

AD-A262 727



①

**Estimating Scene Properties
by Analyzing Color Histograms
with Physics-Based Models**

Carol L. Novak

December 1992
CMU-CS-92-222

School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213



*Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy*

Thesis Committee:
Steven Shafer, Chair
Takeo Kanade
Katsushi Ikeuchi
Glenn Healey, University of California, Irvine

DISTRIBUTION STATEMENT A

Approved for public release;
Distribution Unlimited

Copyright © 1992 Carol L. Novak

This research was sponsored by the Avionics Laboratory, Wright Research and Development Center, Aeronautical Systems Division (AFSC), U.S. Air Force, Wright-Patterson AFB, Ohio 45433-6543 under Contract F33615-90-C-1465, ARPA Order No. 7597. C.L. Novak was supported by a fellowship from the Air Force Office of Scientific Research under Contract F49620-86-C-0127.

The views and conclusions contained in this document are those of the author and should not be interpreted as representing the official policies, either expressed or implied, of the U.S. Government.

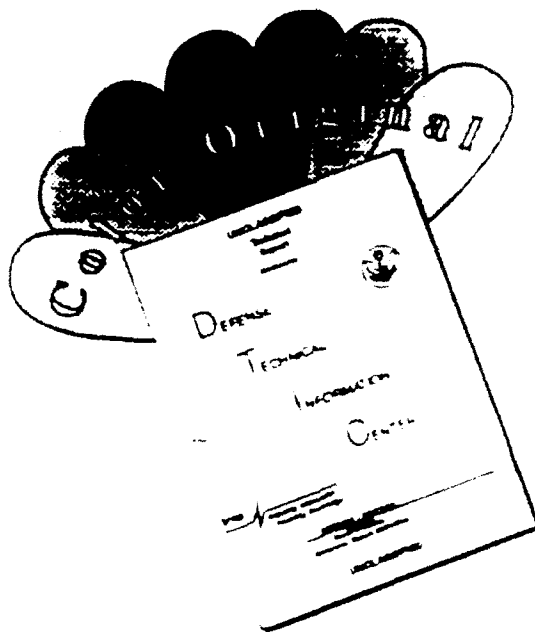
93 4 06 080

93-07198



1497

DISCLAIMER NOTICE



THIS DOCUMENT IS BEST QUALITY AVAILABLE. THE COPY FURNISHED TO DTIC CONTAINED A SIGNIFICANT NUMBER OF COLOR PAGES WHICH DO NOT REPRODUCE LEGIBLY ON BLACK AND WHITE MICROFICHE.

**Carnegie
Mellon**

School of Computer Science

**DOCTORAL THESIS
in the field of
Computer Science**

*Estimating Scene Properties
by Analyzing Color Histograms
with Physics-Based Models*

CAROL L. NOVAK

Accession For	
NTIS GRA&I	☒
DTIC TAB	☐
Unannounced	☐
Justification	
By <i>per letter</i>	
Distribution/	
Availability Codes	
Dist	A. 12 122/or Special
<i>A-1</i>	

Submitted in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy

ACCEPTED:

Steven Ash Shfer

THESIS COMMITTEE CHAIR

4 December 1992

DATE

Meris

DEPARTMENT HEAD

12/10/92

DATE

APPROVED:

T. Try

DEAN

1/26/92

DATE

Abstract

The goal of machine vision is to allow intelligent systems to describe the world around them by the interpretation of images. The difficulty is that vision is a very complex process, since images may contain shadows, highlights, interreflections, and other phenomena. Images are created through the interaction of light with the world; therefore, any vision system that is to understand images must have a model of those interactions. By using physics-based models to describe image formation, we can analyze images in a systematic way.

In applying physical models to machine vision, one of the key tools has been color histogram analysis. A color histogram shows the variation of colors observed within the scene. In the mid-1980s, it was recognized that the color variation for a single inhomogeneous surface can be modeled as a regular physical process with a planar distribution in color space. The identification of this plane and the vectors that define it leads directly to an analysis of object color and illumination color. However there is much more to be said about color histograms. The colors do not fall randomly in a plane, but form clusters at specific points in color space. The colors in the histogram relate not only to the color of the object and the illumination, but also to non-color properties of surface roughness and imaging geometry.

