as well. The diffuse component is reflected by the pigment particles embedded in the surface. It exhibits a specific color depending on the chromaticities of the surface pigment and the scene illuminant. The diffuse component is supposed to be scattered equally in all directions but with an intensity varying in accordance with the imaging geometry. The ambient component is due to the background illumination which is generally assumed to be uniformly incident from the environment and reflected equally in all

[1]The notion of surface smoothness is formulated in various ways. In Horn's work [BrHo85], smoothness is realized as a minimization of the gradient of the surface. In another work [HoBr86], a more sophisticated formulation is used. A smooth surface is considered as an integrable one, i.e. $z_{xy} = z_{yx}$ where z is the depth map of the image.

[2]The reflected intensity is at a constant ratio to the illuminating intensity.

[3]The reflected intensity is independent of the viewing angle. They are also known as Lambertian reflectors.

directions. It does not contribute to the variation of the reflected intensity across the surface. The specular component represents the reflection of light at the surface interface. It appears as highlights which are concentrated about the specular direction. The chromaticity of the specular component is related to the scene illuminant but independent of the surface pigment.

Our strategy is to uncover the the specular component in a picture, and then remove it from the image resulting in an image containing only the diffuse and the ambient components. Since ambient reflection is constant with respect to imaging geometry and diffuse reflection is usually assumed to be Lambertian [Shaf84a,WeHe66], with the simple reflectivity function describing the two reflection processes, the shape of the objects can be recovered.

We would also like to move from the monochromatic world to the color world. One reason is that specular image component is easier to detect in a color image as it exhibits a consistent color across the whole image. Another reason is that a color image represents a more realistic picture of the world.

1.2. The Scope of Investigation

The diffuse component appears as intensity variation in an image. Unfortunately, the specular component is also present as intensity variation. Since the intensity profile is the only information available in a black-and-white picture, given a monochromatic image there is no way to separate the two features without assumptions about their patterns of variation. Usually, slow variation and sharp peaks are the commonly employed

assumptions about the intensity profiles of the diffuse and specular image components respectively. Nonetheless, the assumption of slow intensity variation inevitably limits the class of surfaces to relatively flat ones. The second assumption restricts the world to that of optically smooth surfaces that reflect highlights concentrated around the specular direction. Therefore, it is preferable to be able to uncover the specular image component without subjecting the method to these restrictive assumptions.

The distinctiveness of specular reflection can be recognized when data from multiple chromatic channels is brought together. Since specular reflection is an interface phenomenon[4], its occurrence has no relationship to the surface reflectance[5]. In an image, the specular component should show uniformity in spectral features across differently colored regions. Thus, in spite of its intensity variation being obscured in the monochromatic image, the spectral characteristics of the specular component can be calculated by comparing the data from different chromatic channels. Hence, to understand the specular component in an image, we move from the black-and-white to the color world.

The smoothness requirement of the surfaces is still observed in the thesis. Although color gives additional information for understanding the specular component, on the other hand, it also adds an extra dimension of complexity to the analysis. To avoid color changes that obscure other physical events, the investigation is restricted to smooth surfaces in order to avoid any orientation variations being obscured by color edges. However, purely specular surfaces are excluded from consideration. As these surfaces reflect specularities only, removal of the specular image component would thus result in a void of shape

[4]When light strikes a surface, part of it is reflected at the air-material interface due to the difference in the refractive indices across the interface. See [Shaf84b]

[5]When light traverses the bulk of a surface, it undergoes scattering, absorption, and reemission by the colorant particles on the way. The degree of interaction differs with the wavelength of the light. When the light reemerges from the surface, it shows a different spectrum with respect to the incident light. The ratio between the reflected and the incident spectra is called the surface reflectance. See [Shaf84b].

information.

In summary, this thesis addresses the problem of shape-from-shading in a scene domain with the following characteristics:

(1) color images.

(2) partially specularly reflecting surfaces.

(3) smooth surfaces.

(4) single distant point source illumination.

(5) light source of unknown location and color.

(6) unknown ambient background illumination.

Figure 1.1 shows a typical scene within the scope of the thesis.

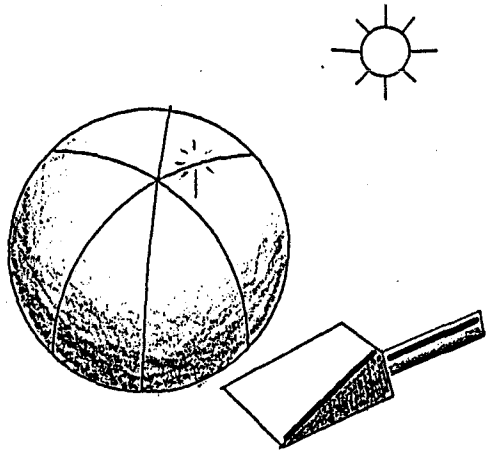

Figure 1.1 : A typical scene consisting of smooth surfaces with differently colored patches. Components due to diffuse and specular reflections are reflected off the surfaces.

1.3. The Strategy

Specular reflection is difficult to model analytically. Any shape computation depending upon specular reflectance is expected to be complicated. In order to avoid the problem, we prefer to remove the specular component from the shading information, to obtain an image consisting of only the diffuse and the ambient components, on which the shape computation is realizable. As ambient reflection is constant with respect to the image geometry, and diffuse reflection is usually assumed to be Lambertian [Shaf84a,WeHe66], a shape-from-shading algorithm using the Lambertian reflection model augmented with a constant ambient term can be applied to the image resulting from the specularity removal process to recover the surface shape.

1.3.1. Removing the Specular Image Component

The first step in our method is to remove the specular component from the shading information. As a color image contains sufficient information about the spectral characteristics of the specular component, we can combine the information of the intensity variations in multiple chromatic channels to resolve for the color of the specular component.

In the thesis, Shafer's Dichromatic Model [Shaf84b] for color reflection is used as an analytical tool to understand the relationships between different image components. The model has two merits. First, it is applicable to a wide class of surfaces, namely those of optically inhomogeneous materials[6]. This fits with our restriction that no region is purely

[6]They are composed of colorant particles embedded in a bulk of optically transparent medium. Light is reflected specularly at the interface and diffusely by the body colorants. See Chapter 2.

specular. Second, the model is based on the theories of optical physics. In addition, the model simplifies computation on color reflection as it describes the phenomenon of reflection as a linear process. Reflected light is interpreted as a sum of independent reflection components[7]. In an image, while the diffuse component is responsible for the diffuse shading, the specular component accounts for the highlights, and their combination yields the reflected light. Since the various components are mingled together in a linear combination, we can employ linear algebra to help in solving for their properties, in particular, the specularity color and the diffuse component intensity.

Applying the linear transformation from light mixtures to color coordinates[8], the reflected color can be described as a linear combination of the color of the respective reflection components. Consequently, pixels corresponding to a uniformly colored region are contained in a dichromatic plane[9] defined by the color vectors of the specular, diffuse and ambient components. As regions of different body reflectances reflect light in different colors, they should show different colors in their diffuse components. It thus follows that they are contained in different dichromatic planes. However, there is one uniformity among the regions. Since specularities arise from interface reflection, they are not affected by the spectrally-biased body reflectance. Therefore, every region should reflect specularities of the same color[10]. As the color of the specular image component is common to all the regions, it should be obtainable as the intersection line of all the dichromatic planes.

A specularity removal algorithm exploiting the invariance of the specularity color across an image has been developed. The algorithm calculates the dichromatic planes in an

[7]Reflected light is a linear combination of the ambient, the diffuse and the specular reflection components.

[8]See [Shaf82] for *Spectral Projection*. It is also discussed in Section 3.3.2.

[9]It is a plane in color space. Here *Dichromatic* means that the plane is defined by the chromaticity of *two* components, namely the interface and body reflection components.

[10]It is the color of the light source [Shaf84b]. Also see [EgHi79].

RGB color space. Unlike Shafer's algorithm which requires a prior image segmentation into uniformly colored regions, our algorithm successfully avoids the segmentation problem by using only local information to calculate the planes. Then the common intersection of the planes is found yielding the specularity color.

After the specularity color has been found, it is used to filter out the specular component. This can be done by resolving each pixel color into two components: one with the specularity color, the other with the orthogonally complementary color. With the former component removed, an image without specularities is obtained. This resultant image preserves the relative intensity of the diffuse and ambient components, and is good enough for shape-from-shading computation. In the thesis, a shape-from-shading algorithm which handles Lambertian reflection under a point source plus an ambient illumination of unknown intensity is presented.

1.3.2. The Shape-from-Shading Computation

While the existing shape-from-shading algorithms restricted the problem domain to surfaces of uniform Lambertian reflectance, we relax the scope to include regionally constant reflectance. Hence, we have to handle images with color regions and thus color boundaries. Moreover, as non-zero ambient illumination is allowed, discontinuities at color boundaries are further complicated by the difference in the ambient reflection components across the boundary[11]. In our problem setting where surface smoothness is assumed, the discontinuity across a color boundary is attributed to the changes in the body reflectance

[11]Surface reflectances are different on both sides of a color boundary, so the two regions reflect the ambient

and ambient component, but not the surface orientation. So, we get around the problem by handling each color region as an independent smooth surface of uniform body reflectance and under constant illumination. The requirements of surface smoothness and constant point source illumination are, however, still observed at the color edges by propagating the corresponding information over as boundary constraints.

Like Horn [HoBr86], we approach shape-from-shading as a variational problem[12] [Wein52,CoHi53,HoBr86]. Suppose we make a guess about the local surface orientation, reflectance, and illumination. We can then calculate the predicted reflected intensity based on the guess. Very likely, our guess will be quite far off, and the predicted intensity so obtained will show a significant discrepancy from the measured one. Since the discrepancy is the manifestation of the error in the guess, we can formulate the problem as a minimization of the discrepancy so that the guess is forced towards the correct solution.

However, just minimizing the discrepancy would not be sufficient to make the problem well-posed[13] because a measured intensity can be caused by an infinite number of combinations of orientation, illumination and reflectance. We have to use some constraints to resolve the ambiguity in order to arrive at a unique interpretation of the image. In the thesis, the requirement of smooth surfaces, constant regional reflectances and invariant illumination are adopted to constrain the problem to a resolvable one.

Mostly, the information about the light source in shape-from-shading computation is known a priori. Brooks and Horn, in [BrHo85], treated the intensity and direction of the

light in different chromaticities.

[12]A variational problem is a one in which it seeks to extremize a functional so that the problem is forced to the optimal solution. Some constraints are usually incorporated into the functional to restrict the set of candidate solutions to a plausible one.

[13]A variational problem is said to be well-posed if extremization of the functional would lead to a unique solution.

illumination as unknowns and derived an iterative scheme to compute the local surface orientation and the source direction and intensity. Although their method iterates locally to evaluate the local orientation, the computation for the source requires global scans over the whole image data.

Aiming at a true local method, the unknown source problem is formulated in such a way that a scheme facilitating local iteration towards both the source and the surface orientation can be derived. This can be done by recasting the problem of shape computation as a variational problem which takes the unknown source as a function rather than a global variable. A constancy constraint is adopted to force the source direction and intensity to be globally constant[14]. The ambient reflection component also introduces unknown parameters in our shape-from-shading problem. In our formulation, the ambient component is handled in a similar way. Consequently, a formulation with the source and the ambient component treated as local quantities can be derived, leading to a local computational scheme which enables a network implementation.

1.3.3. Summary

Our method involves two steps. The first step is to remove the specular component. Then a shape-from-shading algorithm is applied. Since the specular component exhibits a consistent color across an image, different color regions can be brought together to discover the spectral characteristics of the image specularities. Knowing the specularity color, it is possible to remove the specular component from the original color image and obtain an

[14]The unknown source is defined as a function of $s(x, y)$. Any departure of $s(x, y)$ from a constant function is penalized heavily so as to inhibit the s value from varying with respect to (x, y). This is discussed in details in Section 4.2.3.

image consisting of only the diffuse and the ambient components. Shafer's Dichromatic Reflection Model is used as an analytical tool. It portraits reflectioon as a linear process, which leads to the possibility of using linear algebra for the removal of image specularities. As the diffuse component is assumed to be Lambertian, and the ambient component is assumed to be constant, the next step then computes the local orientation using the Lambertian reflection model augmented with a constant ambient term. The shape-from-shading problem is formulated using variational principles while subjected to the constraints of smooth surfaces, constant regional reflectances and constant illumination. Without assuming knowledge about the source location and intensity, but recognizing the effect of regionally constant body reflectances and non-zero ambient illumination, the problem is handled in a more general setting relative to the existing algorithms. A numerical scheme has been derived based on the variational formulation. A local method which enables a network implementation formed the main objective in the design of the scheme.

Chapter 2

Physical Properties of Reflection

In the thesis, scenes of partially specularly reflecting surfaces are investigated. These surfaces reflect both specularly and diffusely. Many common materials reflecting in this way can be described as optically inhomogeneous -- for example, most paints, varnishes, paper, ceramics and plastics [Shaf84b]. Therefore, we focus our attention to the optically inhomogeneous materials in our work. This section presents an account of the specular and diffuse reflections of inhomogeneous materials.

Inhomogeneous materials can be understood as comprised of an optically transparent medium that constitutes the bulk of the material. Embedded in it are the particles of a colorant that produce scattering and coloration [Shaf84b]. When incident light is reflected off a surface of inhomogeneous material, two processes take place. Due to the difference in the refractive indices across the air-material interface, the incident light is partially reflected at the interface. This is *interface reflection*. The residual light penetrates through the interface. Traversing the medium, the light is scattered or absorbed by the colorant on its way. Eventually, some light reemerges through the interface producing *body reflection*. Figure 2.1 illustrates the reflection phenomena occurring on inhomogeneous materials.

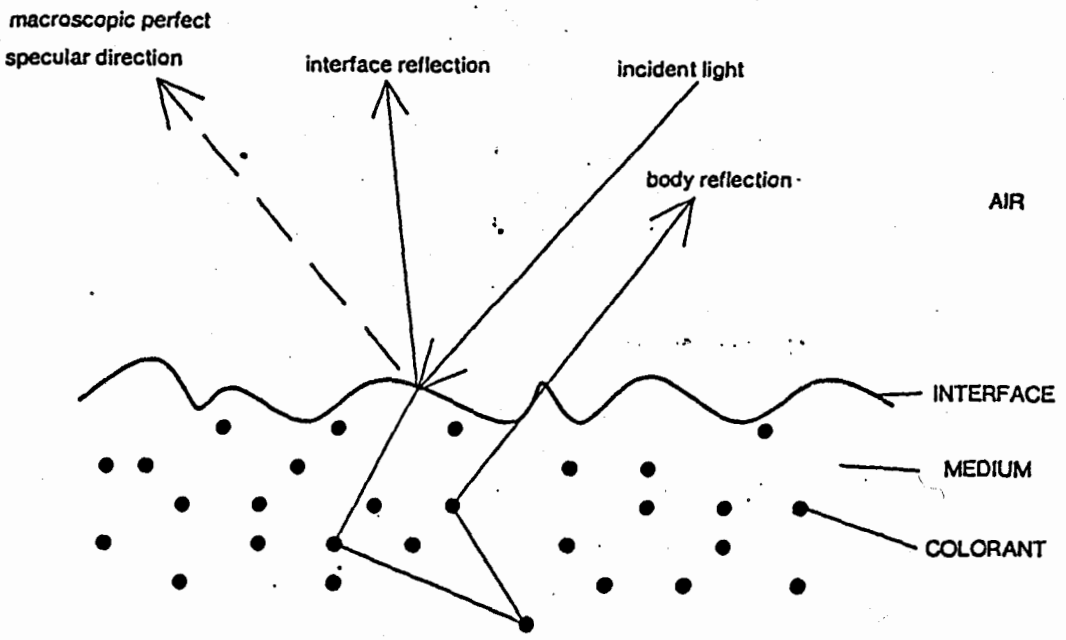

Figure 2.1 : Reflection of Light from an Inhomogeneous Material. (Figure 2-1 [Shaf84b])

Interface reflection is governed by Fresnel's laws. Figure 2.2 depicts the ray diagram of interface reflection.

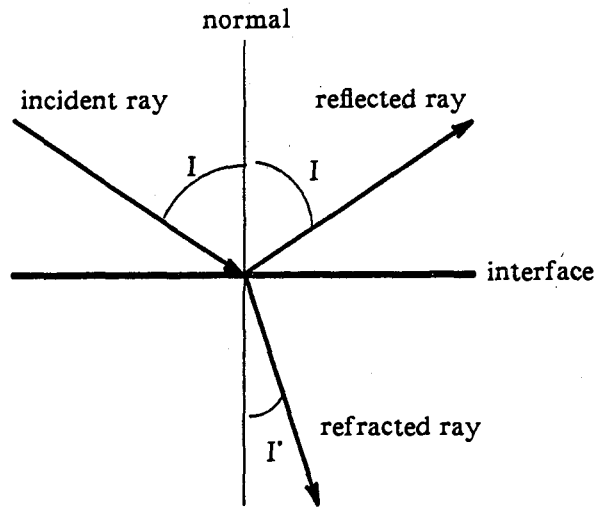

Figure 2.2 : Reflection and Refraction at a plane surface.

Fresnel described the relationship among the reflected light, incident angle and refractive indices in the equations [Long73,Bril80]:

$$\frac{A_r^{\parallel}}{A_i^{\parallel}} = \frac{\tan(I - I')}{\tan(I + I')}$$

$$\frac{A_r^{\perp}}{A_i^{\perp}} = \frac{-\sin(I - I')}{\sin(I + I')}$$

where

- A_i and A_r are the amplitudes of the incident and the reflected light respectively.
- The superscripts $^{\parallel}$ and $^{\perp}$ indicate the light components in and perpendicular to the plane of incidence.
- By the Snell's law, $\eta \sin I = \eta' \sin I'$, where η and η' are the refractive indices of the two media respectively.

So, at an angle of incidence, Fresnel's equations relate the reflection coefficient to the refractive indices. As refractive index depends on wavelength, the reflection coefficient is thus also a function of wavelength. But since refractive index is usually relatively constant across

the visible spectrum, for example, the refractive index of acrylic plastic varies only 1.3% between the ends of the visible spectrum, the reflection coefficient is generally assumed to be constant with respect to wavelength. Thus the reflected light is usually said to have the same color as the incident light [Shaf84b].

However, the color of the body reflection is generally different from that of the incident light. In the surface body, the light undergoes scattering, absorption and reemission upon interacting with the colorant particles. The colorant pigment usually exhibits selective absorption towards different wavelengths. Thus, in general, interaction with the colorant particles results in a spectrally biased reflection. The body reflection therefore exhibits a different color from that of the incident light.

Optically smooth surfaces reflect light along the perfect specular direction. However, real surfaces are always rough. The analytical model assumes that a roughened surface is comprised of small, randomly disposed mirror-like facets, and the mean surface is the one that we observe macroscopically [ToSp67]. Each facet reflects the incident light along its own local perfect specular direction. As the facets are disposed about the mean surface, the local specular direction differs from the macroscopic specular direction but scatters about it. This is the reason specularities usually scatter over a range about the specular direction [Shaf84a]. When light illuminates the facets at an oblique angle, masking and shadowing [ToSp67] of one facet by adjacent ones may occur. See Figure 2.3.

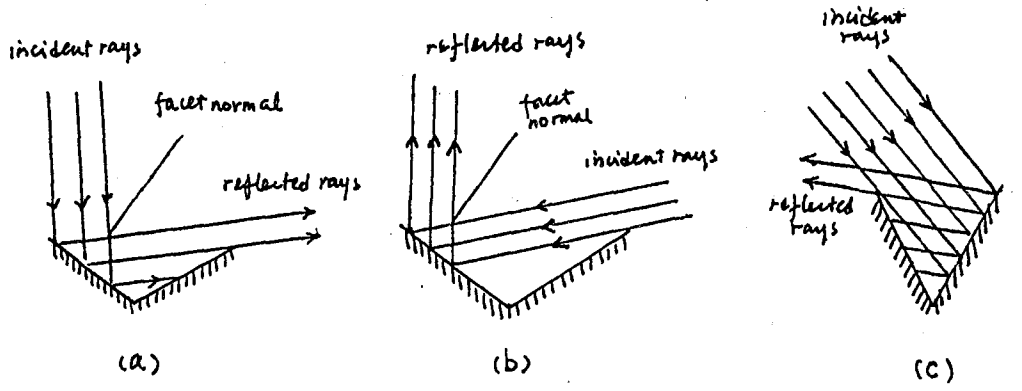

Figure 2.3 : (a) masking, (b) shadowing, and (c) simultaneous masking-shadowing.

This constitutes the geometrical attenuation of the specular reflection. As the attenuation is asymmetric about the specular direction, a peak of reflected intensity is usually observed beyond the specular angle. Figure 2.4 shows some experimental data about the off-specular peak.

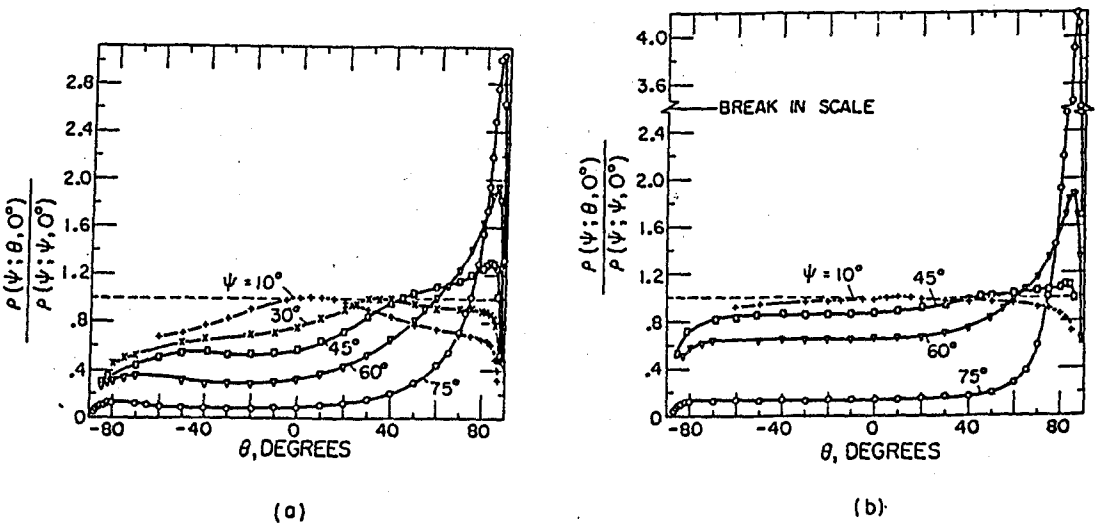

Figure 2.4 : Bidirectional reflectance distributions in the plane of incidence for various angles of incidence ψ, λ=0.5μ. (a) Aluminum (2024-T4), aluminum coated, σ_m =1.3μ. (b) Magnesium oxide ceramic, σ_m =1.9μ. (Figure 2 [ToSp67])

While specular reflection is highly directional, body reflection, contrarily, is usually assumed isotropic[15] [SHaf84b]. Its magnitude depends directly on the intensity of illumination.

[15] i.e. the intensity is independent of the viewing direction. Also see [EgHi79].

Chapter 3

Related Work

As this thesis investigates how the variations in image color tell about the surface shape, the reflectance functions which relate these two quantities are interesting to us. In particular, the complication of expressing the specular reflection coefficient analytically leads us to the idea of removing the image specularities for shape computation. In this chapter, some reflectance functions are reviewed. Then a brief account of the Dichromatic Color Reflection Model is presented. The algorithm due to Shafer [Shaf84b] is discussed. Lastly, the shape-from-shading algorithms are surveyed. As the approach based on variational principles is adopted to handle the shape-from-shading computation, our survey was done primarily on the work based on variational formulations.

3.1. Reflectance Functions

As mentioned in the previous chapter, the image shading provides a good constraint on surface orientation. The diffuse reflection component tells how much the surface turns away from the light source, whereas, the highlights are reflected off the surface at the

characteristic specular angle. If we bring in these constraining factors to solve the shape computation, we have to know how they relate to the surface orientation in a given illumination environment. In this section, we investigate some reflectance models that represent attempts to describe accurately the reflectivity of an illuminated surface.

Phong [Phon75] proposed a shading model describing both the diffuse and specular reflection effects. The shading at a point p on an object surface is expressed as:

$$S_p = C_p [\cos i \ (1-d) + d] + W(i) \cos^n \alpha$$

where

- C_p is the reflection coefficient of the object at point p.
- i is the angle of incidence.
- d is the environmental diffuse reflection coefficient.
- $W(i)$ gives the intensity ratio of the highlight and the incident light.
- α is the off-specular angle.
- n is the power which models the specular reflected light.

The function $W(i)$ and the power n express the specular reflection characteristics of the material. For a highly reflective material, the values of both $W(i)$ and n are large.

Based on empirical measurements of some paints, Horn [Horn75] arrived at an equation describing the reflectivity as:

$$\phi(i,e,g) = \frac{1}{2} s \ (n+1)(2 \cos i \cos e - \cos g)^n + (1-s) \cos i$$

where

- i, e, g are the angles of incidence, reflection and the phase angle respectively. See Figure 3.1 for their definitions.
- s lies between 0 and 1, it determines the fraction of light reflected off the surface interface.
- n determines the sharpness of the specularity peak.

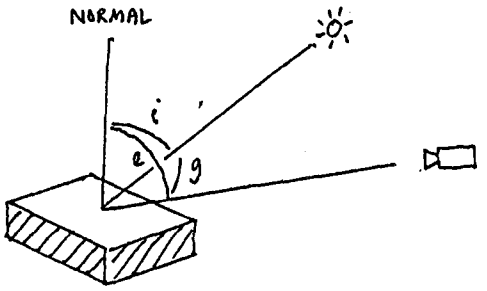

Figure 3.1 : Definition of the angles of incidence i, reflection r and the phase angle g.

Although these models capture the phenomenological characteristics of specular and diffuse reflections quite well, they lack the theoretical background of optical physics. For example, the sharpness of the highlight peaks is modeled by the empirically adjusted parameter n, and no physical justification is made. Not surprisingly, these models cannot predict the commonly observed off-specular peak phenomenon. In the following, we look at two reflectance functions which have been developed on the basis of geometrical optics.

3.1.1. Torrance and Sparrow's Theory of Roughened Surfaces

In their work [ToSp67], Torrance and Sparrow proposed an analytical model which assumes that a surface consists of small, randomly disposed, mirror-like facets. Diffuse reflection of the surface arises from the multiple reflections and internal scattering. Specular

reflection from these facets and the diffuse reflection are postulated as the basic mechanisms of the reflection process.

Consider the geometry at the reflecting surface as shown in Figure 3.2.

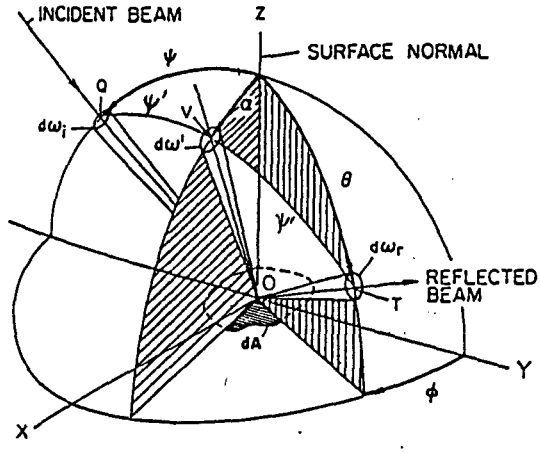

Figure 3.2 : Spatial angles of incident and reflected flux. (Figure 1 [ToSp67])

The composition of the reflected flux dN_r from the surface is expressed as a sum of the specular component $dN_{r,s}$ and the diffuse component $dN_{r,d}$.

$$dN_r(\psi;\theta,\phi) = dN_{r,s}(\psi;\theta,\phi) + dN_{r,d}(\psi)$$

The diffuse component varies directly as the flux incident on the surface, i.e. the projection of the radiance N_i on a unit area of the surface.

$$dN_{r,d}(\psi) = a\ N_i \cos\psi, \quad \text{where } a \text{ is a constant.}$$

Assuming a Gaussian probability distribution of the facets about the mean surface, the probability of the facet orienting α away from the mean surface is:

$P(\alpha) = b\, e^{-(c^2\alpha^2)}$, where b and c are constants.

If f is the area of each facet, and N_i is the incident radiance, the flux incident upon the facets whose normals lie within $d\,\omega'$ is given by:

$$d\,\Phi_i = f\ N_i \cos\psi' P(\alpha)\, d\,\omega'\, dA\ d\,\omega_i$$

As the reflection is governed by Fresnel's laws, the fraction of flux being reflected is given by the Fresnel's reflectance.

$F(\psi',\eta)$ where η is the refractive index.

Masking and Shadowing represent a geometrical attenuation factor of the reflection. Torrance and Sparrow assumed every facet comprises one side of a symmetric V-groove cavity on the surface. In Figure 3.3, the cross-section of the V-groove shows how the masking and shadowing occur.

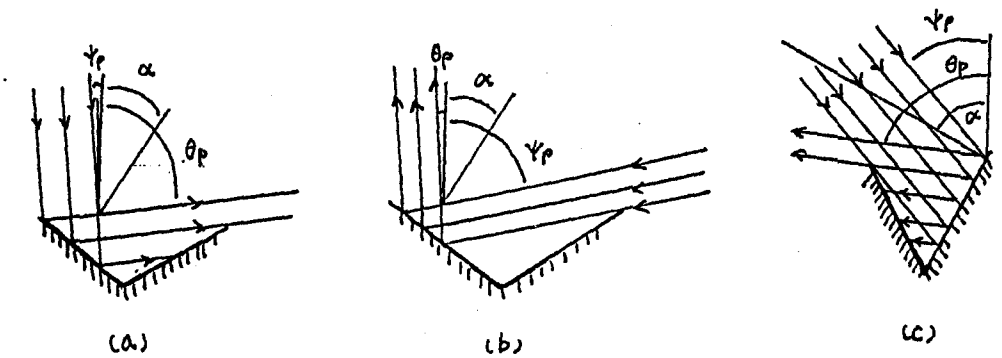

Figure 3.3 : The cross-section of a V-groove shows (a) the masking, (b) the shadowing and (c) the simultaneous masking-shadowing.

Depending on the angles of incidence ψ_p and reflection θ_p in the cross-section, the reflection

is attenuated accordingly. Let $G(\psi_p, \theta_p)$ be the attenuation factor, the reflected flux is expressed as:

$$\begin{aligned} d\Phi_r &= G(\psi_p, \theta_p) F(\psi', \eta) d\Phi_i \\ &= f\ N_i\ G(\psi_p, \theta_p) F(\psi', \eta) \cos\psi' P(\alpha) d\omega' dA\ d\omega_i \end{aligned}$$

Since the reflected flux can also be written in terms of the radiance $dN_{r,s}$,

$$d\Phi_r = dN_{r,s}(\psi;\theta,\phi) \cos\theta\, dA\ d\omega_r$$

and $d\omega'$ can be expressed as $d\omega_r / 4\cos\psi'$,

$$dN_{r,s}(\psi;\theta,\phi) = \left(f\ N_i \frac{d\omega_i}{4} \right) F(\psi', \eta) \frac{G(\psi_p, \theta_p)}{\cos\theta} P(\alpha)$$

Therefore

$$dN_r(\psi;\theta,\phi) = \left(b\ f\ N_i \frac{d\omega_i}{4} \right) F(\psi', \eta) \frac{G(\psi_p, \theta_p)}{\cos\theta} e^{-c^2\alpha^2} + a\ N_i \cos\psi$$

The geometrical attenuation factor $G(\psi_p, \theta_p)$ is a complicated expression. Consider the geometry of masking in a V-groove cavity as shown in Figure 3.4, using the reflection triangle, G can be evaluated as

$$\begin{aligned} G(\psi_p, \theta_p) &= 1 - \frac{m}{l} \\ &= 1 - \frac{1 - \sqrt{1 - A^2}}{A} \end{aligned}$$

$$A = \frac{\sin^2\theta_p - \cos^2\left(\dfrac{\theta_p - \psi_p}{2} \right)}{\cos^2\left(\dfrac{\theta_p - \psi_p}{2} \right) - \cos^2(\theta_p - \psi_p) \sin^2\theta_p}$$

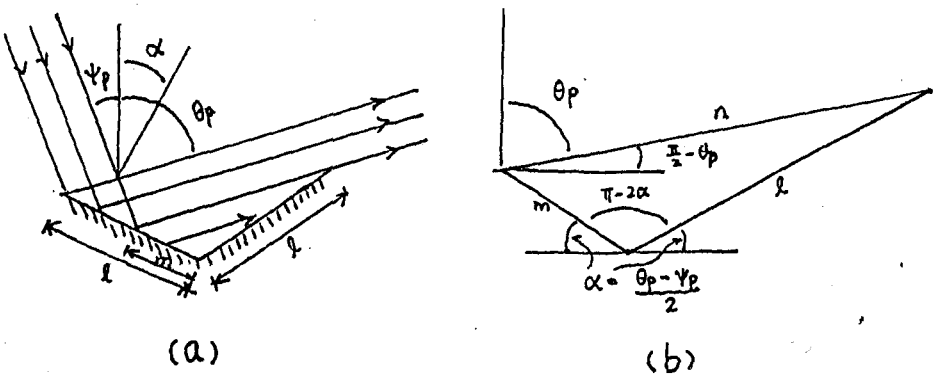

Figure 3.4 : (a) masking in a V-groove cavity. (b) the reflection triangle.

In the cases of shadowing and simultaneous masking-shadowing, the geometry and the reflection triangles are similar. The same equation for the G factor is derived. At certain ranges of incident angles, the reflection is free from masking or shadowing. Torrance and Sparrow enumerated those G values in their paper [ToSp67]. The following table summerizes the results.

Incidence angle : $0 \leq \psi_p \leq \pi/4$		
Reflection angle	Attenuation	Formula
$-\pi/2 \leq \theta_p \leq \frac{\psi_p - \pi}{3}$	masking / masking-shadowing	$G(\psi_p, \theta_p)$
$\frac{\psi_p - \pi}{3} \leq \theta_p \leq \frac{\psi_p + \pi}{3}$	free	1
$\frac{\psi_p + \pi}{3} \leq \theta_p \leq \pi/2$	masking	$G(\psi_p, \theta_p)$

Incidence angle : $\pi/4 \leq \psi_p \leq \pi/2$		
Reflection angle	Attenuation	Formula
$-\pi/2 \leq \theta_p \leq -\psi_p$	masking-shadowing	$G(\psi_p, \theta_p)$
$-\psi_p \leq \theta_p \leq 3\psi_p - \pi$	shadowing	$G(\theta_p, \psi_p)$
$3\psi_p - \pi \leq \theta_p \leq \frac{\psi_p + \pi}{3}$	free	1
$\frac{\psi_p + \pi}{3} \leq \theta_p \leq \pi/2$	masking	$G(\psi_p, \theta_p)$

Torrance and Sparrow's model is based on geometrical optics. It represents an analytical model for the reflection process of roughened surfaces. The model predicts the off-specular peak in very good agreement with the experimental findings [ToSp67].

3.1.2. Cook and Torrance's Reflectance Model

Cook and Torrance [CoTo82] developed a reflectance model that treats reflection as consisting of three components: ambient, diffuse and specular. The ambient component is a

constant term, independent of the imaging geometry. The diffuse component is a Lambertian term, varying as the cosine of the incident angle. The specular component, however, is a very complicated term involving the factors of the surface facet slope distribution, the geometric attenuation (G factor), and the Fresnel's reflectance. In their paper, they expressed the reflected intensity as:

$$I_r = I_a R_a + I_i \cos\psi \, d\omega_i \, (s R_s + d R_d)$$

where

I_a is the intensity of the ambient illumination.

R_a is the ambient reflectance.

I_i is the incident intensity due to the light source.

R_s, R_d are the specular and diffuse reflectances.

s, d are the relative weights of the specular and diffuse components, $s + d = 1$.

Cook and Torrance chose to adopt the Beckmann distribution function [BeSp63] to model the facet slope distribution. In contrast to the Gaussian model [ToSp67], the Beckmann function has an advantage that no arbitrary constant is needed.

$$P(\alpha) = \frac{1}{m^2 \cos^4\alpha} e^{-\frac{\tan^2\alpha}{m^2}}$$

where

m is the *rms* measure of the slope of the facets.

They also used a simplified expression for the G factor[16].

$$G(\psi;\theta,\phi) = \min\left\{1, \frac{2\cos\alpha\cos\theta}{\cos\psi'}, \frac{2\cos\alpha\cos\psi}{\cos\psi'}\right\}$$

Then they were able to write R_s as[17]

[16] Also see [Blin77,Blin78].

[17] *c.f.* Torrance and Sparrow's work [ToSp67].

$$R_s = \frac{1}{\pi \cos\psi} F(\psi',\eta) \frac{G(\psi;\theta,\phi)}{\cos\theta} \frac{1}{m^2 \cos^4\alpha} e^{-\frac{\tan^2\alpha}{m^2}}$$

3.2. Highlight Detection

[Pell86] and [GJT87] deal with the problem of specularity detection. In [Pell86], the specular component is assumed to correspond to a sinusoidal peak in the intensity profile; and the locations of specular reflectances are marked as the simultaneous peaks in all the three chromatic signals. Pellicano uses a local differential operator to find the zero-crossings in the first differential of the intensity image. Zero-crossings of different resolution scales are then *OR*ed together to capture the peaks of various widths. After that, a concavity test is conducted to discrimate those zero-crossings corresponding to the intensity troughs from those corresponding to the peaks. The true highlights are then found as the simultaneous peaks in the *R*, *G*, and *B* signals by *AND*ing the corresponding zero-crossing maps together.

[GJT87] also addresses the problem of highlight identification. Gershon observes that the reflected light shows a color shift when the reflecting surface transits from a diffuse region to a highlight area. In the color-constant space (*C*-space) [Gers87][18], the cluster corresponding to the pixels in the transition region looks like a "dog-leg" structure. Based on this observation, Gershon derives an algorithm which segments an image into uniformly colored regions and then looks for the particular "dog-leg" color shift between all the adjacent region pairs. (See section 6.3 for more detailed discussion on Gershon's method.)

[18]The R, G, B values are transformed into a three-dimensional color-constant space where the chromatic effect of the illuminant is discounted.

Both [Pell86] and [GJT87] present their methods for identifying highlight regions in an image. (See section 6.3 for the comments on them.) However, neither of them addresses the issue of removing image specularities. Furthermore, we argue that finding the highlight regions does not fit very well into our scheme. The reason is explained in the following.

Inside a highlight region, the pixel chromaticity is a mixture of the ambient, the diffuse, and the specular spectra. The composition of the mixture is determined by the geometric parameters of the imaging system and also by the diffuse and specular reflectances of the surface. It is not easy at all to recover the information of the specularities in the presence of so many unknown parameters.

Shafer's method of separating the diffuse and specular components [Shaf84b] represents a more closely related work. His method computes the diffuse and specular components without requiring knowledge about the imaging geometry nor the surface reflectance. In this thesis, we derive a specularity removal algorithm based on Shafer's model of color reflection. Our method is, however, more general than Shafer's method.

3.3. The Dichromatic Reflection Model

Shafer, in his work [Shaf84b,Shaf84a], addressed the problem of separating the components due to the diffuse and specular reflections respectively in a color image. He postulated reflection as a linear process in a simple mathematical model, called *Dichromatic Reflection Model*. Using the spectral projection [Shaf84b,Shaf84a,Shaf82], which is the process whereby pixel values are computed from the spectral power distribution, *SPD*, of the measured light, he transformed the problems of color image analysis into ones in color

space. The problems encoded in terms of the color space parameters by then become nicely manageable using linear algebra. Working on the encoded version of the original problem in color space, Shafer claimed to be able to separate the various components which in the whole account for the perceived color image.

Shafer assumed the illumination to consist of a light source plus an ambient light of low intensity and possibly with a different color than the light source. The reflected light is composed of three parts. One part is due to the interface reflection, constituting the highlights. The second part is due to the body reflection. The third part is caused by reflection due to the ambient illumination. It is a constant quantity as the ambient light is incident and reflected equally in all directions.

In Shafer's model, each of the reflection components has a constant relative *SPD*, i.e. the highlight has a constant color, the body exhibits a constant color, and the ambient reflection adds a constant color to the reflected light. Thus the patterns in the perceived image can be explained as variations of the relative intensity levels of the respective reflection components. Such a variation is attributed to the imaging geometry. So, Shafer postulated that the reflected light is a sum of independent reflection components, and each component corresponds to light of a specific color but its intensity is modulated by the imaging geometry. Consequently, viewing the chromatic and geometric features of the imaged scene independently, Shafer arrived at an irradiance equation expressing the reflected light as a linear combination of the characteristic colors of the respective reflection components. The irradiance equation is stated as:

$$\begin{aligned} L(\lambda,i,e,g) &= L_i(\lambda,i,e,g) + L_b(\lambda,i,e,g) + L_a(\lambda) \\ &= m_i(i,e,g)\,C_i(\lambda) + m_b(i,e,g)\,C_b(\lambda) + C_a(\lambda) \end{aligned}$$

It says that the total radiance L of the reflected light is composed of three independent parts:

(1) the radiance L_i due to the interface reflection.

(2) the radiance L_b due to the body reflection.

(3) the radiance L_a due to the ambient reflection.

and each of these components can be decomposed into two parts:

composition

SPD's of L_i, L_b and L_a which are represented by c_i, c_b and c_a respectively. They are independent of the imaging geometry.

magnitude

geometric scale factors m_i and m_b which depend on the geometry but are independent of the wavelength.

3.3.1. The Dichromatic Model in Color Space

The chromatic information of a beam of light can be adequately represented by a vector in a color space whose basis is composed of orthogonal primary colors. Thus we can have a spectral projection function:

$f : \mathbf{L} \rightarrow \mathbf{C}$.

where

L is the set of possible spectral composition.

C is the color space of dimension n described by the n orthogonal primary colors.

f is a linear mapping function from **L** to **C**.

If $L \in \mathbf{L}$ then there exists one and only one $c \in \mathbf{C}$. such that $f(L)=c$ and the mapping is

linear.

The Dichromatic Model states irradiance as:

$$L = m_i C_i + m_b C_b + C_a$$

Now, apply the linearity of spectral projection,

$$f(L) = m_i f(C_i) + m_b f(C_b) + f(C_a)$$

Rewrite it as:

$$c = m_i c_i + m_b c_b + c_a$$

where

$$c = f(L)$$
$$c_i = f(C_i)$$
$$c_b = f(C_b)$$
$$c_a = f(C_a)$$

In this form, the reflected color is a linear combination of the three color vectors: c_i, c_b and c_a where

- c_i is the characteristic color of interface reflection, or equivalently, the color of incident light.
- c_b is the characteristic color of body reflection.
- c_a is the color of ambient reflection.

Let us look at the coefficients of the various color vectors. The m_i and m_b are scale factors of c_i and c_b. They vary according to the imaging geometry. c_a, on the other hand, is a constant quantity. So, the locus of c in the color space C is a parallelogram plane described by c_i and c_b with the lowest corner displaced from the origin by c_a (see Figure 3.5). Within the parallelogram, the position of any color is determined by its m_i and m_b.