The color histogram for an object illuminated by a single source may be broken into two clusters of pixels, the body reflection cluster and the highlight cluster. The highlight cluster consists of those pixels that correspond to the highlight area of the object, whereas the body reflection cluster consists of those pixels that do not show an appreciable amount of highlight. The length of the highlight cluster is strongly dependent upon the roughness of the object, since smoother objects have brighter highlights. The width of the highlight cluster is also strongly dependent upon roughness, because highlights on rough surfaces cover a greater range of reflectance angles. The position where the highlight and body clusters intersect shows the magnitude of the body reflection component in the highlight pixels, which depends upon the imaging geometry. However, the intersection is also affected to some degree by the object's roughness, whereas the length and width measurements are in turn affected by imaging geometry. An understanding of the interplay of these factors also shows how they affect estimates of the illumination color. The result is a set of interdependent relationships between scene parameters and histogram features.

Furthermore, analysis of color histograms may be extended to images that contain interreflection. Applying the dichromatic reflection model to interreflection between two objects shows how color, imaging geometry, and surface roughness affect the color histogram. By measuring the dimensions of the resulting clusters, additional information about scene characteristics may be obtained that is not available in simpler scenes.

This thesis describes an algorithm for analyzing color histograms that yields estimates of the properties of object roughness, object albedo, and object color, as well as illumination position, illumination intensity and illumination color. The relationship between these scene parameters and the histogram measurements is complex and cannot be solved analytically. Instead, the solution is obtained by interpolating between histogram measurements from images with known scene parameters. The algorithm is tested on simulated images, and the values recovered by the program compare very favorably with the actual parameters used to simulate the scene. The algorithm is also tested on real images, and the calculated values are reasonably close to estimates of the scene parameters obtained by other means.

The algorithm presented here works very quickly, and requires only a single color image. This method may be applied to such varied tasks as surface inspection and object recognition. It demonstrates the importance of considering laws of reflection in explaining color variation, and shows why color is such a valuable feature. This type of analysis of color appearance brings us closer to the goal of allowing machines to interpret images.

Acknowledgments

I owe many people a heart-felt thank you for all their help and support over my years here. First of all, I want to thank my advisor, Steve Shafer. He has given me wonderful advice on writing, speaking, and researching. He has tirelessly corrected draft after draft of my papers, including this thesis. He is always full of ideas for solving problems, and always received my own ideas and results with great enthusiasm. Steve is a great teacher in every way.

I also want to thank the other members of my committee, Takeo Kanade, Katsushi Ikeuchi, and Glenn Healey. They each have provided very insightful advice, by bringing their own perspective to this research. I wish to thank all the members of the VASC (Vision and Autonomous Systems Center) group, who together have provided a fantastic environment for doing research. It has been a great experience to work with such a talented (and fun) group of people.

I owe many thanks to John Krumm, Reg Willson and Ron Stone who are fellow users of the Calibrated Imaging Lab (CIL). Reg wrote a huge amount of software to run the lab, which I used to take the pictures for this thesis. All three of them have provided great discussions on technical issues, as well as a lot of encouragement. I must also express my appreciation to Jim Moody and Bill Ross, our group's technical support. They have cheerfully helped me track down many, many hardware and software problems.

I want to thank all my office-mates, past and present. Gudrun Klinker, Larry Matthies, Carlo Tomasi, Jill Crisman, Keith Gremban, Jennie Kay, Conrad Poelman, and Todd Williamson were always able to find the answer for any question, from undocumented software "features" to mathematical derivations. They have been great at providing advice and support over the years.

Everyone else in the School of Computer Science also deserves a thank you for providing such a great department for graduate students to become computer scientists. I want to thank Sharon Burks (who I am convinced has never had a bad day) and everyone else on the fourth floor who has been so helpful at handling all the practical details of this department. I also want to thank my friends and fellow students Allan Heydon and Bruce Horn for all the moral support they've provided over the years.

Finally I want to thank my husband, Hunter Gordon, who has always provided the support, advice, and enthusiasm to keep me going. Both of our families also deserve recognition for their encouragement and patience over the years.

To my parents, who inspired me to become a scientist;
To my advisor, who showed me how;
And to my husband who never let me doubt that I could do it.

Contents

1	Introduction	17
1.1	The Use of Color Histograms in Image Understanding	18
1.2	Physics-Based Models for Image Understanding	21
1.3	A More Complete Analysis of Color Histograms	25
1.4	Thesis Outline	27
2	Physics-Based Models	29
2.1	Dichromatic Reflection Model	29
2.2	Body Reflection Model	32
2.3	Surface Reflection Model	32
3	Structure of the Color Histogram	35
3.1	The Body Reflection Cluster	36
3.2	The Highlight Cluster	38
3.2.1	Direction of Highlight Cluster	39
3.2.2	Length of Highlight Cluster	42
3.2.3	Width of Highlight Cluster	46
3.2.4	Intersection Point of Highlight and Body Reflection Clusters	51
4	Color Histogram Structure when Interreflection is Present	57
4.1	Modeling Interreflection	57