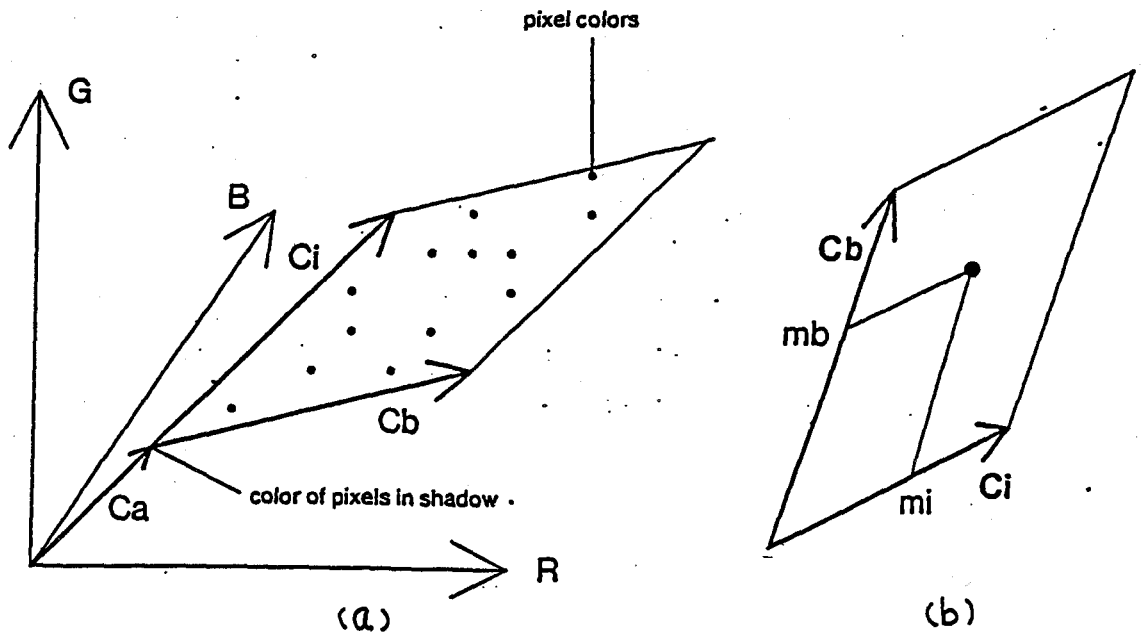

Figure 3.5 : (a) Pixel values on a surface lie on a parallelogram in color space. (Figure 4-1 [Shaf84b]) (b) Position within the parallelogram is determined by m_i and m_b. (Figure 3-2 [Shaf84b])

The color space model provides a powerful analytical tool for interpreting the color of an image. A surface of consistent body reflectance would have its spectral projection lying in a parallelogram floating in the color space. The parallelogram is defined by the three characteristic color vectors -- c_i . c_b and c_a .

The generality of the model well deserves appreciation. There is no assumption about the imaging geometry nor is there any about the surface curvature. The model is not based on any reflectance model. No distribution of the light sources is assumed. Whether it is a point source, extended source or multiple sources, the model applies equally well. Even when the distribution of illumination varies, it does not affect the power of the model in describing and analysing the situation. The model is general and applicable to a wide range of problems. In the thesis, we use the model as an analytical tool in solving for the color of the specular reflection.

3.3.2. The Spectral Projection

Shafer studied the color reflection of a scene in the three-dimensional *RGB* color space. He used linear spectral projection to map the measured light mixture to the color coordinates in the *RGB* space. Specifically, the transformation he used is [Shaf84b,Shaf84a]:

$$c = f(L) = \begin{bmatrix} r \\ g \\ b \end{bmatrix} = \begin{bmatrix} \int L(\lambda)\bar{r}(\lambda)\,d\lambda \\ \int L(\lambda)\bar{g}(\lambda)\,d\lambda \\ \int L(\lambda)\bar{b}(\lambda)\,d\lambda \end{bmatrix}$$

where

- L is the measured light mixture.
- c is the color vector specified by the components r, g and b.
- f is the spectral projection transformation.
- $\bar{r}$, $\bar{g}$, $\bar{b}$ are the responsivity functions of the camera system in the red, green and blue spectra.

The vector space of light mixture has an infinite number of dimensions [Shaf82]. A light mixture L can be considered as a linear combination of the pulse functions

$$\delta_x(\lambda) = \begin{cases} 1, \lambda = x \\ 0, \lambda \neq x \end{cases}$$

They form the basis functions of the vector space. As there are infinite number of these pulse functions[19], the space has an infinite number of dimensions. It is shown that a trichromatic system[20] is adequate to distinguish lights of different colors [Gers84]. So, it is justified to use the linear spectral projection to map the infinite-dimensional light components to the three-dimensional color coordinates.

[19]One pulse function for each λ.

[20]Also see the trichromatic theory of color vision suggested by Thomas Young. The hypothesis were supported by J. Maxwell, H. von Helmholtz and confirmed by E. MacNichol and G. Wald.

3.3.3. Shafer's Algorithm for Separating the Diffuse and Specular Components

A linear spectral projection codes the image pixels as color points in the color space. In the model, each pixel value is a linear combination of independent color vectors. The color points together sweep out a parallelogram in the color space. By this observation, Shafer suggested a simple algorithm [Shaf84b] for computing the intrinsic images, m_i and m_b, of an imaged surface under no ambient illumination. The algorithm is presented as follows:

Algorithm:

(1) Project the pixel values into a color space as a set of color points.

(2) Fit a plane to these points, with the restriction that the plane must pass through the origin.

(3) Fit a parallelogram on this plane with the lowest corner at the origin. The sides are c_i and c_b respectively.

(4) At each pixel, express its color as a linear combination of c_i and c_b. The coefficients of the combination are the values for m_i and m_b.

With the plane-fitting and parallelogram-fitting operations relaxed, the algorithm can be extended to cope with ambient illumination as well [Shaf84b].

Shafer's algorithm is amazingly neat and simple. However, it is still quite restrictive in handling the complexity of real world images. In practice, there is usually more than one color surface in a scene. So, there is more than one color space parallelogram. Moreover, the distribution of the color points may not be as easy to model as expected. When extending Shafer's algorithm to these cases, difficulties are met.

First of all, Shafer's algorithm relies on fitting planes to the points. When several surfaces happen to occur in the same scene, more than one plane will occur in the color space. Fitting a plane through all the points is a standard problem having standard solutions. Nevertheless, finding different planes so that every point lies on one plane is no less than a segmentation problem. In his paper [Shaf84b], Shafer did make a note that the model in color space assumes a prior segmentation of the image into groups of pixels corresponding to each surface respectively.

Fitting a parallelogram on the plane to the points is another crucial step in Shafer's algorithm. Shafer commented that the distribution of pixel values within the parallelogram must not be pathological[21] [Shaf84b] in order to be able to fit the parallelogram. What he meant is that most of the pixel points should lie close to either the c_i or the c_b axis. Normally, such non-pathological distributions correspond to the cases where sharp and well-defined highlights are reflected off shiny surfaces. The reflection is dominated by the specular component once inside the highlight region and falls to mere diffuse reflection once out of it. However, in practice, many surfaces are rough enough to give an extensively broad highlight area. Those highlights are neither sharp nor well-defined. In those cases, the distribution of the pixel values would be pathological. Moreover, it may happen that only part of a surface is imaged. The whole distribution of pixel values is not available in the image. It is then likely to run into a pathological data sample again.

3.4. Shape-from-Shading Computation

[21] See [KSK87] for non-pathological examples.

As we know, surfaces facing the light source are bright whereas those turning away from the source appear dark. Under fixed lighting, the reflected intensity varies as the imaging geometry. People working on the problem want to recover the information about the surface orientation from the image of reflected intensity. In this area of research, Horn has contributed a lot. Let us review his work in the following paragraphs.

In his doctoral thesis [Horn75], Horn addressed the problem of shape-from-shading. He showed that shape can be recovered from the shading information if the reflectance function and the position of the light source are known. He formulated the relationship between the image irradiance, the reflectance, the source and the gradient of the surface in a system of partial differential equations. Horn wrote the image irradiance equation in the form of a first-order non-linear partial differential equation in two independent variables, x and y:

$$F(x,y,z,p,q) = I(r)\,\phi(i,e,g) - E(r) = 0$$

where

$I(r)$	is the incident light intensity at the image point r,
$\phi(i,e,g)$	is the reflectance function of the imaging angles i,e,g,
$E(r)$	is the intensity measured at the image point r.

Then he sought to solve the equivalent set of ordinary differential equations:

$$\begin{aligned}
\dot{x} &= \lambda F_p \\
\dot{y} &= \lambda F_q \\
\dot{z} &= \lambda (p F_p + q F_q) \\
\dot{p} &= \lambda (-F_x - p F_z) \\
\dot{q} &= \lambda (-F_y - q F_z)
\end{aligned}$$

where

$$\lambda = \frac{1}{\sqrt{F_p^2 + F_q^2 + (p F_p + q F_q)^2}} \quad .$$

The dot denotes differentiation with respect to s, a parameter which varies with the

distance along a characteristic strip[22]. When these equations are integrated numerically along the characteristic strips, characteristic curves can be traced out on the surface and hence the information of depth and surface orientation is obtainable. See Figure 3.6.

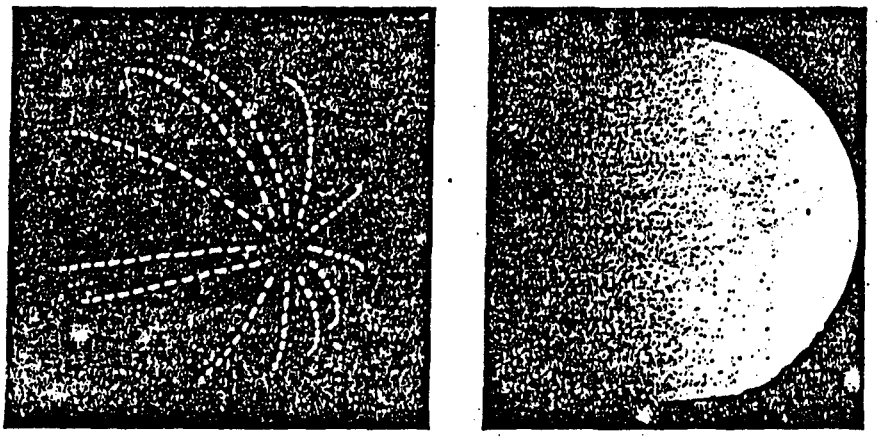

Figure 3.6 : Image of a sphere and the characteristic curves obtained from the shading. (Figure 4-2 [Horn75])

In [Horn77], Horn reformulated his work in terms of gradient space, which makes it much simpler to understand. The work combined his previous shape-from-shading method with geometric arguments in gradient space[23]. Horn introduced the reflectance map [Horn77,Marr82] as a tool for computing surface orientation. For some surfaces, mathematical models are possible for analytical determination of the reflectance function. However, such techniques are usually difficult to use in practice, so reflectance functions are in general determined empirically. For a given type of surface and distribution of light sources, the

[22]See [Gara64,CoHi62].

[23]It is a representation of surface orientation popularized by Huffman and Mackworth in their work [Huff71,Mack73]. A brief account can also be found in [Marr82,IkHo81,Horn77]. The surface patch with gradient $z_x = p$, $z_y = q$ is mapped to the gradient space point at coordinate (p, q). Geometrically, we can think of this as the projection of the Gaussian sphere from its center onto the tangential plane near to the viewer. See [IkHo81].

surface reflects in accordance with the local orientation. Horn recorded the image intensity as a single-valued function in gradient space. This is used as the reflectance map for reading the orientation from the image intensity. Figure 3.7 shows an example.

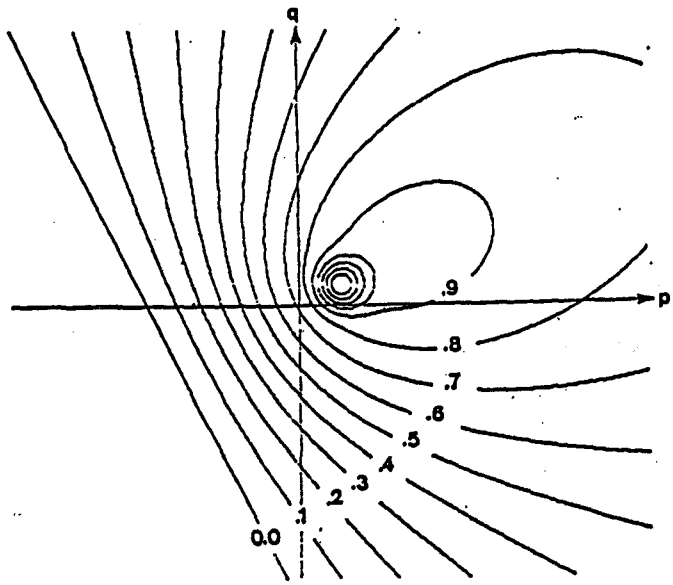

Figure 3.7 : Contours for $\phi(i,e,g) = \frac{1}{2} s\,(n+1)\,(2\cos i\cos e - \cos g)^n + (1-s)\cos i$. This is the reflectance map for a surface with both a diffuse and a specular component of reflectivity illuminated by a single point-source. (Figure 7 [Horn77])

Based on the previous work, Horn and Ikeuchi [IkHo81] developed an iterative method for computing shape from shading using occluding boundary information. The previously used gradient space is insufficient for representing the gradient at occluding boundaries[24] as it becomes unbounded when the surface turns at a right angle to the viewer. Horn and Ikeuchi employed the stereographic plane [IkHo81,Soho41] to express the orientation of surface patches. In the stereographic plane, the whole visible hemisphere of the Gaussian sphere is projected onto a circle only twice as big (see Figure 3.8).

[24]At the occluding boundaries, the surface turns away into the hidden side of the viewed object. Points on the equator of the Guassian sphere correspond to the occluding boundary when the viewer is at the north. With the gradient space projection, the equator maps to the infinity.

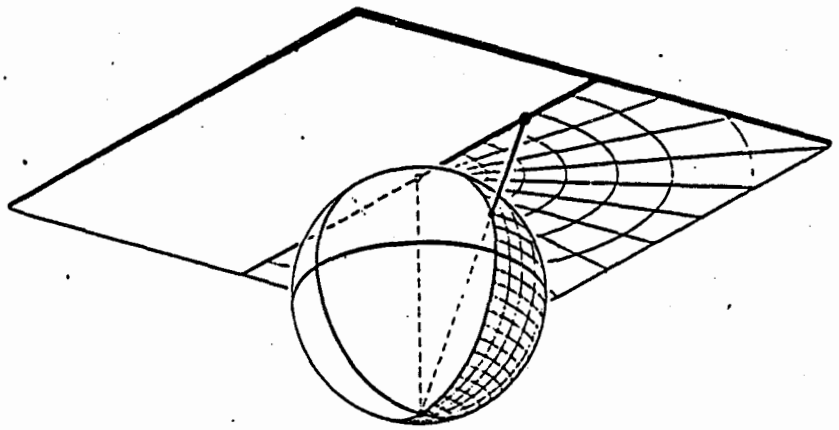

Figure 3.8 : The stereographic mapping projects each point on the surface of the sphere, along a ray from one pole, onto a plane tangent to the opposite pole. (Figure 7 [IkHo81])

Adopting the smoothness constraint which requires minimum slope of the surface, Horn and Ikeuchi formulated the shape-from-shading as a problem of minimizing the functional:

$$\int\int_{\Omega} (f_x^2 + f_y^2) + (g_x^2 + g_y^2) + \lambda [E(x,y) - R(f,g)]^2 \, dx \, dy$$

where

- f, g are the stereographic representation of the surface orientation.
- E is the observed image intensity.
- R is the reading from the reflectance map.

Here the first two terms in the integrand measure the departure from the smooth surface and the last term measures the error in brightness estimation, whereas λ is the weight factor between the smoothness constraint and the error of image irradiance.

Minimization of an integral is a problem in variational calculus [Wein52,CoHi53]. When the integrand F is a function of the independent variables f and g as well as their first partial derivatives f_x, f_y, g_x and g_y, it is sufficient to solve the associated Euler equa-

tions:

$$F_f - \partial/\partial x(F_{f_x}) - \partial/\partial y(F_{f_y}) = 0$$
$$F_g - \partial/\partial x(F_{g_x}) - \partial/\partial y(F_{g_y}) = 0$$

Applying these formulas, we have:

$$\nabla^2 f - \lambda[E(x,y) - R_s(f,g)]\partial R/\partial f = 0$$
$$\nabla^2 g - \lambda[E(x,y) - R_s(f,g)]\partial R/\partial g = 0$$

As a way to solve these equations, Horn and Ikeuchi derived a discrete version of the problem using the finite element method, and constructed an iterative scheme to compute the numerical solution.

In his work [Pent84], Pentland developed a non-iterative local operation to recover the shape from shading with an unknown point source. His method assumes the surface is locally equal-curvatures[25]. Commenting against the restrictiveness of the equal-curvatures assumption and the incapability of the non-iterative method to propagate the occluding boundary constraint across the image, Horn and Brooks extended their previous work to show how the problem can be solved when the reflectance map is not available but is known to have a given form with some unknown parameters.

In [BrHo85], Horn and Brooks proposed an iterative scheme to calculate the surface normals and the source direction and intensity as well. They sought to minimize the brightness estimation error and the departure from surface smoothness. In their method, the functional being minimized is:

[25]Principal curvatures are the curvatures κ_1 and κ_2 that occur along the directions of maximum and minimum surface curvature. For an equal-curvatures surface, $|\kappa_1| = |\kappa_2|$ everywhere.

$$I = \int\int_{\Omega} (E - n \cdot s)^2 + \lambda (n_x^2 + n_y^2) + \mu (n^2 - 1)\, dx\, dy$$

where

- n is the surface normal.
- s specifies the source direction and strength.
- μ is the Langrangian multiplier function used to force n to be unit vector.

The associated Euler equation is then solved for the surface normals. To compute the source, we evaluate the partial derivative of the functional I with respect to s.

$$I_s = -\int\int_{\Omega} 2 (E - n \cdot s)\, n\, dx\, dy = 0$$

From this, we obtain the source s as

$$s = \left[\int\int_{\Omega} n\, n^T\, dx\, dy \right]^{-1} \int\int_{\Omega} E\, n\, dx\, dy$$

In the next year, Horn and Brooks published another paper. This time, they developed a method that enforces the integrability constraint:

$$z_{xy} = z_{yx}$$

Instead of using the regularization method that minimizes the term $n_x^2 + n_y^2$, they changed to minimize the functional:

$$\int\int_{\Omega} (E - R)^2 + \lambda (n \cdot \hat{z})^2 ([n\ n_x\ \hat{x}] + [n\ n_y \hat{y}])^2 + \mu (n^2 - 1)\, dx\, dy$$

This method represents an improvement over the previous one as it converges to a better solution, i.e. closer to the true solution in many cases.

The variational approach provides a mathematical framework enabling a good treatment of shape-from-shading problems. In the thesis, we extend the existing method to color images. A variational formulation is derived which leads to a local and parallel[26] scheme

solving for the local surface orientation and the light source simultaneously. The method also takes care of the irregularities at the color boundaries.

[26]Contrasting to the global method for computing the source by Horn and Brooks in [BrHo85].

Chapter 4
The Algorithms

Our method for solving shape-from-shading for scenes of non-Lambertian reflectors involves two steps. The first one is to remove the specular component from an image. The second step then calculates the shape from the resultant image. In the following sections, the problem is broken down into these two subproblems, namely the specularity removal and the shape-from-shading.

4.1. An Algorithm for Specularity Removal

Shafer's algorithm [Shaf84b] computes the specular and diffuse reflection components of an image. His algorithm could have been used to uncover the diffuse image component for shape-from-shading computation. However, the algorithm assumes prior image segmentation into uniformly colored regions, and non-pathological pixel distributions are required for calculating the specularity and the body colors. As a matter of fact, these pathological distributions occur quite often in real images. Moreover, image segmentation is a difficult problem. In the thesis, we do not count on prior image segmentation nor, to the same extent,

non-pathological pixel distributions. Our algorithm is expected to be more robust in practice for removing image specularities.

The specularity removal algorithm consists of two steps. First, it computes the specularity color. Then, it uses it to filter out the specular component from the original image.

4.1.1. Computing the Specularity Color

Regions of different body reflectance reflect light of different colors. They sweep out different dichromatic planes in color space. The reflected colors, however, share a common specular component. The specular component arises from interface reflection, so it is not affected by the spectrally-biased body reflectance: and thus every region should reflect specularities of the same color. Since the specular component color is common to all the regions, it will similarly be common to all the planes swept out in color space. Therefore as shown in Figure 4.1, it can be calculated as the intersection line of the color planes generated by the different regions. As a result, instead of segmenting the image into different color regions and using each to compute its c_i, c_b, m_i and m_b — as Shafer's algorithm does — we combine the information across regions to compute just c_i, the common specular component.

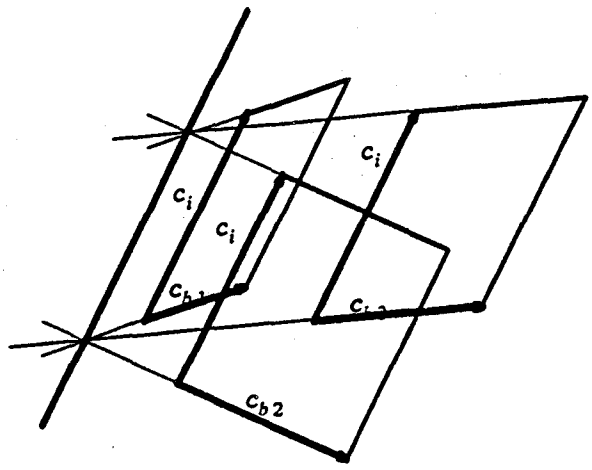

Figure 4.1 : The color planes in the *RGB* space. The intersection of the planes represents the color vector of the specularity color.

Before the color planes are intersected to discover c_i, they have to be computed. This involves calculating dichromatic planes fitting the color points in the color space. Without knowledge about which planes the points belong to, finding planes to all the points would be no easier than a general segmentation problem. However, we can make use of the spatial distribution pattern of the points in the color space to avoid running into the hard segmentation problem. Here, we make two observations:

(1) By the Dichromatic Reflection Model, the image pixels in the same dichromatic region are projected onto the same plane in the color space.

(2) Color patches in the 3-dimensional object world are continuous. They should produce continuous regions in the image space.

At each image pixel, therefore, we can pick a neighborhood ρ that lies in a single dichromatic region. When projecting ρ into the color space, it resides on a single plane. Furthermore, all the pixels in that dichromatic region should project to the same plane. So,

at each pixel in a dichromatic region, if we pick its neighbors and map them to the color space, their best-fitting plane will be biased towards the global plane which corresponds to the region.

Thus, in order to get the color planes, we do not need to do a general region segmentation. Instead, we can make use of the adjacency structure in the image space. We compute the local best-fitting plane of each pixel in the color space. As each local plane represents a suggestion to the global dichromatic plane, the histograms of the local planes then yield the global plane.

Since every dichromatic parallelogram has one of its sides defined by the color vector of the specular image component, the parallelogram planes should all contain the vector representing the specularity color. This is a very nice property as the intersection of the planes will reveal the common vector among them which should be the specularity color vector.

By exploiting the constancy of the ambient component, we can actually discount the ambient component from affecting the above computation. Thus, the same arguments can also be applied to the scenes with non-zero ambient illumination. With non-zero ambient illumination, a dichromatic plane can be expressed as:

$$c = c_a + m_b\, c_b + m_i\, c_i$$

Despite the three component colors — c_a, c_b and c_i — being mixed together, c_a can be sorted out from the other two. Indeed, as c_a contributes a constant term to c only, while c_b and c_i define the plane orientation, c_a takes part only in the displacement of the plane from the origin. Therefore, the effect of c_a can be discounted by dropping the term about the displacement of the fitting plane from the origin. The resultant orientation plane, however,

still preserves the linear combination of c_b and c_i.

Thus, c_i is expected to be defined as the intersection of the orientation planes obtained by the above process. In our method, a least-squares fit is used to find the line of intersection as the line closest to perpendicular to all the plane normals.

4.1.1.1. Finding Local Plane Orientations using a Least-Squares Fit

This section presents a method of calculating the dichromatic plane orientations in color space. The computation is conducted using local information. At every pixel, the local plane orientation is calculated and it, according to the above discussion, is biased towards the global orientation. When working together, these local orientations vote for the orientations of the global dichromatic planes. In the thesis, a least-squares fit [BFR81,PFTV86] is used to find the local plane orientations.

The problem is stated as:

> Given a pixel p_{ij} located at the $(i,j)^{th}$ grid position in the image space, a $n \times n$ neighborhood is then picked and its projection into the RGB space is denoted as the color points c_{kl}, where $k = i - \lfloor n/2 \rfloor, \ldots, i + \lfloor n/2 \rfloor$, $l = j - \lfloor n/2 \rfloor, \ldots, j + \lfloor n/2 \rfloor$. Find the plane which best fits the color points.

We find the best-fitting plane using the least-squares method. Let us express a plane in the RGB space as $\pi_{(b)}$.

$$\pi_{(b)}: \; b = \alpha r + \beta g + \delta$$

where

r, g, b are the coordinates of a point on $\pi_{(b)}$.

The best plane corresponds to the one with the sum of the squares of the residues in the b-direction minimized.

$$\chi^2 = \sum_{i,j} (b_{ij} - \alpha r_{ij} - \beta g_{ij} - \delta)^2$$

At χ^2's minimum, we have

$$0 = \frac{\partial \chi^2}{\partial \alpha} = -\sum_{i,j} (b_{ij} - \alpha r_{ij} - \beta g_{ij} - \delta) r_{ij}$$

$$0 = \frac{\partial \chi^2}{\partial \beta} = -\sum_{i,j} (b_{ij} - \alpha r_{ij} - \beta g_{ij} - \delta) g_{ij}$$

$$0 = \frac{\partial \chi^2}{\partial \delta} = -\sum_{i,j} (b_{ij} - \alpha r_{ij} - \beta g_{ij} - \delta)$$

Defining terms: $S_{ab} = \sum_{i,j} a_{ij} b_{ij}$, we rewrite the system as:

$$\begin{bmatrix} S_{rr} & S_{rg} & S_r \\ S_{rg} & S_{gg} & S_g \\ S_r & S_g & n^2 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \delta \end{bmatrix} = \begin{bmatrix} S_{rb} \\ S_{gb} \\ S_b \end{bmatrix}$$

Denoting the matrices as $S_{(b)}$, $x_{(b)}$ and $s_{(b)}$ respectively, we obtain α, β and δ in x as:

$$S_{(b)} x_{(b)} = s_{(b)}$$
$$x_{(b)} = S_{(b)}^{-1} s_{(b)}$$

and the least-squares estimation error is calculated as:

$$\begin{bmatrix} \sigma_\alpha^2 \\ \sigma_\beta^2 \\ \sigma_\delta^2 \end{bmatrix} = \sigma^2 \sum_{i,j} \left(\frac{\partial x_{(b)}}{\partial b_{ij}} \right)^{(2)} = \sigma^2 \sum_{i,j} \left(S_{(b)}^{-1} \begin{bmatrix} r_{ij} \\ g_{ij} \\ 1 \end{bmatrix} \right)^{(2)}$$

where

the notation $u^{(2)}$ means $u^{(2)} = [\, x^2 \; y^2 \; z^2 \,]^T$ if $u = [\, x \; y \; z \,]^T$.

Therefore, we can obtain the plane normal $n_{(b)}$ and the associated error term $\sigma_{(b)}$ as follows:

$$n_{(b)} = \frac{[\, \alpha \; \beta \; -1 \,]^T}{\sqrt{\alpha^2 + \beta^2 + 1}}$$

$$\sigma_{(b)}^2 = \frac{\sigma_\alpha^2 + \sigma_\beta^2}{\alpha^2 + \beta^2 + 1}$$

However, we can also express the plane as $\pi_{(g)}$.

$$\pi_{(g)} : \; g = \gamma b + \alpha r + \delta$$

Minimizing the sum of the squares of the residues in the g-direction, we obtain the least-squares fitting plane as:

$$S_{(g)} = \begin{bmatrix} S_{bb} & S_{rb} & S_b \\ S_{rb} & S_{rr} & S_r \\ S_b & S_r & n^2 \end{bmatrix} \quad s_{(g)} = \begin{bmatrix} S_{gb} \\ S_{rg} \\ S_g \end{bmatrix} \quad x_{(g)} = \begin{bmatrix} \gamma \\ \alpha \\ \delta \end{bmatrix}$$

$$x_{(g)} = S_{(g)}^{-1} s_{(g)}$$

$$\begin{bmatrix} \sigma_\gamma^2 \\ \sigma_\alpha^2 \\ \sigma_\delta^2 \end{bmatrix} = \sigma^2 \sum_{i,j} \left(\frac{\partial x_{(g)}}{\partial g_{ij}} \right)^{(2)} = \sigma^2 \sum_{i,j} \left(S_{(g)}^{-1} \begin{bmatrix} b_{ij} \\ r_{ij} \\ 1 \end{bmatrix} \right)^{(2)}$$

$$n_{(g)} = \frac{[\, \alpha \; -1 \; \gamma \,]^T}{\sqrt{\alpha^2 + 1 + \gamma^2}}$$

$$\sigma_{(g)}^2 = \frac{\sigma_\alpha^2 + \sigma_\gamma^2}{\alpha^2 + 1 + \gamma^2}$$

Similarly, expressing the plane as $\pi_{(r)}$,

$$\pi_{(r)}: r = \beta g + \gamma b + \delta$$

we obtain:

$$S_{(r)} = \begin{bmatrix} S_{gg} & S_{gb} & S_g \\ S_{gb} & S_{bb} & S_b \\ S_g & S_b & n^2 \end{bmatrix} \quad s_{(r)} = \begin{bmatrix} S_{rg} \\ S_{rb} \\ S_r \end{bmatrix} \quad x_{(r)} = \begin{bmatrix} \beta \\ \gamma \\ \delta \end{bmatrix}$$

$$x_{(r)} = S_{(r)}^{-1} s_{(r)}$$

$$\begin{bmatrix} \sigma_\beta^2 \\ \sigma_\gamma^2 \\ \sigma_\delta^2 \end{bmatrix} = \sigma^2 \sum_{i,j} \left(\frac{\partial x_{(r)}}{\partial r_{ij}} \right)^{(2)} = \sigma^2 \sum_{i,j} \left(S_{(r)}^{-1} \begin{bmatrix} g_{ij} \\ b_{ij} \\ 1 \end{bmatrix} \right)^{(2)}$$

$$n_{(r)} = \frac{[-1\ \beta\ \gamma]^T}{\sqrt{1+\beta^2+\gamma^2}}$$

$$\sigma_{(r)}^2 = \frac{\sigma_\beta^2 + \sigma_\gamma^2}{1+\beta^2+\gamma^2}$$

No matter whether we express the plane as $\pi_{(r)}$, $\pi_{(g)}$ or $\pi_{(b)}$, the best plane is the one with the smallest least-squares error. For example, the best-fitting plane will be $\pi_{(b)}$ if $\sigma_{(b)}$ is the smallest among the three error terms: $\sigma_{(r)}$, $\sigma_{(g)}$, $\sigma_{(b)}$; and $n_{(b)}$ will be the corresponding surface normal.

4.1.1.2. Intersecting the Planes to find the Specularity Color

The specularity color lies in the common intersection of the dichromatic orientation planes. As the modes of distribution of the local orientation planes yield the global dichromatic orientation planes, the common intersection among the global planes may as well be calculated by intersecting the local planes found in the modal classes. However, due to the errors occurring in the imaging process and the plane calculation, the local planes may not intersect nicely. Nevertheless, the line closest to perpendicular to all the plane normals should be good enough to suggest their intersection, and thus the intersection of the global dichromatic orientation planes. From that, the specularity color follows.

For computing the intersection of the local orientation planes, we have the problem stated as:

> Given a set of planes denoted by plane normals $n_i = [\ \alpha_i\ \beta_i\ \gamma_i\ ]$, $i = 1,\ldots,m$, find the line closest to perpendicular to all the plane normals.

Again, we find the best line using least-squares techniques. Choose the vector $s = [\ r\ g\ 1\ ]^T$ on the line closest to perpendicular to all n_i. Realizing the closeness in the sense of least-squares, the sum of the squares of the dot products between s and the plane normals should be minimum.

$$\chi^2 = \sum_i (\ s \cdot n_i\ )^2 = \sum_i (\ r\ \alpha_i + g\ \beta_i + \gamma_i\ )^2$$

At its minimum,

$$0 = \frac{\partial \chi^2}{\partial r} = 2 \sum_i (\ r\ \alpha_i + g\ \beta_i + \gamma_i\ )\ \alpha_i$$

$$0 = \frac{\partial \chi^2}{\partial g} = 2 \sum_i (\ r\ \alpha_i + g\ \beta_i + \gamma_i\ )\ \beta_i$$

Denoting $\sum_i a_i\ b_i$ as S_{ab},

$$\begin{bmatrix} S_{\alpha\alpha} & S_{\alpha\beta} \\ S_{\alpha\beta} & S_{\beta\beta} \end{bmatrix} \begin{bmatrix} r \\ g \end{bmatrix} = - \begin{bmatrix} S_{\alpha\gamma} \\ S_{\beta\gamma} \end{bmatrix}$$

$$\begin{bmatrix} r \\ g \end{bmatrix} = - \begin{bmatrix} S_{\alpha\alpha} & S_{\alpha\beta} \\ S_{\alpha\beta} & S_{\beta\beta} \end{bmatrix}^{-1} \begin{bmatrix} S_{\alpha\gamma} \\ S_{\beta\gamma} \end{bmatrix}$$

Now, we have s specifying the line closest to perpendicular to all the plane normals. We can always normalize s by:

$$s = \frac{[r \; g \; 1]^T}{\sqrt{r^2 + g^2 + 1}}$$

4.1.2. Removing the Specular Component

Knowing the specularity color, it becomes possible to get rid of the image specularities. Our method is to remove all the image components which bear the specularity color. Thus, those components whose color coincides with the specularity color will be filtered out completely. The image specularities, thereupon, will be removed successfully. However, the diffuse and ambient components may also be partially filtered out due to their color decomposition yielding some specularity-colored subcomponents. Nevertheless, the residual subcomponents still preserve the relative intensity changes; and that does provide sufficient information for shape computation as far as shape-from-shading is concerned.

The problem considered in this section is:

Given a color image with each pixel specified by its red, green and blue intensities, and the specularity color denoted by the unit vector $\hat{c}_i = [\, r \;\; g \;\; b \,]^T$, find the image with the specular component removed while the relative intensity changes are preserved.

In the dichromatic reflection model, the reflected color is a linear combination of three components.

$$c = c_a + m_b\, c_b + m_i\, c_i$$

By the linearity of color coordinate computation, we can decompose the color terms in the equation into two orthogonal components: one is parallel to c_i, and the other is perpendicular to it, as shown in Figure 4.2.

$$c^{\parallel} + c^{\perp} = c_a{}^{\parallel} + c_a{}^{\perp} + m_b\,(c_b{}^{\parallel} + c_b{}^{\perp}) + m_i\, c_i$$

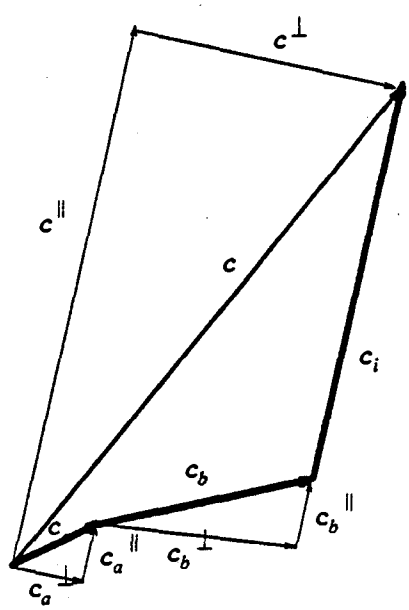

Figure 4.2 : The color terms of the reflected light are decomposed in two orthogonal components: one is parallel to c_i, and the other is perpendicular to it.

From the above equation, it can be observed that all the components which are parallel to c_i can be isolated out, but the factor m_b which governs the diffuse intensity variation is still captured in the residual components.

$$c^{\perp} = c_a{}^{\perp} + m_b\, c_b{}^{\perp}$$

Figure 4.3 shows that the result of removing the components parallel to c_i. Note that the relative diffuse intensity is still preserved.

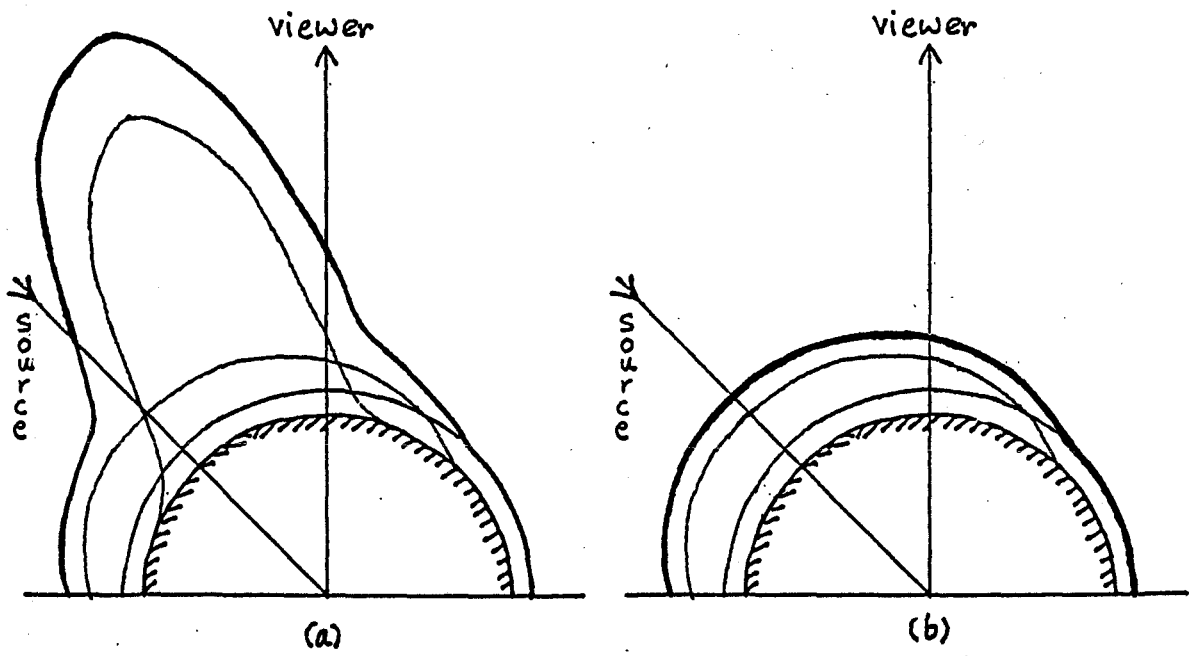

Figure 4.3 : (a) shows the intensity reflected by a spherical surface. (b) shows the case after the specularities are removed. The thick lines trace the intensity profile of the reflected light whereas the thin lines show the profiles of the various components.

$c^{\parallel}$ can be calculated from the given $\hat{c}_i$.

$$c^{\parallel} = (c \cdot \hat{c}_i)\,\hat{c}_i$$

Thus, $c^{\perp}$ is simply

$$c^{\perp} = c - c^{\parallel} = c - (c \cdot \hat{c}_i)\,\hat{c}_i$$

By the above equations, $c - (c \cdot \hat{c}_i)\,\hat{c}_i$ yields an image due to the ambient and diffuse

components while the image specularities are successfully filtered out.

In summary, an algorithm for specularity removal has been presented. It consists of a local method for computing the dichromatic planes. The specularity color is then evaluated as the common intersection of them. Lastly, the image specularities are removed leaving an image consisting of only the diffuse and ambient components. In the next section, a shape-from-shading algorithm is presented. It is a method adopted from Horn's and Brooks's work [HoBr86], but modified to adapt to the image features of reflectance edges and non-zero ambient component. As a result, it is able to handle the images passed down from the specularity removal process.

4.2. A Shape-from-Shading Algorithm for Scenes with Reflectance Edges and Non-zero Ambient Illumination

4.2.1. The Irradiance Equation

Using the method discussed in the previous section, the image specularities can be subtracted away from the original image. As a result, information about the diffuse component is preserved without being obscured by the image specularities. In addition, a Lambertain reflection model becomes sufficient to understand the resultant image.

$$c^{\perp} = c_a{}^{\perp} + m_b\, c_b{}^{\perp}$$

Under a fixed illumination environment (in our case, a constant ambient illumination plus a fixed point source), the resultant image intensity directly relates to the surface orientations. As the relationship is assumed to be Lambertain in the cases of diffuse reflection, the

coefficient m_b of the diffuse component in the above equation can be written as:

$$m_b = n \cdot s$$

where

n is the surface normal

s is the directional vector of the light source

As intensity corresponds to the magnitude of the color vector, change in the light intensity thus does not alter its color. It also means that the relative intensities of the chromatic signals do not change with respect to each other. Hence, any intensity change can be captured equally well by any one of the chromatic signals. Since it is the intensity variation that reflects the surface shape, we can reduce our input from a color image to just one chromatic signal. Therefore, we have a simplified irradiance equation with the color terms replaced by intensity variables.

$$I = I_a + (n \cdot s) I_b$$

Without loss in expressive power, we include I_b into s to obtain a further simplified irradiance equation. We also rewrite I as L and I_a as a to avoid confusion in the use of variables. Consequently, we have the irradiance equation:

$$L = a + n \cdot s$$

4.2.2. The Variational Formulation for Known Illumination

Our specularity removal algorithm yields an image whose intensity is expressible by the simple irradiance equation:

$$L(x,y) = a + n(x,y) \cdot s$$

With the intensity map and the irradiance equation, we want to find the most probable surface described by $n(x,y)$, which, under the illumination environment specified by a and s, yields the given intensity image.

In practice, the image data will be corrupted by noise. So, instead of insisting on the equality of the measured intensity $E(x,y)$ and the predicted irradiance $L(x,y)$, we choose to compute $n(x,y)$ by minimizing the brightness errors:

$$\int\int_{\Omega} (E(x,y) - L(x,y))^2 \, dx \, dy$$

Horn and Brooks derived a numerical method in their work [HoBr86]. They solved the shape-from-shading problem using variational calculus. Their method is employed in our algorithm in computing $n(x,y)$. To start, let us see how the computation of $n(x,y)$ is formulated using variational principles.

Horn and Brooks compute $n(x,y)$ by minimizing the above brightness errors. To make the problem well-posed, $n(x,y)$ is required to correspond to a smooth surface. The integrability property of twice-differentiable surfaces is used to enforce surface smoothness. In other words, a surface defined by a function $z(x,y)$ is said to be smooth if

$$z_{xy} = z_{yx}$$

To enforce that, a penalty term based on the departure from perfect integrability is used. It is written as $(z_{xy} - z_{yx})$. However, as the derivatives become unbounded near the occluding boundaries, Horn and Brooks modified the penalty expression so that it enables incorporation of the occluding boundary information. The modified expression reads as follows:

$$I = (n \cdot \hat{z})([n\, n_x\, \hat{x}] + [n\, n_y\, \hat{y}])$$

Another constraint that forces the surface normals to be unit vectors is written as

$$n^2 = 1$$

Putting all these together, the extremum of the following functional will give the most probable field of surface normals $n(x,y)$.

$$\int\int_{\Omega} F \; dx\, dy = \int\int_{\Omega} (E-L)^2 + \lambda I^2 + \mu(n^2-1) \; dx\, dy$$

where

- λ is a scalar weighting the relative importance between the brightness error and the integrability penalty.
- μ is a Lagrangian multiplier function used to impose the constraint that n be a unit vector.

The extremum of the functional can be obtained by evaluating the associated Euler equation [CoHi53].

$$\frac{\partial F}{\partial n} - \frac{\partial}{\partial x}\left[\frac{\partial F}{\partial n_x}\right] - \frac{\partial}{\partial y}\left[\frac{\partial F}{\partial n_y}\right] = 0$$

which is simplied to read [HoBr86]:

$$-(E-L)L_n + \lambda I\,(I'\hat{z} + 2(n\cdot\hat{z})I'_n) + \mu n + 2\lambda I k + \lambda(n\cdot\hat{z})^2(I'_x(n\times\hat{x}) + I'_y(n\times\hat{y})) = 0$$

where

$$k = (n_x\cdot\hat{z})(n\times\hat{x}) + (n_y\cdot\hat{z})(n\times\hat{y})$$
$$I' = [n\; n_x\; \hat{x}] + [n\; n_y\; \hat{y}]$$

Horn and Brooks [HoBr86] developed a numerical method to solve the above nonlinear partial differential equation. We employ their method to calculate the field of surface normals $n(x,y)$. However, there are additional problems we have to consider in deriving our variational formulation. First of all, besides the surface normals $n(x,y)$, the illumination parameters a and s are unknown quantities as well. Thus, the above formulation has to be modified so that evaluation of $n(x,y)$ does not rely on a priori evaluation for a and s.