4.2	The Secondary Highlight Cluster	59
4.2.1	Length of the Secondary Highlight Cluster	61
4.2.2	Direction of the Secondary Highlight Cluster	62
4.2.3	Intersection Point of the Secondary Highlight Cluster	62
4.3	The Interreflection Cluster	67
4.3.1	Surface to Body Reflection	67
4.3.2	Body to Body Reflection	67
4.3.3	Body to Surface Reflection	71
4.3.4	Analysis of the Interreflection Cluster	72
5	Analysis of the Color Histogram for an Ideal Sensor	75
5.1	Properties of an Ideal Sensor	75
5.2	Knowns and Unknowns	77
5.3	Analytic Solution	79
5.4	Ideal vs. Measured	80
5.5	Approximate Solution	87
5.6	Generating Lookup Tables	89
5.7	Calculating Roughness, Phase Angle, and Illumination Intensity	92
5.8	Calculating Illumination Chromaticity	95
6	Extending the Analysis to a Realistic Sensor Model	103
6.1	Modeling Camera Noise	104
6.2	Modeling Camera Clipping	107
7	Application of the Algorithm to Real Images	111
7.1	Additional Problems with Real Cameras	111
7.1.1	Nonlinearity and Color Imbalance	112

7.1.2 Chromatic Aberration	113
7.1.3 Blooming	116
7.2 Estimating Phase Angle	117
7.3 Estimating Illumination Intensity	120
7.4 Estimating Roughness	123
8 Summary and Conclusions	125
8.1 Contributions	125
8.2 Directions for Future Research	127
8.3 Applications	129
8.4 Concluding Remarks	130
A Color Figures	131

•

•

•

•

List of Figures

1.1	Gaussian histogram as described by Haralick and Kelly, 1969	19
1.2	Planar histogram as described by Shafer, 1984	19
1.3	T-shaped histogram as described by Klinker, 1987	20
2.1	Process of dichromatic reflection	30
2.2	Angles that define imaging geometry	31
2.3	Dichromatic reflection components in a single pixel	31
2.4	The facet model of surface reflection	33
3.1	Histogram of a single object	36
3.2	Dependence of highlight cluster direction on object roughness	40
3.3	A skewed highlight	41
3.4	Dependence of highlight cluster length on object roughness	43
3.5	Dependence of length on roughness for different indices of refraction n	44
3.6	Dependence of length on imaging geometry	45
3.7	Dependence of length on roughness for different imaging geometries	45
3.8	Dependence of highlight cluster width on object roughness	47
3.9	Dependence of width on incident light intensity	49
3.10	Illustration of why width varies with imaging geometry	49
3.11	Dependence of width on roughness for different imaging geometries	50
3.12	Illustration of how highlight cluster varies with changing imaging geometry	52

3.13	Dependence of intersection point on imaging geometry	53
3.14	Dependence of intersection on geometry for different roughness values	54
3.15	Dependence of histogram features upon scene parameters	55
4.1	Interreflection components, by path	59
4.2	Interreflection components, by hue	60
4.3	Histogram for interreflection	60
4.4	Location of a secondary highlight	63
4.5	Angle at which secondary highlight occurs	64
4.6	Dependence of secondary highlight intersection point upon geometry	65
4.7	Histogram where primary and secondary highlight clusters have crossed over	65
4.8	Dependence of intersection upon geometry for different roughness values	66
4.9	Spectral reflectances of four colors	68
4.10	The color overlap and its effect	70
5.1	Skewed length measurement	81
5.2	Ideal vs. measured length	82
5.3	Variation in ideal length with changes in roughness and phase angle	82
5.4	Variation in measured length with changes in roughness and phase angle	83
5.5	Histogram with a negative intersection	84
5.6	Variation of ideal intersection values	85
5.7	Variation of measured intersection values	85
5.8	Histogram with a width greater than one	86
5.9	Variation of ideal width values	86
5.10	Variation of measured width values	87
5.11	Surface of roughness and phase angles within L-W-I space	89
5.12	Range of scene parameters used to generate lookup table	90