Secondly, we are dealing with color images. Although our specularity removal algorithm removes the image specularities, reflectance edges still remain. Thus, the invariance assumption on surface reflectance no longer holds in our case. This leads to discontinuities in the terms a and s in our irradiance equation when stepping across reflectance edges. In the above formulation, however, a and s are globally constant terms.

4.2.3. Incorporating the Unknown Illumination

As long as the illumination parameters are considered independent of x, y, n, n_x, and n_y, the above variational formulation is still valid. To evaluate a and s, we minimize the brightness errors with respect to them. As a and s are constants, no variational calculus is needed.

$$\frac{\partial}{\partial a}\int\int_{\Omega}(E-a-n.s)^2\,dx\,dy = \int\int_{\Omega}(E-a-n.s)\,dx\,dy = 0$$

$$\int\int_{\Omega}(E-n\cdot s)\,dx\,dy = a\int\int_{\Omega}dx\,dy$$

Let A be the area of the image space. Then

$$a = \frac{1}{A}\int\int_{\Omega}(E-n.s)\,dx\,dy$$

For s, we have

$$\frac{\partial}{\partial s}\int\int_{\Omega}(E-a-n.s)^2\,dx\,dy = \int\int_{\Omega}(E-a-n.s)n\,dx\,dy = 0$$

$$\int\int_{\Omega}(E-a)n\,dx\,dy = \int\int_{\Omega}(n\cdot s)n\,dx\,dy$$

Noting $(a^T b)c \equiv (c\,a^T)b$, we have

$$s = \left[\iint_\Omega n n^T \, dx \, dy \right]^{-1} \iint_\Omega (E - a) n \, dx \, dy$$

Now, all the unknowns -- a, s, $n(x,y)$ -- can be evaluated based on each other. As the Horn and Brooks method iterates towards the $n(x,y)$ satisfying the above Euler equation, we can include the following iterative formulas for a and s so that at each iteration, better estimates for all of a, s and $n(x,y)$ are evaluated.

$$a^{(k+1)} = \frac{\epsilon^2}{A} \sum_{i,j \in \Omega} (E_{ij} - n_{ij}^{(k+1)} \cdot s^{(k)})$$

$$s^{(k+1)} = \left[\sum_{i,j \in \Omega} (n_{ij}^{(k+1)} n_{ij}^{(k+1)T}) \right]^{-1} \sum_{i,j \in \Omega} ((E_{ij} - a^{(k)}) n_{ij}^{(k+1)})$$

This method represents a straightforward adaptation of Horn and Brooks's method to handle unknown illumination parameters. However, there is an argument against it. The relaxation method which calculates the new estimates for a and s requires global scans. This is undesirable because we then cannot avoid processing the image data sequentially. This impairs the parallelism of the relaxation method. Thus, we prefer a scheme in which n, a and s are computed iteratively within a local neighborhood only.

We still approach the problem as an extremization problem. Like the previous scheme, this one also minimizes the brightness errors; it also imposes the integrability penalty and the unit normal constraint as before. This time, we treat a and s in a different way. They are not considered constants but functions of x and y. Their constancy is enforced in another way, however.

We try to find the function values of $a(x,y)$ and $b(x,y)$ that minimize the brightness error. Since without imposing any constraints on $a(x,y)$ and $b(x,y)$ the problem

would not be well-posed, we introduce the constraint terms enforcing the functions to be constant across the image.

Constraint for $a(x,y)$:

$$a_x = 0 \qquad a_y = 0$$

Constraint for $s(x,y)$:

$$s_x = 0 \qquad s_y = 0$$

The penalty terms $(a_x^2 + a_y^2)$ and $(s_x^2 + s_y^2)$ are used to inhibit departure from these constancy requirements. Incorporating these terms into the functional, we minimize

$$\begin{aligned} &\int\int_\Omega F \, dx \, dy \\ &= \int\int_\Omega (E-L)^2 + \lambda_1 I^2 + \lambda_2(a_x^2+a_y^2) + \lambda_3(s_x^2+s_y^2) + \mu(n^2-1)\, dx \, dy \end{aligned}$$

The Euler differential equations for this functional are:

$$\frac{\partial F}{\partial n} - \frac{\partial}{\partial x}\left[\frac{\partial F}{\partial n_x}\right] - \frac{\partial}{\partial y}\left[\frac{\partial F}{\partial n_y}\right] = 0$$

$$\frac{\partial F}{\partial a} - \frac{\partial}{\partial x}\left[\frac{\partial F}{\partial a_x}\right] - \frac{\partial}{\partial y}\left[\frac{\partial F}{\partial a_y}\right] = 0$$

$$\frac{\partial F}{\partial s} - \frac{\partial}{\partial x}\left[\frac{\partial F}{\partial s_x}\right] - \frac{\partial}{\partial y}\left[\frac{\partial F}{\partial s_y}\right] = 0$$

which are simplified to read:

$$\begin{aligned} &-(E-L)L_n + \lambda_1 I\,(I'\hat{z} + 2(n\cdot\hat{z})I'_n) \\ &\quad + \mu n + 2\lambda_1 I k + \lambda_1 (n\cdot\hat{z})^2 (I'_x (n\times\hat{x}) + I'_y (n\times\hat{y})) = 0 \end{aligned}$$

where

$$k = (n_x\cdot\hat{z})(n\times\hat{x}) + (n_y\cdot\hat{z})(n\times\hat{y})$$

$$I' = [n\ n_x\ \hat{x}] + [n\ n_y\ \hat{y}]$$

$$(E - a - n \cdot s) + \lambda_2 \nabla^2 a = 0$$
$$(E - a - n \cdot s) n + \lambda_3 \nabla^2 s = 0$$

where

$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$ is the Laplacian operator.

To enforce the constancy requirements on a and s, we choose to make the scale factors λ_2 and λ_3 very large so that deviations from the constancy requirements are penalized very heavily. In the Euler equations, these large scale factors keep any variations in a and s to a very small magnitude.

4.2.4. The Iterative Scheme

For the terms $\nabla^2 a$ and $\nabla^2 s$, we approximate the Laplacian operator by

$$\{\nabla^2 f\}_{ij} = \frac{4}{\epsilon^2} (\bar{f}_{ij} - f_{ij})$$
$$\bar{f}_{ij} = \frac{1}{4} (f_{i,j+1} + f_{i,j-1} + f_{i+1,j} + f_{i-1,j})$$

where

ϵ is the grid distance.

Hence, we have the iterative formulas for a and s as follows:

$$a_{ij}^{(k+1)} = \bar{a}_{ij}^{(k)} + \frac{\epsilon}{4\lambda_2} (E_{ij} - a_{ij}^{(k)} - n_{ij}^{(k)} \cdot s_{ij}^{(k)})$$
$$s_{ij}^{(k+1)} = \bar{s}_{ij}^{(k)} + \frac{\epsilon}{4\lambda_3} (E_{ij} - a_{ij}^{(k)} - n_{ij}^{(k)} \cdot s_{ij}^{(k)}) n_{ij}^{(k)}$$

Adding these iterative formulas to Horn and Brooks's scheme, we have the following itera-

tive scheme.

$$\bar{h}_{ij}^{(k)} = \tfrac{1}{2}(n_{i+1,j}^{(k)} + n_{i-1,j}^{(k)})$$

$$\bar{v}_{ij}^{(k)} = \tfrac{1}{2}(n_{i,j+1}^{(k)} + n_{i,j-1}^{(k)})$$

$$M_{x\,ij}^{(k)} = (n_{ij}^{(k)} \times \hat{x})(n_{ij}^{(k)} \times \hat{x})^T$$

$$M_{y\,ij}^{(k)} = (n_{ij}^{(k)} \times \hat{y})(n_{ij}^{(k)} \times \hat{y})^T$$

$$J_{x\,ij}^{(k)} = (n_{x\,ij}^{(k)} \times n_{y\,ij}^{(k)} + n_{ij}^{(k)} \times n_{yx\,ij}^{(k)}) \cdot \hat{y}$$

$$J_{y\,ij}^{(k)} = (n_{y\,ij}^{(k)} \times n_{x\,ij}^{(k)} + n_{ij}^{(k)} \times n_{xy\,ij}^{(k)}) \cdot \hat{x}$$

$$l_{ij}^{(k)} = J_{x\,ij}^{(k)}(n_{ij}^{(k)} \times \hat{x}) + J_{y\,ij}^{(k)}(n_{ij}^{(k)} \times \hat{y})$$

$$\Gamma_{ij}^{(k)} = [n_{ij}^{(k)} n_{x\,ij}^{(k)} \hat{x}] + [n_{ij}^{(k)} n_{y\,ij}^{(k)} \hat{y}]$$

$$I_{ij}^{(k)} = (n_{ij}^{(k)} \cdot \hat{z})\, \Gamma_{ij}^{(k)}$$

$$j_{ij}^{(k)} = \Gamma_{ij}^{(k)} \hat{z} + 2(n_{ij}^{(k)} \cdot \hat{z}) I'_{n\,ij}^{(k)} - 3 I_{ij}^{(k)} n_{ij}^{(k)}$$

$$k_{ij}^{(k)} = (n_{x\,ij}^{(k)} \cdot \hat{z})(n_{ij}^{(k)} \times \hat{x}) + (n_{y\,ij}^{(k)} \cdot \hat{z})(n_{ij}^{(k)} \times \hat{y})$$

$$L_{n\,ij}^{\perp(k)} = n_{ij}^{(k)} \times (s_{ij}^{(k)} \times n_{ij}^{(k)}), \text{ is a component of } L_{n\,ij}^{(k)} \text{ perpendicular to } n_{ij}^{(k)}.$$

$$r_{ij}^{(k)} = (M_{x\,ij}^{(k)} \bar{h}_{ij}^{(k)} + M_{y\,ij}^{(k)} \bar{v}_{ij}^{(k)}) - \frac{\epsilon^2}{2} l_{ij}^{(k)} - \frac{\epsilon^2}{(n_{ij}^{(k)} \cdot \hat{z})} \Gamma_{ij}^{(k)} k_{ij}^{(k)}$$

$$- \frac{\epsilon^2}{2(n_{ij}^{(k)} \cdot \hat{z})} \Gamma_{ij}^{(k)} j_{ij}^{(k)} + \frac{\epsilon^2}{2\lambda (n_{ij}^{(k)} \cdot \hat{z})^2} (E_{ij} - a_{ij}^{(k)} - n_{ij}^{(k)} \cdot s_{ij}^{(k)}) L_{n\,ij}^{\perp(k)}$$

$$p_{ij}^{(k)} = r_{ij}^{(k)} + \frac{1}{(n_{ij}^{(k)} \cdot \hat{z})^2} [r_{ij}^{(k)} n_{ij}^{(k)} \hat{z}](n_{ij}^{(k)} \times \hat{z})$$

$$\nu_{ij}^{(k)} = \sqrt{1 - p_{ij}^{(k)2}}$$

$$n_{ij}^{(k+1)} = \begin{cases} p_{ij}^{(k)} + \nu_{ij}^{(k)} n_{ij}^{(k)}, \\ p_{ij}^{(k)} - \nu_{ij}^{(k)} n_{ij}^{(k)}, \\ \textit{depending on which gives } n_{ij}^{(k+1)2} = 1. \end{cases}$$

$$\bar{a}_{ij}^{(k)} = \frac{1}{4}(a_{i,j-1}^{(k)} + a_{i,j+1}^{(k)} + a_{i-1,j}^{(k)} + a_{i+1,j}^{(k)})$$

$$\bar{s}_{ij}^{(k)} = \frac{1}{4}(s_{i,j-1}^{(k)} + s_{i,j+1}^{(k)} + s_{i-1,j}^{(k)} + s_{i+1,j}^{(k)})$$

$$a_{ij}^{(k+1)} = \bar{a}_{ij}^{(k)} + \frac{\epsilon}{4\lambda_2}(E_{ij} - a_{ij}^{(k)} - n_{ij}^{(k)} \cdot s_{ij}^{(k)})$$

$$s_{ij}^{(k+1)} = \bar{s}_{ij}^{(k)} + \frac{\epsilon}{4\lambda_3}\left(E_{ij} - a_{ij}^{(k)} - n_{ij}^{(k)} \cdot s_{ij}^{(k)}\right) n_{ij}^{(k)}$$

This scheme operates in a neighborhood around a pixel. It iterates locally towards the solutions to n, a and s. The scheme represents a relaxation method with a high degree of parallelism.

4.2.5. Dealing with the Reflectance Edges

The above relaxation method should work fine within a region of uniform reflectance. However, the method breaks down at the reflectance edges mainly because the constancy assumptions on a and s no longer hold at those places.

Recall the irradiance equation $L = a + n \cdot s$, in which a represents the reflected intensity due to the ambient reflection, and s is a composite term packed with the directional vector, the magnitude of the point source and the surface reflectance. As the surface reflectances differ across a reflectance edge, the terms a and s can no longer be assumed constant. In fact, they undergo a step change across a reflectance edge. Consequently, the constancy assumptions used in the variational formulation on a and s become invalid at the edges leading to failure of the relaxation method.

The relaxation method, however, works fine within a single region, and fails only at the reflectance boundaries. Thus, we can avoid the problem by running the relaxation method on each region separately. That means each region computes its own a, s and n

with the smoothness requirement on n and the constancy requirements on a and s satisfied. At the reflectance edges, information about n and $\hat{s}$ are passed in from the neighboring regions as boundary conditions whereas the propagation of a and $|s|$ are not. Consequently, the regions are knitted together by propagating the surface smoothness constraint and the constant point source direction. At the same time, the constancy constraints on a and $|s|$ are confined to single regions.

This leads to modifications to the iterative formulas for a and s. The formulas for n are fine because the smoothness constraint holds over the entire surface and across different regions. However, for a and s, we estimate them from the local neighborhood in the following way.

For the pixel p_{ij} located at the $(i,j)^{th}$ grid position, the local information about a and s supplied by the pixel p_{kl} in its neighborhood is propagated as:

$$a_{kl} = \begin{cases} a_{kl}, & \text{if } p_{ij} \text{ and } p_{kl} \textit{ are in the same region}, \\ a_{ij}, & \textit{otherwise}. \end{cases}$$

$$s_{kl} = \begin{cases} s_{kl}, & \text{if } p_{ij} \text{ and } p_{kl} \textit{ are in the same region}, \\ |s_{ij}|\,\hat{s}_{kl}, & \textit{otherwise}. \end{cases}$$

This means that if p_{kl} is located in a different region, then the information about a from p_{kl} is simply ignored. Similarly, only the directional aspect of vector s from p_{kl} is adopted while the magnitude is discarded.

In summary, the previous section (Section 4.1) has presented the specularity removal algorithm; and we have an shape-from-shading algorithm in this section, which is able to handle the images passed down from the specularity removal process. Most of these images contain features of reflectance edges and non-zero ambient component. Besides, the

information about the source of illumination is not known a priori. In the algorithm, a variational formulation which deals with unknown ambient and source terms has been laid down. Reflectance edges are handled. From the formulation, a local iterative method has been derived.

Chapter 5

Implementation Results

The algorithms that calculate the specularity color of a scene and remove the specular component from an image have been implemented. Many practical complications have been uncovered and difficulties have been resolved. Since the algorithms have already been explained in the previous chapters, we will concentrate mainly on the implementation details and the corresponding results in this section.

5.1. The Imaging System

First of all, let us look at the set-up of the imaging process. In the experiments, the objects imaged were optically inhomogeneous. Two plastic jugs and a leather bag were used in one experiment, and a vividly colored ball was used in another. While the plastic jugs represent glossy surfaces reflecting strong highlights, the leather bag shows the effect of a comparatively rough surface. In the second experiment, the simple geometry of the beach ball provides a good illustration of how well the algorithm performs in recovering the Lambertian image of a sphere.

The scene illumination was provided by a tungsten bulb. It was chosen to provide a single-color illuminant as required by the dichromatic reflection model. It also agrees with the point source assumption required by the shape-from-shading algorithm. The background illumination was from ordinary office lighting. It was reasonably uniform from all directions and thus fits the model of constant ambient illumination.

We took the pictures using an SC505 VSP Labs CCD camera, on which an Apollo MC TV lens was mounted. The camera was coupled with an International Imaging Systems (IIS) Model 75F which supports a video digitizer. In this setting, the scene intensity captured by the camera was delivered to the IIS and a digitized image was then obtained. In each experiment, three pictures were taken using the red (No. 25), green (No. 58) and blue (No. 47B) Kodak Wratten gelatin filters. In addition, an infra-red filter was employed. Figure 5.1 shows the *RGB* image recorded in Experiment 1.

Figure 5.1 : The scene.

5.2. The Specularity Color

The program computes the specularity color by first finding the dichromatic planes in color space and then intersecting them to obtain the common color vector. Finding the dichromatic plane locally was not as easy as expected. There are practical problems to tackle. First, the local cluster of color points may not be large enough to determine the local plane orientation. Second, even if it is sufficiently scattered, it may not show a planar distribution. In the following, we will discuss these practical problems and their solutions.

By the dichromatic reflection model, neighboring pixels in an image should sweep out a portion of a dichromatic plane in color space. It is demonstrated in Figure 5.2. Thus, given a group of neighboring image pixels, it should be possible to calculate the dichromatic plane by locally fitting a plane to the cluster of the corresponding color points.

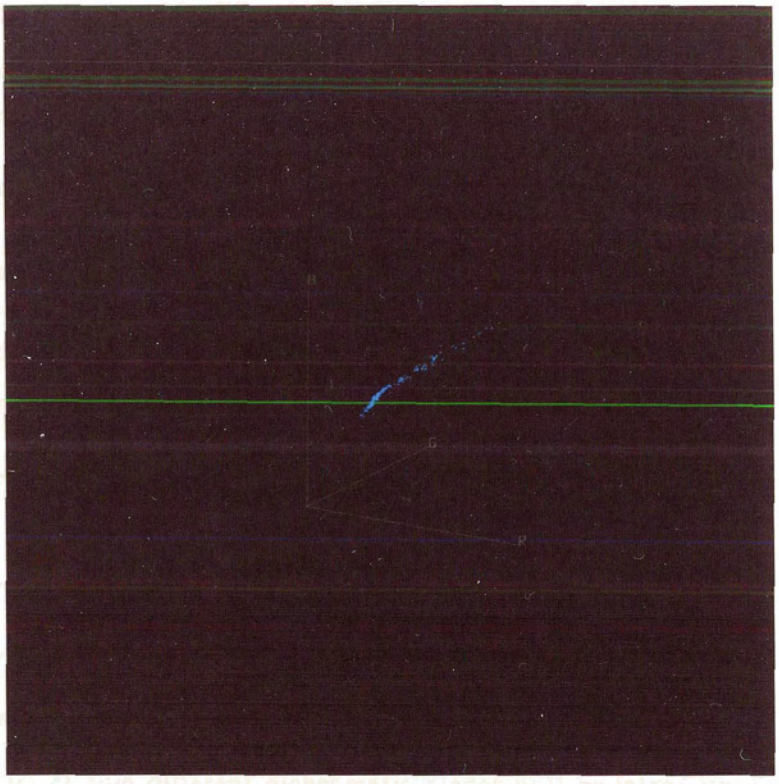

Figure 5.2 : A patch in the blue region in Figure 5.1 are projected to the color space as a set of color points shown in this figure. They sweep out a portion of the dichromatic plane. The cluster's locally fitting plane should be consistent with the global dichromatic plane to which it belongs.

To guarantee successful local fitting, it is necessary that the neighborhood of the pixel under consideration be large enough. Otherwise, the corresponding cluster of color points will not be scattered enough to reveal any information about the dichromatic plane. On the other hand, the neighborhood should be kept small so that it will not extend over two different color regions. A dynamic windowing method was employed to resolve the

problem. It starts with a small window around the pixel under consideration and extends progressively outwards. Once the windowed pixels yield a large enough cluster, it uses it for the subsequent local fitting. The following algorithm has been implemented:

---

(1) for $w = 1$ to *MaxWindowSize* do (2) through (4);

(2) if p_{ij} is a color edge point, and $|i-i_0| \leqslant w$ and $|j-j_0| \leqslant w$, then return(*NoSuitableWindow*);

(3) $p_1 = \frac{1}{4w^2} \sum p_{ij}$, for $|i-i_0| \leqslant w$ and $|j-j_0| \leqslant w$;

$p_2 = \frac{1}{4w^2} \sum p_{ij}^2$, for $|i-i_0| \leqslant w$ and $|j-j_0| \leqslant w$;

(4) if $(p_2 - p_1^2) >$ *ClusterThreshold* then return(w);

(5) return(*NoSuitableWindow*).

The program progressively widens the window from size 3 up to $2\times$*MaxWindowSize* $+1$. At each window size, the variance of distribution of the corresponding cluster is tested to determine if it is large enough (step (3) and (4)). However, if the window extends across a color edge, clusters of two different dichromatic planes will be mixed up. So, the program stops and returns *NoSuitableWindow* at that point, as in step (2). Besides, the program cannot dilate the window without limit. When the maximum window size is reached, and yet, the cluster is still too small, it also stops and returns *NoSuitableWindow* too, as in step (5).

After a suitable window has been selected, the program proceeds to calculate the dichromatic plane from the windowed pixels. Nevertheless, there is a practical complication we have to deal with in this step. In real images, the local fitting plane usually fails to show the global dichromatic plane. This is because the cluster is rarely planar enough for a stable

fit to be possible. So even a small error in pixel intensity arising in the imaging process may displace the fitting plane to such a great extent that the fitting result will no longer represent reliable information about the global dichromatic plane. Figure 5.3 shows a plot of the image pixels in the *RGB* space. Note that the color points are largely clustered in a linear pattern. Fitting a plane to a cluster is thus very unstable.

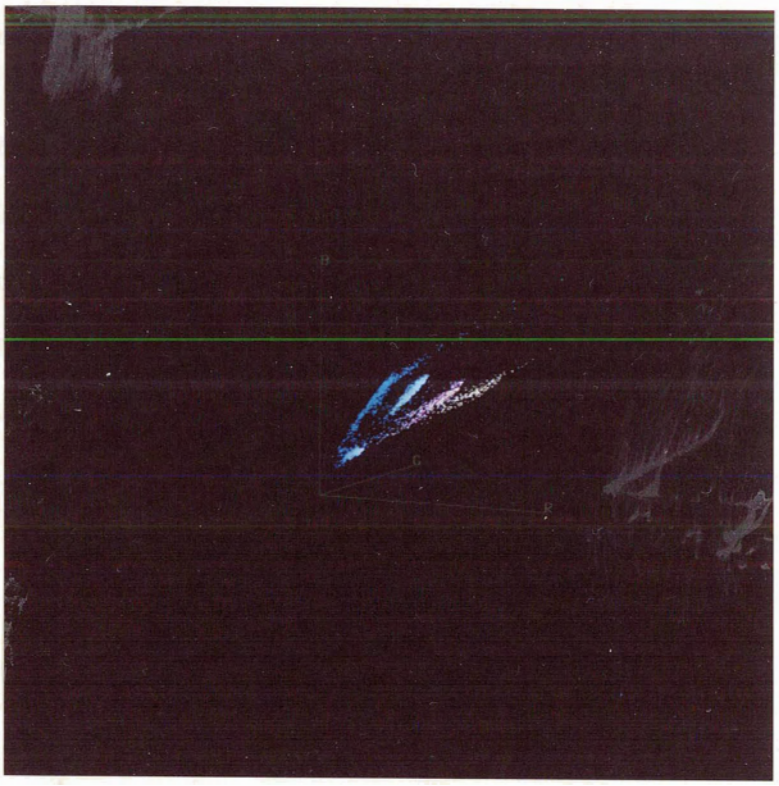

Figure 5.3 : It is a *RGB* plot of the image in Figure 5.1.

In our implementation, the difficulty is resolved by augmenting the straightforward plane-fitting method with a procedure that computes the local plane indirectly when the plane-fitting fails on a non-planer cluster distribution. At first, the program attempts the direct approach of fitting a plane to the cluster of the windowed pixels. If the cluster is distributed in a good planar pattern, the fitting will yield a stable plane and the error estimate

will be small. In that case, the plane fitting is considered successful. Otherwise, it is considered unsuccessful because the cluster is not a good planar distribution. However, if it is not planar, there is likely to be linear scattering. If the color points do yield a dichromatic plane but the local plane fitting fails only because the local cluster distribution is not planar enough, then the linear scattering of the clusters should differ from each other but still lie on the same plane. This variation in the linear scattering among local pixel clusters provides information about the dichromatic plane.

Therefore, in the cases where direct plane-fitting fails, the program finds the line that best represents the longitudinal direction of the cluster. Then it looks at the neighboring clusters and collects information about their fitting planes and lines. Finally, the local plane is calculated as the one that contains all the fitting lines in the neighborhood that is consistent with the fitting planes there too. Algorithmically, it can be presented as follows:

---

(1) At a pixel, select the suitable window size so that the corresponding cluster is large enough for local fitting operations.

(2) Fit a plane to the cluster. If the estimation error is small, the fitting result is reliable and thus can be returned as an estimation for the dichromatic plane.

(3) Otherwise, fit a line to the cluster. Then select a neighborhood so that,
   (a) there is successful plane-fitting, or
   (b) the line-fitting results vary enough to calculate the dichromatic plane.

(4) Within the neighborhood, if there are successful fitting planes, make sure that the planes are consistent with the fitting lines. If they are, return the fitting planes as the estimation for the local dichromatic plane.

(5) Otherwise, compute the plane containing all the local fitting lines and return it as the result of local dichromatic plane estimation.

The algorithm has been implemented. With the pixel distribution as described in Figure 5.3, the plane-fitting was found to be very unstable. So, instead of fitting planes, the program fitted lines to the local clusters. Figure 5.4 shows the line-fitting results.

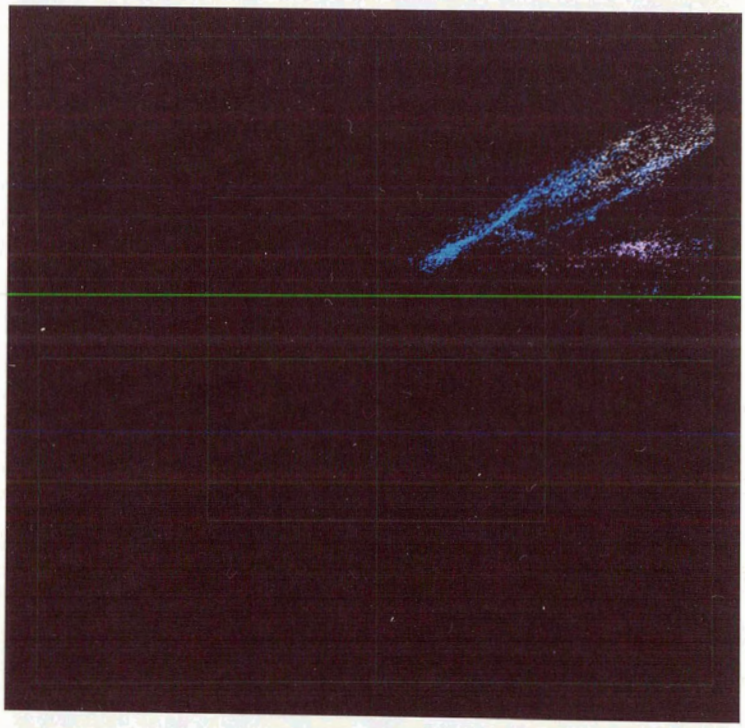

Figure 5.4(a)

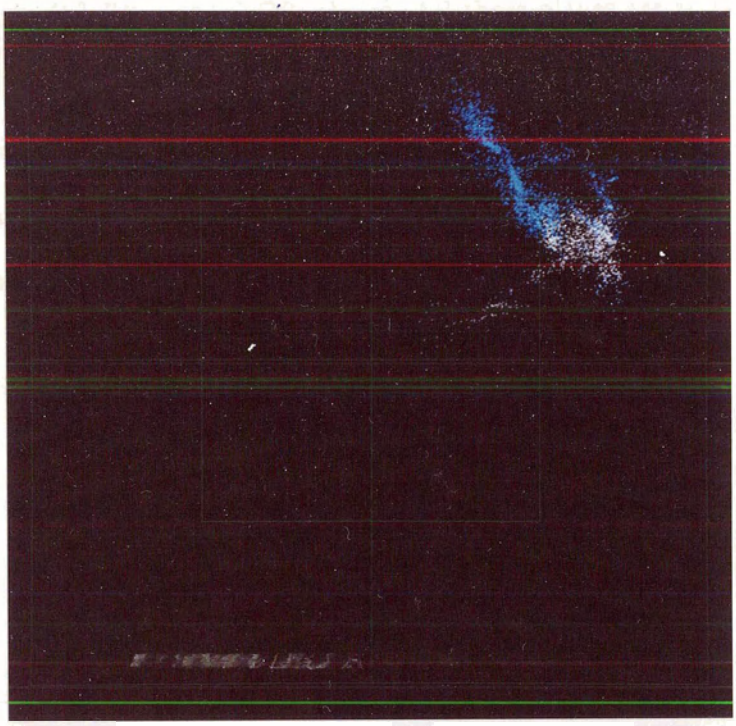

Figure 5.4(b)

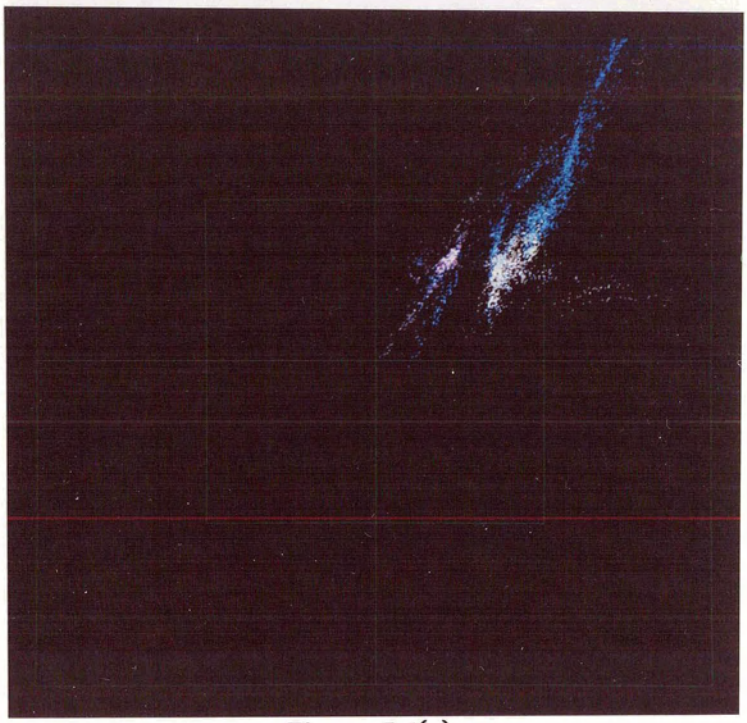

Figure 5.4(c)

Figure 5.4 : Orientation of the fitting lines to the local clusters. They are plotted

on the (a) *RG* , (b) *RB* and (c) *GB* planes. All three planes are at a unit distance from the *RGB* origin.

In the figure, we can see that although each cluster yields only a fitting line, they vary enough from each other to indicate the global dichromatic plane. The result of the local plane calculation is shown in Figure 5.5.

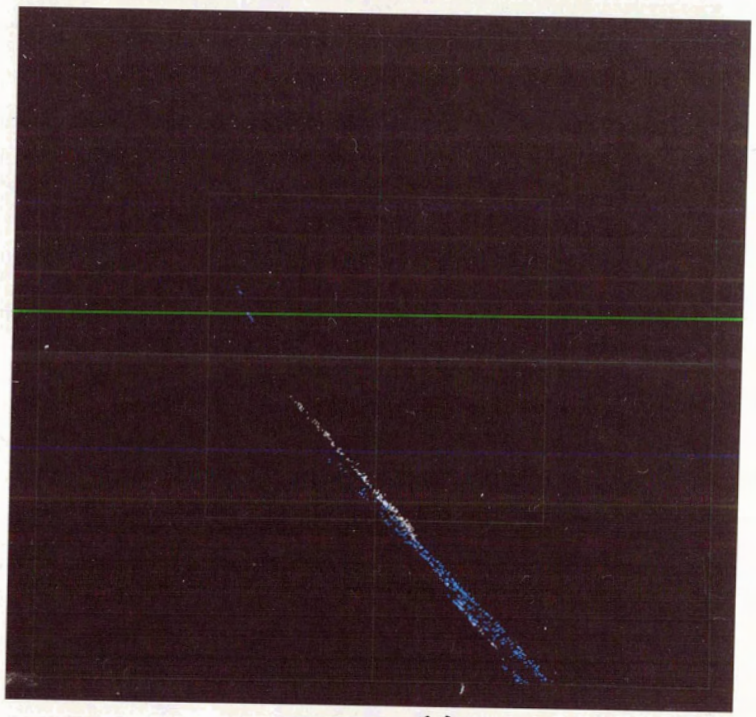

Figure 5.5(a)

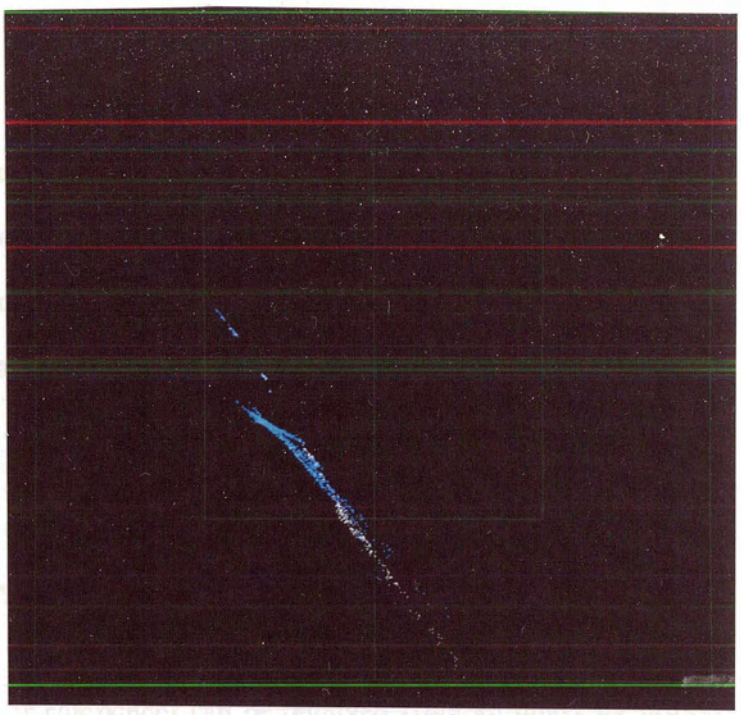

Figure 5.5(b)

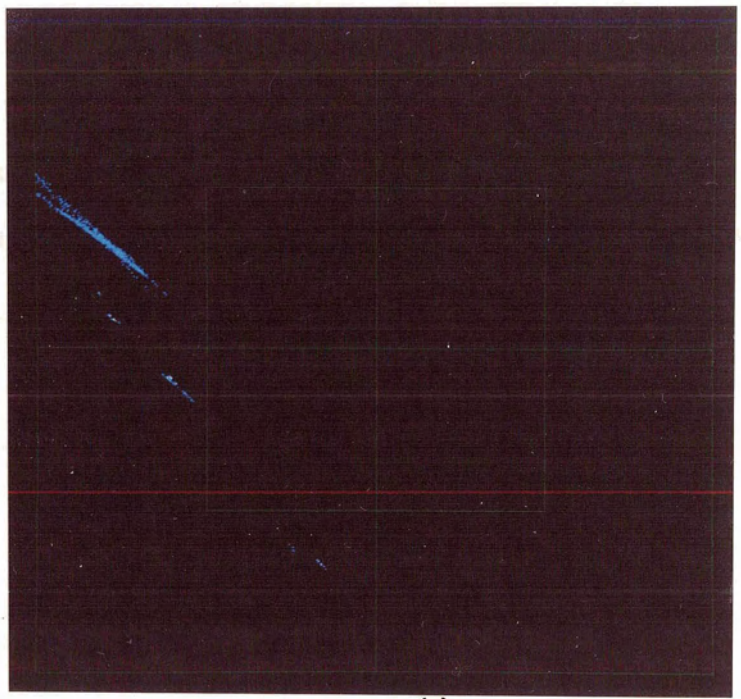

Figure 5.5(c)

Figure 5.5 : Normals to the local fitting planes. They are calculated as the plane

which contains all the local fitting lines in the neighborhood. The plane normals are plotted on the (a) RG, (b) RB, and (c) GB planes.

As explained in Section 4.1.1, the specularity color lies on the line of intersection of the dichromatic planes. With the results of the local plane-fitting, we can calculate the specularity color as the perpendicular to all the local plane normals. In our example, the specularity color was found to lie on the vector: $[\ 0.65\ 0.58\ 0.48\ ]^T$.

5.3. Removal of Specularities

The specular component can be removed from an image by projecting the pixel colors onto the plane orthogonal to the specularity color. The algorithm has been explained in Section 4.1.2. The result of the implementation is quite satisfactory. The following figure shows the result of projecting the pixel colors. Note that the parallelogram patterns of the pixel distribution in Figure 5.3 have been reduced to a set of lines in Figure 5.6. This signifies that the component of specularity color has been removed from the image.

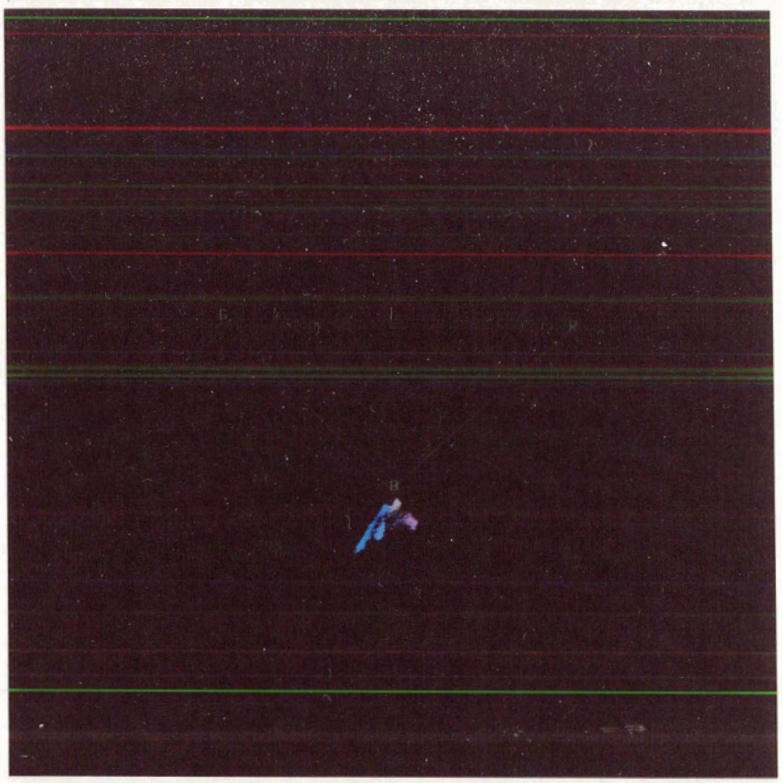

Figure 5.6 : The result of projecting the pixel colors onto the plane orthogonal to the specularity color.

Figure 5.7(a) shows the intensity map of the image with specularities removed. We can see that the highlights have been filtered away successfully. Within a single-color region, the intensity variation is attributed to the body reflection only. Note that the resultant image shows typical Lambertian shading indicating that the specularities have been successfully removed.

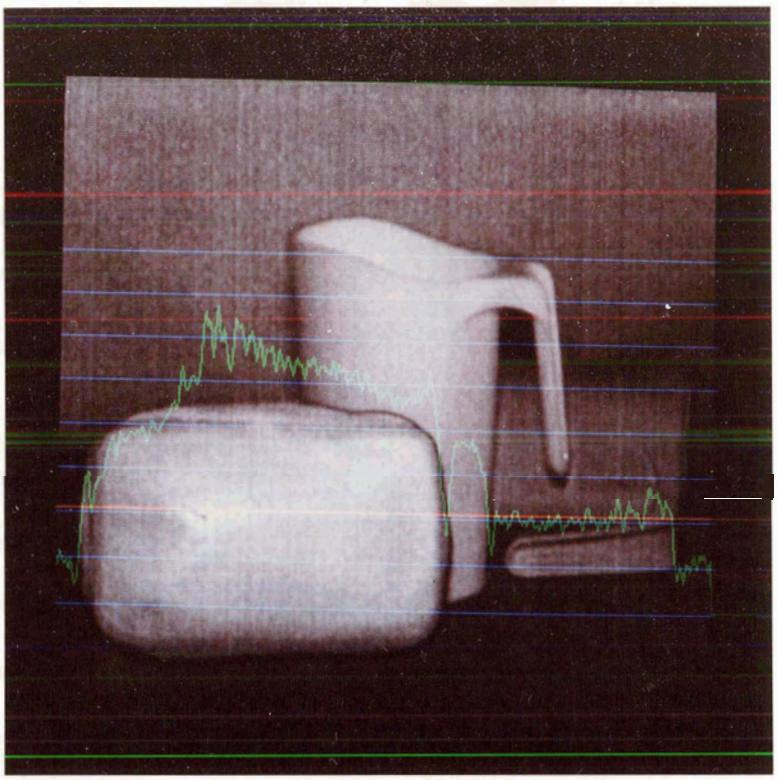

Figure 5.7(a)

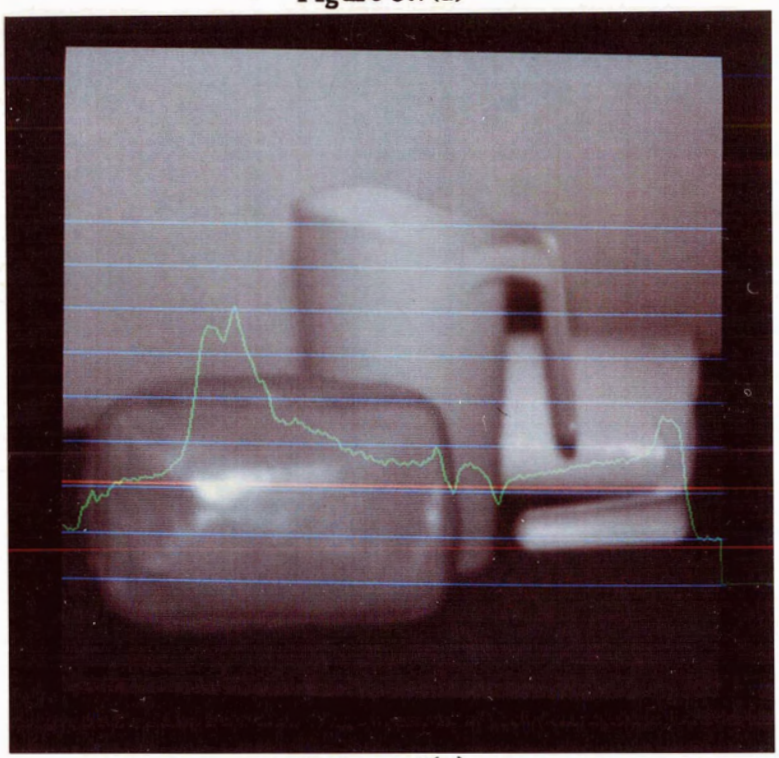

Figure 5.7(b)

Figure 5.7 : (a) is the intensity map of the image in Figure 5.1 with the specular

component removed. (b) is the original intensity map.