5.13	Finding the body cluster	90
5.14	Goodness-of-fit for polynomial functions	93
5.15	Algorithm for calculating roughness, phase angle, and illumination intensity	94
5.16	Results on noise-free data	95
5.17	Variation in cluster skew with changes in roughness and phase angle	96
5.18	Results for calculating skew	97
5.19	Histogram of simulated image	98
5.20	Results for test image	99
5.21	Algorithm for recovering scene parameters	99
6.1	Results on noisy data ($\sigma_c = 1.275$)	106
6.2	Results on noisier data ($\sigma_c = 2.55$)	106
6.3	Purpose of second width measurement	108
6.4	Results on noisy data when clipping is present	109
6.5	Errors for three camera models	109
7.1	Sketch of the effects of chromatic aberration on a simple histogram, as described by Willson and Shafer, 1991	114
7.2	Experimental setup for estimating phase angle	118
7.3	Plot of intersection point vs. phase angle	118
7.4	Results for experiments to calculate phase angle from real images	119
7.5	Measured illumination intensity for each image in the sequence	120
7.6	Results for calculating illumination intensity from real images	121
7.7	Calculating roughness when intensity changes	122
7.8	Calculating phase angle when intensity changes	122
7.9	Curve used for calculating luminance	123
7.10	Table of estimating roughness results	124

A.1 Histogram of shiny billiard ball	133
A.2 Histogram of matte billiard ball	133
A.3 Simulation results	135
A.4 Simulated test image	137
A.5 Image simulated from recovered parameters	137
A.6 Intrinsic images calculated with a skewed highlight color	139
A.7 Intrinsic images calculated with the corrected highlight color	139
A.8 Histogram for shiny ball	141
A.9 Histogram for very shiny pail	141
A.10 Closeup of bloomed area	143
A.11 Image with phase angle of 10 degrees	145
A.12 Image with phase angle of 90 degrees	145
A.13 Histogram for phase angle of 90 degrees	147
A.14 Image of plastic alligator	147
A.15 Image of plastic pumpkin	149
A.16 Five objects with different roughness values	149

Chapter 1

Introduction

The goal of machine vision is to allow intelligent systems to perceive the world around them. Vision is a very rich sense which provides much information, yet it is also complex due to shadows, highlights, interreflections, and other phenomena. In order to understand how light created a particular image, a vision system will need a model of the way light interacts with the world. Therefore, a continuing area of research in machine vision has been the development of ever more sophisticated models of image formation, that are complex enough to capture the richness of optical phenomena seen in real images, and yet simple enough that they can be inverted to recover information about scenes. Since image formation involves the projection of a three dimensional scene into two dimensional image data, the problem is of course under-constrained. Yet the human visual system is able to recover an astonishing amount of scene information from images. So far, machine vision is not nearly so sophisticated, but researchers have made a lot of progress in showing how images may be automatically interpreted to recover useful information about the scene. Color images have proved especially useful for analyzing images for such tasks as object classification and scene segmentation.

This thesis will show how color information from an image can be used to recover interesting properties of the scene. These properties include the color of objects and the color of the illumination, but also non-color properties such as surface roughness and illumination position. This is done by exploiting the structure inherent in color histograms. This thesis will show that the dimensions of the color histogram are related in a precise, quantitative way to scene properties; and that measurements of histogram dimensions may be used to recover estimates of scene parameters. The result is a program that automatically calculates these scene parameters using only a histogram from a single color image.

1.1 The Use of Color Histograms in Image Understanding

Color histograms have long been used by the machine vision community in image understanding. Color has been thought of as an important property of objects, since it can be used for segmentation and classification. Unfortunately color is not uniform for all objects of a similar class, nor even across a single object. That would make the segmentation and classification problem too easy! Color variation has come to be expected in images, and thus vision researchers have been working on modeling this variation.

The color histogram provides a method for representing color variation in a straight-forward way. A histogram is created from the color values at each pixel; it shows for each point in color space how many pixels exhibit that color. In many images, the colors tend to form clusters in the histogram, one for each object in the image. The earliest uses of color histograms modeled the histogram as a Gaussian cluster in color space [Haralick and Kelly, 1969]. By manual or automatic analysis of the histogram, the shape of each cluster is found. Then each pixel in the image is assigned to the cluster that is closest to the pixel color in color space. For example, in a road following task, pixels that correspond to grass could be modeled as having a canonical color of green with some possible deviation from this color. The color variation is then modeled as a probability distribution, so that the further from the characteristic color, the less likely that it is a grass pixel. The problem then is to choose the size and shape of the distribution. If the size is too large, many non-grass pixels will be incorrectly included; if it is too small, a large number of grass pixels will be wrongly rejected. Figure 1.1 shows a diagram of this idea. This type of classification of regions by color histogram analysis is being used today in robotics applications such as autonomous navigation [Crisman and Thorpe, 1991]. However, this approach assumes a random variation in color, and so would not be well suited to *systematic* variations in colors such as would be associated with shading or highlights.