We have also run another set of experiments on an image of a beach ball. The results are shown in Figure 5.8 through Figure 5.14. Note that the pixel distribution in the *RGB* space looks much nicer in this experiment. We relate this to the good variety of surface orientations on the spherical beach ball. Figure 5.14(a) displays the image obtained after the effects of the image specularities have been discounted. The success of specularity removal is illustrated in Figure 5.13 as the parallelogram patterns of the color clusters have been reduced to a set of lines. Compare the the way the image intensity varies in Figure 5.14(a) and (b), the sharp peak in Figure 5.14(b) no longer appears in Figure (a). Instead, Figure (a) shows a typical Lambertian shading when the surface turns away gradually from the source.

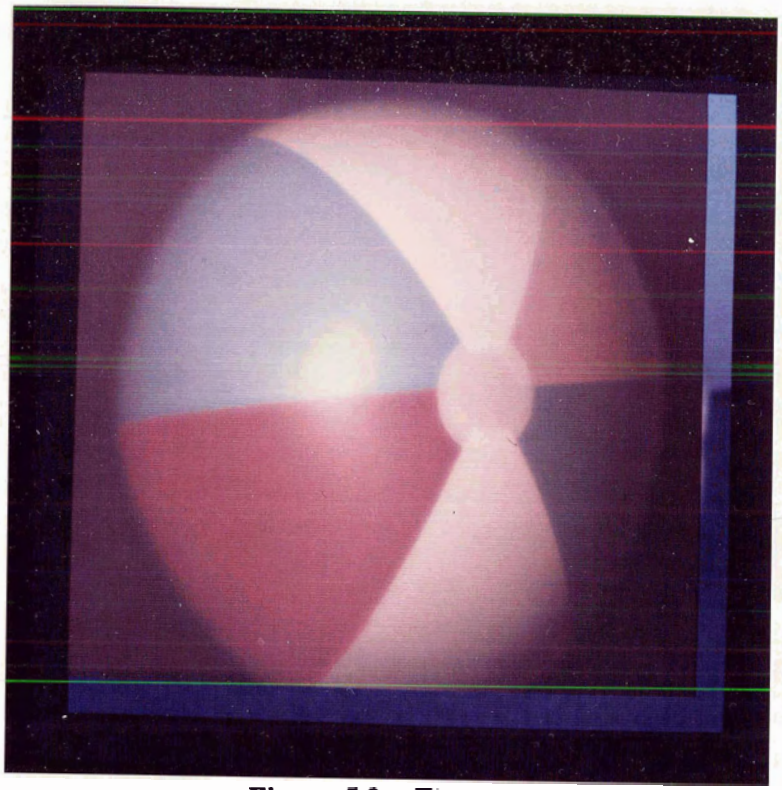

Figure 5.8 : The scene.

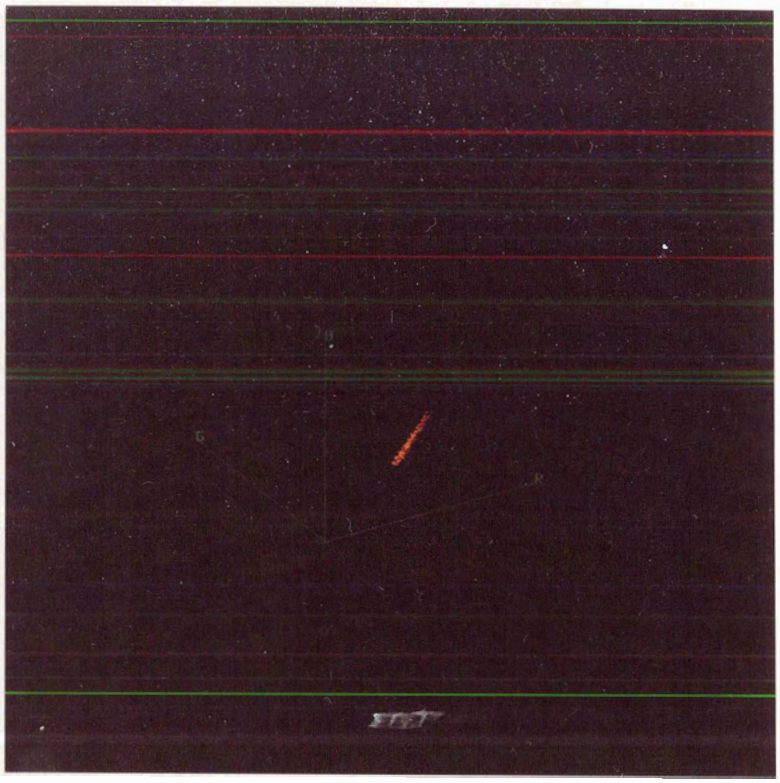

Figure 5.9 : A patch in the red region in Figure 5.8 are projected to the color space as a set of color points shown in this figure. They sweep out a portion of the dichromatic plane. The cluster's locally fitting plane should be consistent with the global dichromatic plane to which it belongs.

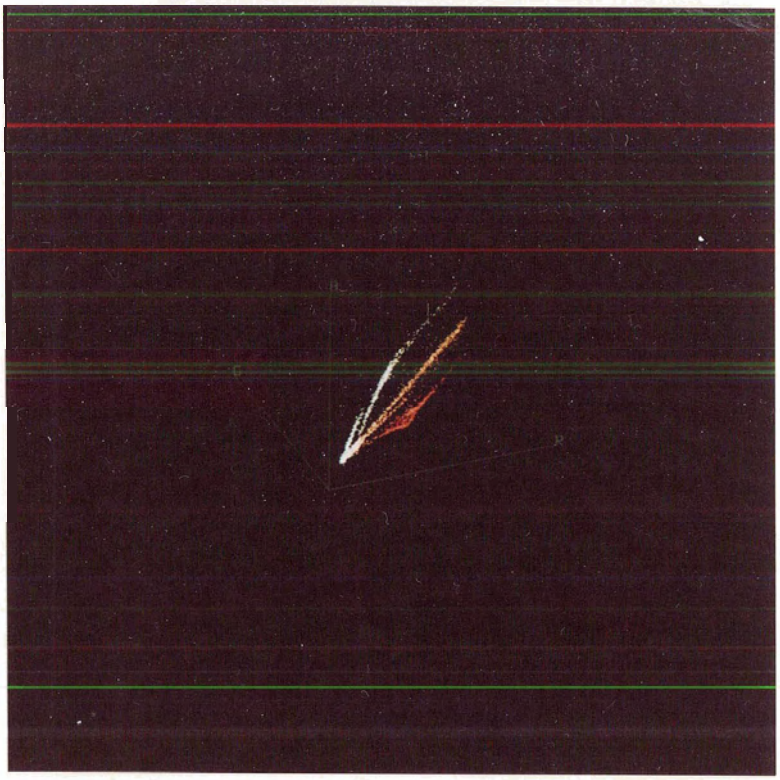

Figure 5.10 : It is a *RGB* plot of the image in Figure 5.8.

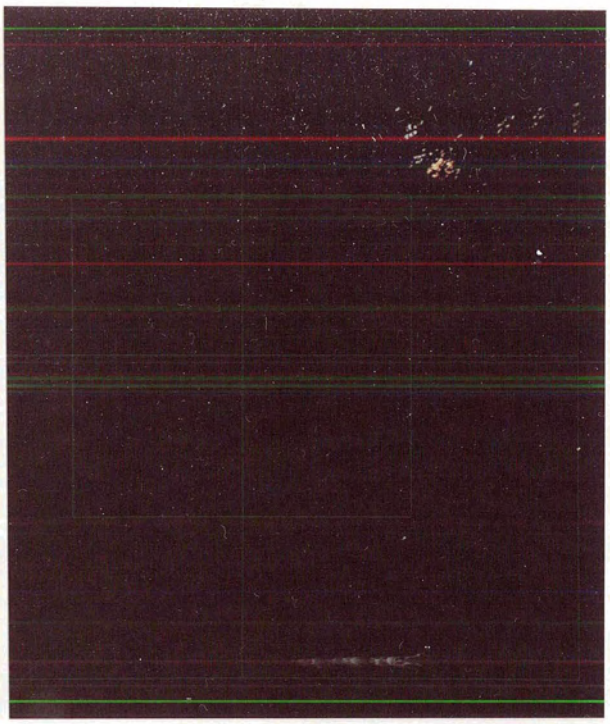

Figure 5.11(a)

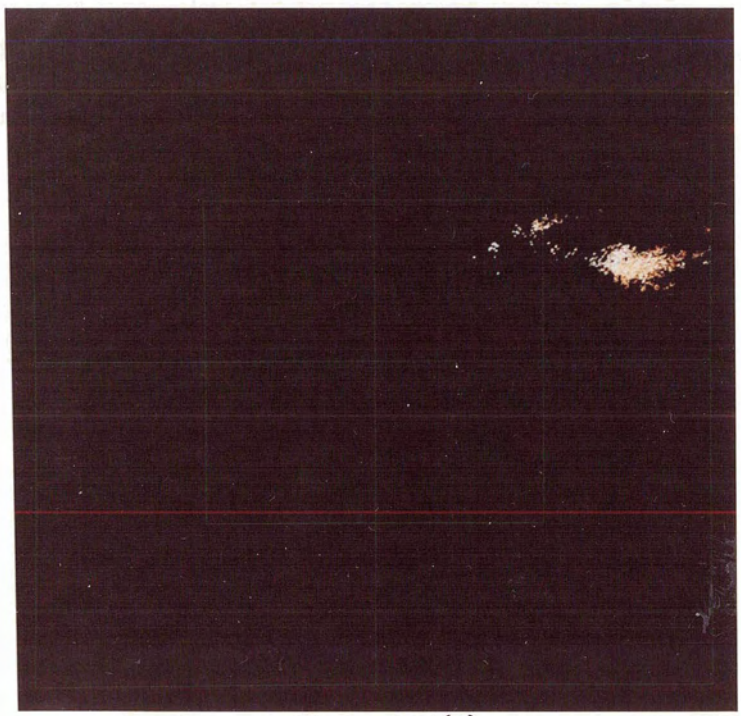

Figure 5.11(b)

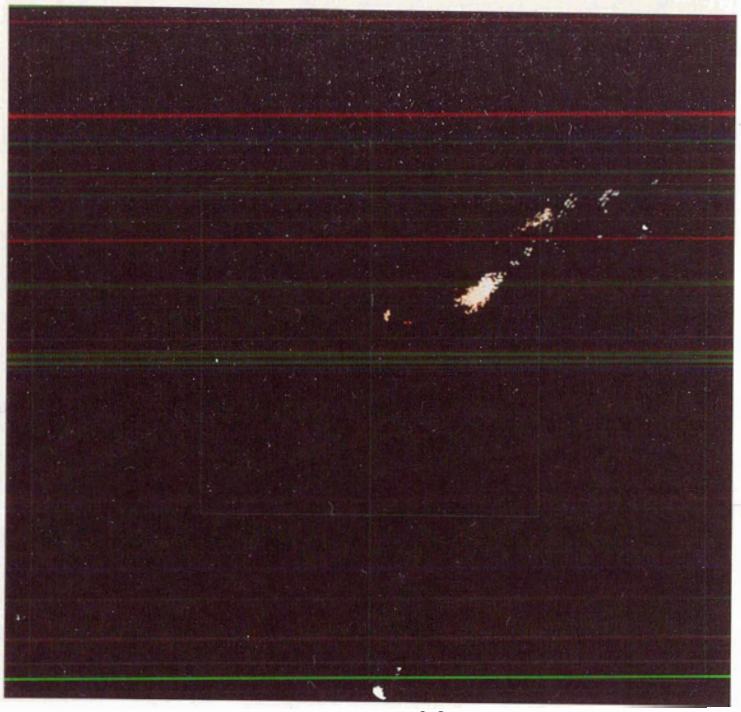

Figure 5.11(c)

Figure 5.11 : It shows the orientation of the fitting lines to the local clusters. They are plotted on the (a) *RG* , (b) *RB* and (c) *GB* planes. All three planes are at a unit distance from the *RGB* origin.

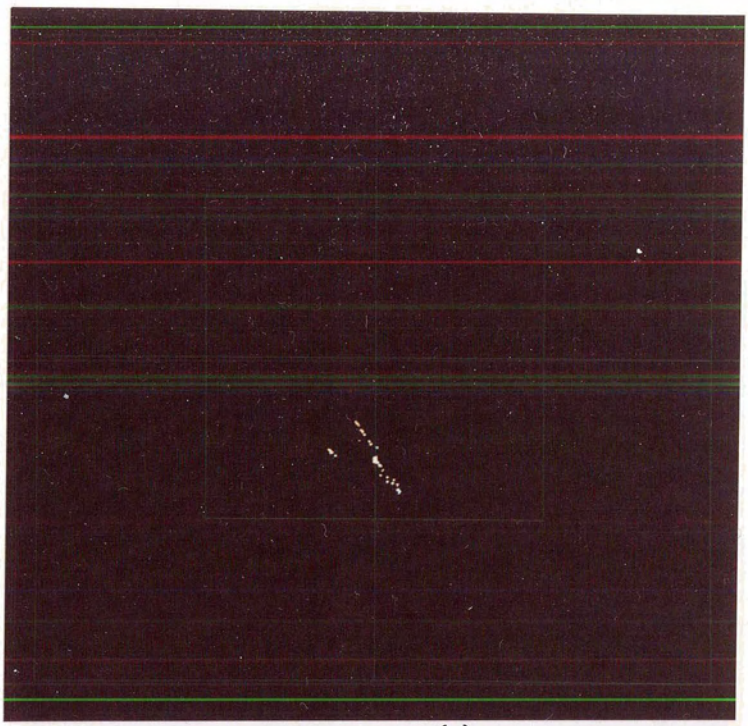

Figure 5.12(a)

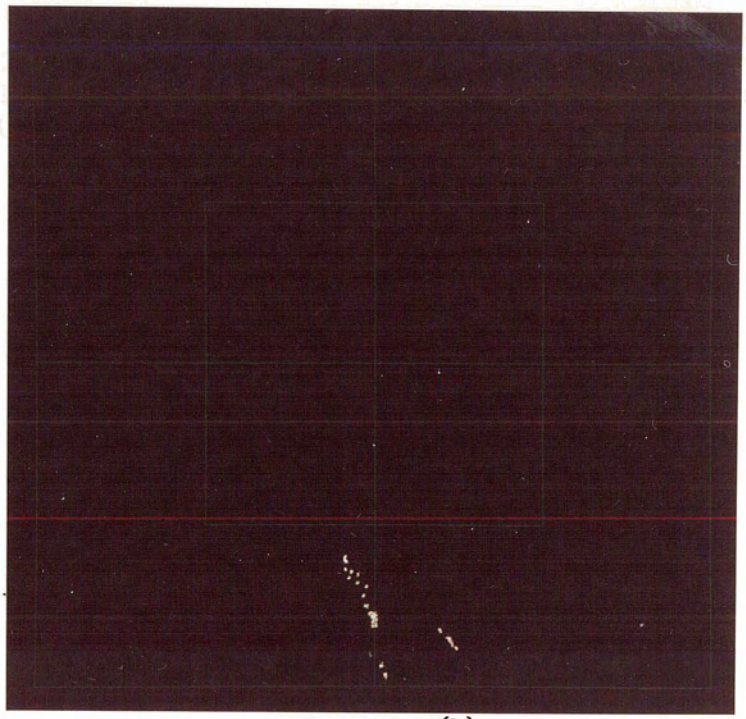

Figure 5.12(b)

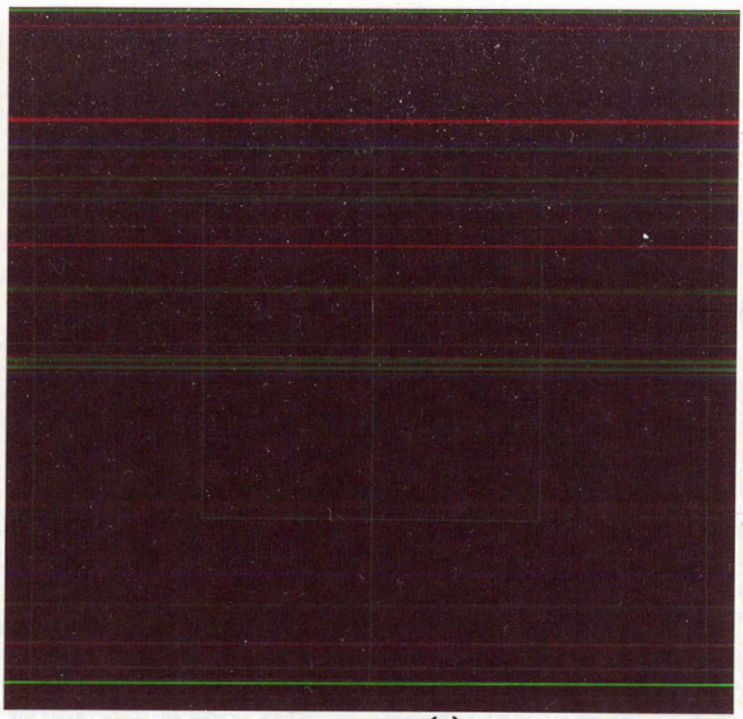

Figure 5.12(c)

Figure 5.12 : It shows the local planes. They are calculated as the plane which contains all the local fitting lines in the neighborhood. The plane normals are plotted on the (a) *RG* , (b) *RB* , and (c) *GB* planes.

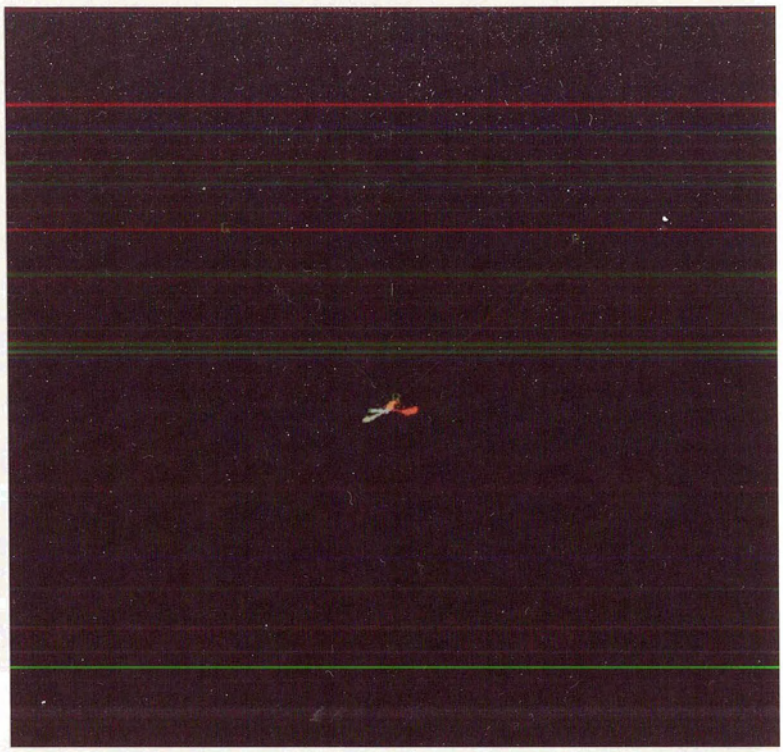

Figure 5.13 : The result of projecting the pixel colors onto the plane orthogonal to the specularity color.

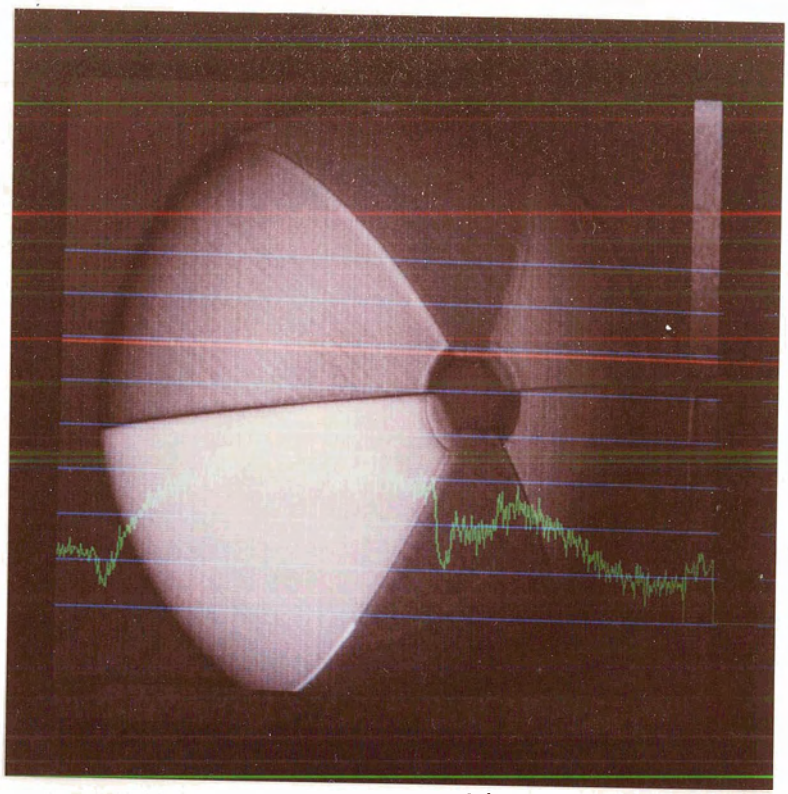

Figure 5.14(a)

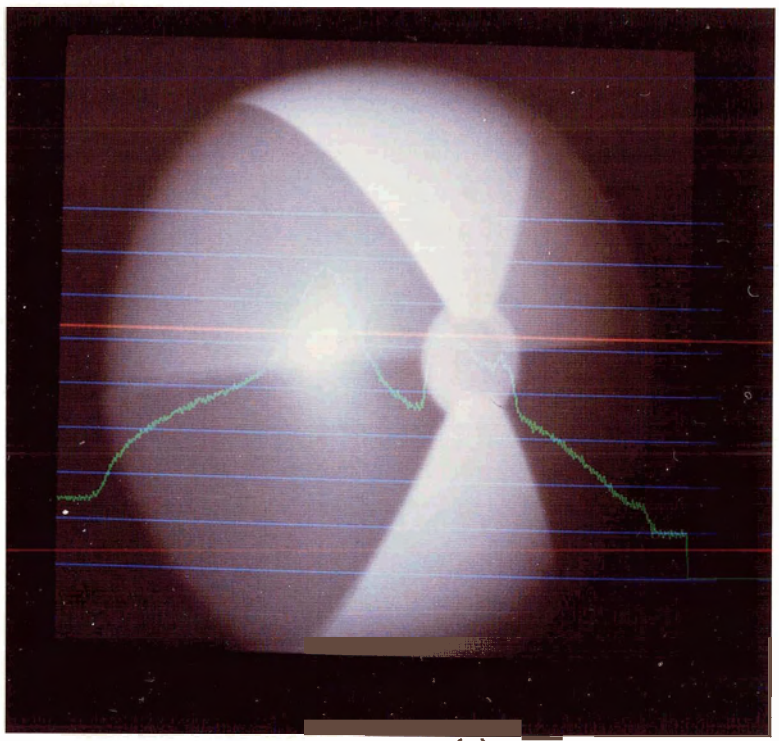

Figure 5.14(b)

Figure 5.14 : (a) is the intensity map of the image in Figure 5.8 with the specular

component removed. (b) is the original intensity map.

Chapter 6
Concluding Remarks

This chapter summaries what we have achieved. Then it reviews our work in relation to some recent results. Following that, we discuss a possible extension of our results in specularity color computation to the problem of highlight detection.

6.1. Research Summary

Based on the Dichromatic Model, light reflected from a surface is projected into color space on a dichromatic plane. We are able to observe the fact that dichromatic planes intersect along a line parallel to the axis of the specularity color. A local method for computing the specularity color has been developed. It avoids the global dichromatic plane calculation, which would otherwise involves difficult image segmentation problems. It is possible to remove the image specularities from the original image after the specularity color has been determined. A projection of the pixel color into a plane orthogonal to the axis of the specularity color readily eliminates all the image specularities.

For the part of shape-from-shading computation, we are able to solve the shape-from-shading problem with unknown source and non-zero ambient illumination. Without assuming a priori knowledge about source location and intensity, but recognizing the regional constancy of body reflectance and ambient component, we handle the problem in a more general setting, compared to the existing algorithms. As we are able to formulate the problem based on local computations, a local relaxation method which enables a parallel network implementation is thus possible.

6.2. Discussion

During the course of my thesis, other researchers have been simultaneously working on related problems and have recently published interesting results. We have observed the fact that dichromatic planes intersect along the axis of illuminant; the papers [KSK87] and [DZLe86] also make use of this observation. We are able to compute the specularity color using local methods so that the need for global segmentation is avoided. Lee [Lee86] also proposes an algorithm for computing the illuminant chromaticity using information from different color regions but requiring no image segmentation, however, his algorithm does not work for the case of ambient illumination.

In spite of the coincidence, we still believe that we have achieved good results in specularity color computation. We argue this in three respects:

(1) the generality of our method.

(2) the global method which involves image segmentation is replaced with local computation.

(3) the ability of our method to handle the ambient illumination.

We also relate our work to recent results on image specularities so as to position our work in the general trend of progress in computer vision on specularity computation.

6.2.1. The Generality of our Method

Our method is general because *(i)* the reflection model on which the method is based is general, and *(ii)* it is not subjected to the glossiness and the shape of the reflecting surface as , for example, Shafer's algorithm is [Shaf84b,KSK87].

As explained in Section 1.3.1, the Dichromatic Reflection Model which we use to understand the relationship between the various reflection components is applicable to a wide class of surfaces, namely the optically inhomogeneous dielectric materials. Most of the partially specular reflectors, including plastics, paints, varnishes and ceramics, are classified into this category. As a matter of fact, these kinds of materials occur so often in our daily life that the generality of the model is well recognized and used to interpret the components in color images. For example, D'Zmura and Lennie in [DZLe86] depicted the reflection response of a surface in a linear model of the diffuse and specular components. Gershon's algorithm [GJT87] detects the color shift phenomenon predicted by the model at the highlight regions. The work in [Lee86] and [KSK87] on specular highlight color computation is developed based on the physical model of inhomogeneous dielectric materials.

Shafer proposed an algorithm in [Shaf84b] for separating the specular and diffuse components. In the work due to Klinker, Shafer and Kanade [KSK87], the algorithm was implemented leading to a successful extraction of the interface and body reflection intrinsic images. However, the method is restrictive (see the comments in Section 3.3.3). Shafer [Shaf84b] comments on the algorithm that the distribution of pixels within the color space parallelogram must not be pathological. By that he meant the pixels should be lying close to the c_i and c_b axes of the parallelogram so that the axes can be calculated accurately. In [KSK87], Klinker, Shafer and Kanade considered the color cluster of the matte and highlight pixels look like a skewed T or comb. The matte and highlight lines are calculated as the two straight edges of the skewed T. Figure 6.1 shows the typical cluster shape.

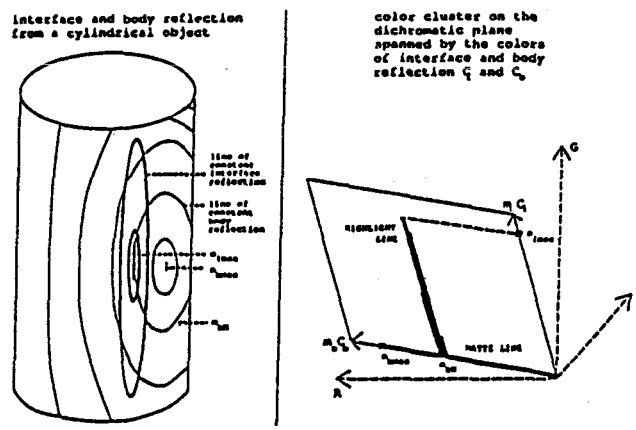

Figure 6.1 : The shape of the color cluster for a cylindrical object. (Figure 2 [KSK87])

We count an argument against the generality of the T-cluster assumption. The Dichromatic Model is general not only because it is applicable to a wide class of reflecting surfaces, but also due to its making no assumption about the surface shape. Naturally, we

will expect the algorithms developed basing on the model also inherit the generality. Consider the T-cluster assumption, the two edges of the T-cluster partitions the pixels into purely matte and purely specular classes. This corresponds to surfaces reflecting purely matte and purely specular light. Therefore, the assumption of skewed-T shape on color clusters inevitably restricts the reflecting surfaces to the class of highly glossy reflectors. In addition, the two edges of the skewed-T must be extending significantly so that the matte and highlight lines can be fitted successfully. This again leads to the requirement for sufficient surface orientations so that the reflected color changes enough to sweep out a distinguishable T in the color space.

Nevertheless, in practice, many surfaces are rough enough to give a broad highlight area. As those highlights are neither sharp nor well-defined, the pixel distribution would be pathological, i.e. the T-cluster assumption is invalid. Besides, very often only part of the surface is imaged. As the whole pixel distribution may not be available in the image, it is likely to run into a pathological data sample again. As a result, these pathological cases lead to failure of the algorithm.

In this respect, our algorithm is relatively more general. We calculate only the dichromatic plane orientation from a color cluster. The specularity color is then obtained by intersecting the local planes. This means that the planarity of the color clusters is sufficient for computing the specularity color. There is no need for extra assumptions about the pixel distribution. As a result, our algorithm is generally applicable to those surfaces described in the Dichromatic Model.

6.2.2. *No Image Segmentation is Needed*

Shafer's algorithm [Shaf84b] assumes the image is segmented a priori into regions of uniform body reflectance. In the implementation of the algorithm [KSK87], the problem of prior segmentation still remains unsolved. The program projects the pixels of a selected image area into the color space and a dichromatic plane is fitted. It then searches within each dichromatic plane for the matte and highlight lines. The crucial step which selects an image region of uniform body reflectance is not done by the program, but done by an image segmentation process performed a priori. However, how the segmentation is done has still not been answered.

D'Zmura and Lennie published their work on mechanisms of color constancy in [DZLe86]. In an attempt to discount the effect of illuminant from the cone signals of a color image, they thought of finding the illuminant using highlights. They observed that surfaces in a scene exhibit different variations in specular and diffuse components, and their response planes intersect along the axis of the illuminant. This could have been the method they used to find the illuminant in order to discount the effect of illuminant so as to achieve color constancy. However, the idea was abandoned as they were held back by the need of image segmentation for determining the loci of responses that correspond to different object. They tended to believe that image segmentation in color vision plays an important role in object discrimination and identification rather than finding the unknown illuminant.

Our algorithm, as described in Section 4.1.1, exploits local information only, it does not require global image segmentation. Without bothering with segmenting the image into regions each of which yields a single dichromatic plane, we calculate the dichromatic planes by letting them be voted by the local planes. In comparison, our method, unlike Shafer's algorithm, is not subjected to the impact of the difficult segmentation problem. Furthermore, we might be able to alleviate D'Zumera's and Lennie's doubt about using segmenta-

tion in finding unknown illuminant from specular highlights.

6.2.3. *The Ability to Handle Ambient Image Component*

As mentioned earlier, Lee was able to propose an algorithm [Lee86] for computing the illuminant color without running into the segmentation problem. However, we argue that our algorithm is more general because we handle ambient image component as well.

Lee's method is also based on the linear reflection model of inhomogeneous materials. As the reflected light is a linear mixture of the specular and diffuse components, in the *CIE(x,y)* chromaticity diagram, the locus of the reflected light from a surface is a straight line connecting the illuminant color point and the surface color point. When several surfaces are imaged, the loci form a radial pattern centered at the illuminant color point. Figure 2 Figure 6.2 shows the *CIE(x,y)* diagram of five differently colored surfaces.

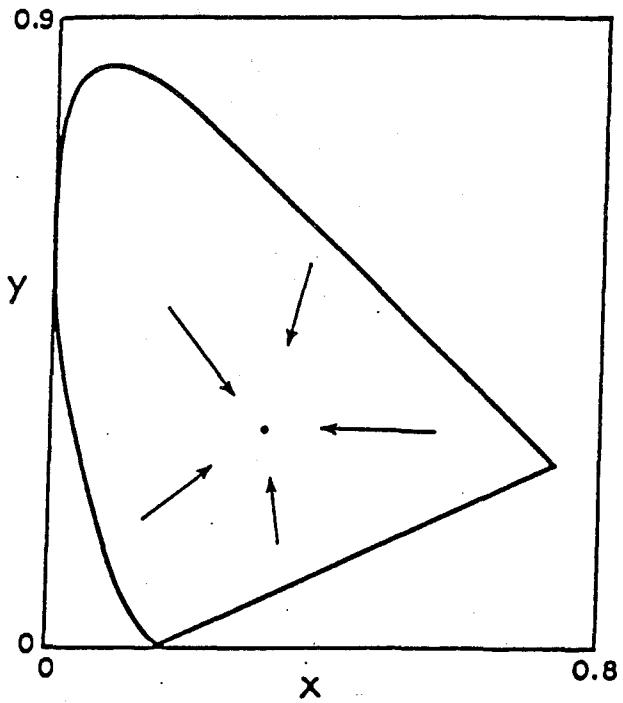

Figure 6.2 : *CIE 1931 x,y* chromaticity diagram showing the ideal loci of chromaticities corresponding to colors from five surfaces of different colors. (Figure 3 [Lee86])

The loci in the *CIE(x,y)* diagram are traced out at the locality of the edge points where the color change is maximum. Since the process is carried out locally near the edge points, no global segementation is required. In this respect, Lee's algorithm shares with our method the credit of requiring no image segmentation. His method, however, fails to handle the ambient image component.

When a surface is imaged under non-zero ambient illumination, the reflected light is added with a constant ambient component. This results in a translational effect on the locus of the color point in the *CIE(x,y)* map. Because different surfaces have different spectral reflectances, they reflect the ambient light in different chromaticities. As a result, their loci in the *CIE(x,y)* map are translated both to different extents and in different directions. The overall radial pattern of the loci, consequently, is no longer preserved and the radial center can no longer meaningfully be used as the illuminant color point.

The major problem of Lee's method is due to the data compression from the tri-chromatic color data into the two-dimensional *CIE* space. The data compression reduces the degree of freedom of the data set from three to two, thus causing loss in information. Our method, on the other hand, uses the three-dimensional color space and does not suffer from the same problem. It is successful in coping with ambient image component (see Section 4.1).

6.3. Related Problem – Highlight Detection

In this section, we consider the use of specularity color computation in another related problem. As discussed in section 3.2, the results of highlight detection [Pell86,GJT87] are not used in this thesis to compute the specularity component. On the contrary, our results in specularity color computation may be used for detecting highlights.

[Pell86] describes a method for detecting highlights as the simultaneous intensity peaks in all the R, G, and B chromatic signals. We argue that the simultaneous peaks criterion is not an exclusive phenomenon of the presence of highlights. For example, a bright white surface marking may reflect more energy in all the chromatic bands and thus will be mistaken for a highlight. However, if we base highlight detection on the physics of reflection, we will have a better understanding of the specular reflection component and thus be able to detect the highlights more accurately.

Gershon's work [GJT87] is based on the dichromatic model of color reflection and thus represents a more accurate method. In the following, we will review Gershon's algorithm and show the possibility of a simpler method when our results in specularity color

computation are applied.

Gershon noticed that in moving from a diffuse region to a highlight area, the specularity color is added to the diffuse color of the surface, resulting in a transition from the diffuse color spectrum to a mixture of the diffuse and specularity spectra. Gershon made use of such a color shift to identify the specular highlights in color images. The color shift appears as a "dog-leg" structure in the C-space [Gers87] with one of the dog-leg points towards the end of the line segment which represents the C-values of perfect reflectors (see Figure 6.3).

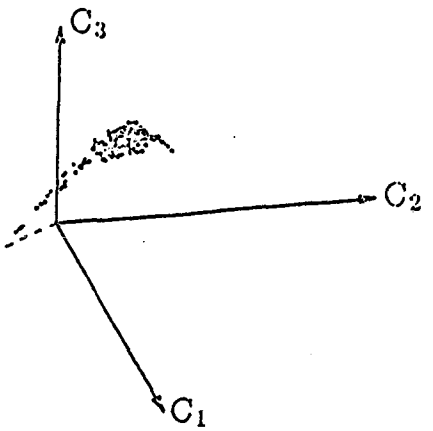

Figure 6.3 : A 3-d scatter plot of a "dog-leg". The dashed line represents the C-values of perfect reflectors. (Figure 2 [GJT87])

As Gershon's method relies on finding the color shift that corresponds to the transition from diffuse to highlight regions, he proposed segmentation as the first step so as to compile the image into uniformly colored regions. Since the "dog-leg" pattern occurs at the diffuse-highlight transition areas, the second step then checks between every pair of adjacent

regions for a "dog-leg" relationship.

In fact, we can have a simpler method to solve the problem. First of all, Gershon's algorithm requires a split-and-merge segementation. Usually, it requires a follow-up adjustment procedure which re-absorbs the spurious small regions resulting from the split-and-merge technique. Since we already have a specularity color algorithm which does not require image segmentation, we prefer to make use of the specularity color to detect highlights. We can employ our algorithm and avoid doing the tedious segmentation process. The purpose of the second part of Gershon's algorithm is to correctly detect the color changes due to specular highlights. As detection of the "dog-leg" pattern involves rather elaborate steps, we would like to replace it with another method as long as the new method also correctly detects the color shift. In fact, we observe that color changes can actually be identified by the shape of the cluster in color space. Therefore, we can look at a color cluster to determine whether it is a color shift due to specular highlights or not. If the cluster shows a linear structure biased towards the specularity color direction, we can conclude that the color is changing due to increasing specular component. That means it is moving into the highlight region.

Thus, we are able to propose a relatively simple algorithm for highlight detection: We use the specularity color computed by the algorithm explained in Section 4.1.1; then lines are fitted to the local clusters; those clusters with their fitting lines running along the axis of specularity color are considered as highlight areas. Although the algorithm looks feasible, it only represents an intuition of using specularity color in highlight detection. Further investigation on its validity and performance is necessary.

References

[Albe72] Albert, A.
Regression and the Moore-Penrose Pseudoinverse
Academic Press, New York, 1972

[BFR81] Burden, R.L., Faires, J.D. and Reynolds, A.C.
Numerical Analysis, 2nd Ed.
Prindle, Weber and Schmidt, Boston, Massachusetts, 1981

[BaTe78] Barrow, H.G. and Tenebaum J.M.
"Recovering Intrinsic Scene Characteristics From Images"
Computer Vision Systems, (Hanson, A.R. and Riseman, E.M. (editor)), 3-26. Academic Press, New York, 1978

[BeSp63] Beckmann, P. and Spizzichino, A.
The Scattering of Electromagnetic Waves from Rough Surfaces
MacMillan, New York, 1-33, 70-98, 1963

[Blin77] Blinn, J.F.
"Models of Light Reflection for Computer Synthesized Pictures"
Computer Graphics, 11, 2, 192-198, 1977

[Blin78] Blinn, J.F.
"Computer Display of Curved Surfaces"
Ph.D. dissertion, Univ. of Utah, Salt Lake City, 1978

[Bril80] Brill, T.B.
Light, Its Interaction with Art and Antiquities
Plenum Press, New York, 1980

[BrHo85] Brooks, M.J. and Horn, B.K.P.
"Shape and Source from Shading"
IJCAI, 1985

[CoHi53] Courant, R. and Hilbert, D.
Methods of Mathematical Physics
Vol.1, Interscience, New York, 1953

[CoHi62] Courant, R. and Hilbert, D.
Methods of Mathematical Physics
Vol.2, Interscence, New York, 1962

[CoTo82] Cook, R.L. and Torrance, K.E.
"A Reflectance Model for Computer Graphics"
ACM Transactions on Graphics, Vol.1, No.1, Jan 1982, 7-24

[DZLe86] D'Zmura, M. and Lennie, P.
"Mechanisms of Color Constancy"
Journal of Optical Society of America A, Vol.3, No.10, Oct 1986, 1662-1672

[EgHi79] Egan, W.G. and Hilgeman, T.W.
Optical Properties of Inhomogeneous Material
Academic Press, New York, 1979

[Gara64] Garabedian, D.R.
Partial Differential Equations
John Wiley and Sons, New York, 1964

[GJT87] Gershon, R., Jepson, A.D. and Tsotsos, J.K.
"Highlight Identification Using Chromatic Information"
Proceedings of First International Conference on Computer Vision, London, England, June 1987, 161-170

[Gers84] Gershon, R.
"Survey on Color: Aspects of Perception and Computation"
Department of Computer Science, Univ. of Toronto, RCBV-TR-84-4, July 1984

[Gers87] Gershon, R.
The Use of Color in Computational Vision.
Ph.D. Thesis, Department of Computer Science, Univ. of Toronto, forthcoming.

[Horn75] Horn, B.K.P.
"Obtaining Shape from Shading Information"

In *The Psychology of Computer Vision* (P.H.Winston, Ed.), McGraw-Hill, New York, 1975

[Horn77] Horn, B.K.P.
"Understanding Image Intensities"
Artif. Intell., 8, No.2, 1977, 201-231

[HoBr86] Horn, B.K.P. and Brooks, M.J.
"The Variational Approach to Shape from Shading"
Computer Vision, Graphics and Image Processing, 33, 174-208, 1986

[Huff71] Huffman, D.A.
"Impossible Objects as Nonsense Sentences"
Machine Intelligence, 6, (Meltzer, R., and Michie, D. (Eds)), Edinburgh Univ. Press, 1971, 295-323

[IkHo81] Ikeuchi, K. and Horn, B.K.P.
"Numerical Shape from Shading and Occluding Boundaries"
Artif. Intell., 17, 1981, 141-185

[KSK87] Klinker, G.J., Shafer, S.A. and Kanade, T.
"Using A Color Reflection Model to Separate Highlights From Object Color"
Proceedings of First International Conference on Computer Vision, London, England, June 1987, 145-150

[Lee86] Lee, H.
"Method for Computing the Scene-Illuminant Chromaticity from Specular Highlights"
Journal of Optical Society of America A, Vol.3, No.10, Oct 1986, 1694-1699

[Long73] Longhurst, R.S.
Geometrical and Physical Optics, 3rd Ed.
Longman, 1973

[Mack73] Mackworth, A.K.
"Interpreting Pictures of Polyhedral Scenes"

Artif. Intell., 4, (1973), 121-137

[Marr82] Marr, D.
Vision
W.H. Freeman and Co., 1982

[PFTV86] Press, W.H., Flannery, B.P., Teukolsky, S.A. and Vetterling, W.T.
Numerical Recipes
Cambridge Univ. Press, New York, 1986

[Pell86] Pellicano, P.N.E.
"Detection of Specularities in Color Images Using Local Operators"
Proceedings of the Graphics Interface / Vision Interface 1986, Vancouver, B.C., May 1986, 370-374

[Pent84] Pentland, A.P.
"Local Shape Analysis"
IEEE Transactions on Pattern Analysis and Machine Intelligence, March 1984, 170-187

[Phon75] Phong, B.T.
"Illumination for Computer Generated Pictures"
Commun. ACM, 18, 6, (June 1975), 311-317

[Prat78] Pratt, W.K.
Digital Image Processing
John Wiley and Sons, New York, 1978

[Shaf82] Shafer, S.A.
"Describing Light Mixtures through Linear Algebra"
J. Optical Soc. Am. 72(2):299-300, Feb 1982

[Shaf84a] Shafer, S.A.
"Optical Phenomena In Computer Vision"
Computer Science Department, University of Rochester, TR-135, March 1984

[Shaf84b] Shafer, S.A.
"Using Color to Separate Reflection Components"

Computer Science Department, University of Rochester. TR, April 1984

[Soho41] Sohon, F.W.
The Stereographic Projection
Chelsea, New York, 1941

[ToSp67] Torrance, K.E. and Sparrow, E.M.
"Theory for Off-specular Reflection from Roughened Surfaces"
J. OPt. Soc. Am. 57, (Sept 1967), 1105-1114

[Wein52] Weinstock, R.
Calculus of Variations, with Applications to Physics and Engineering
McGraw-Hill, New York, 1952

[WeHe66] Wendlandt, W.W. and Hecht, H.G.
Reflectance Spectroscopy
Interscience, New York, 1966