A more powerful approach for handling systematic variations in color is to consider the physics of light reflection from objects. In 1984 Shafer showed that for dielectric materials with highlights, the color histogram associated with a single object forms a plane [Shafer, 1984]. This plane is defined by two color vectors: a body reflection vector and a surface reflection vector. In a diagram accompanying this idea (figure 1.2), he visualized that the histogram would fill out a parallelogram. The paper did not describe how to determine these two vectors. In fact, for any planar histogram, there are many sets of two vectors that could define it.

However, in 1987 Klinker and Gershon independently observed that the color histogram does not

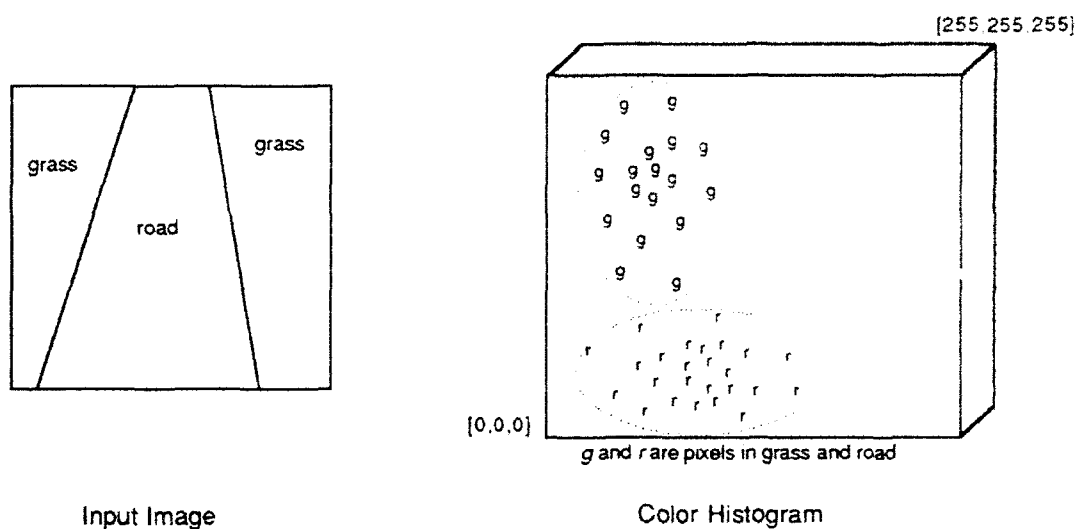


Figure 1.1: Gaussian histogram as described by Haralick and Kelly, 1969

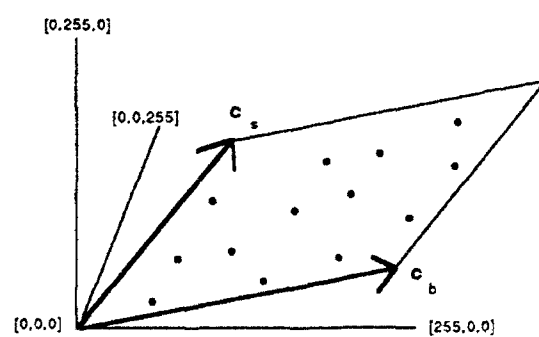


Figure 1.2: Planar histogram as described by Shafer, 1984

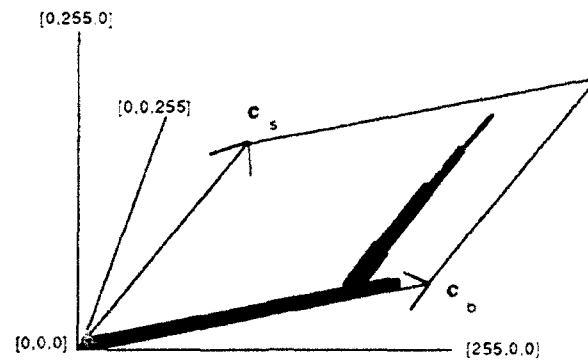


Figure 1.3: T-shaped histogram as described by Klinker, 1987

uniformly fill the plane, but instead forms a T-shape or dog-leg in color space [Klinker *et al.*, 1987], [Gershon, 1987]. Figure 1.3 shows that this color histogram is composed of two linear clusters, one corresponding to pixels that exhibit mostly body reflection and one corresponding to pixels that exhibit mostly surface reflection. This T-shape made it possible to identify characteristic body reflection and illumination colors.

In 1988, Healey showed that the number of dimensions occupied by the histogram may be used to determine material type [Healey, 1988]. Metals do not exhibit any body reflection; they show only surface reflection. Thus the color histogram of a metal object will have only a one-dimensional distribution in color space, in contrast to the two dimensional distribution for dielectrics with highlights. Examining the color histogram allows the automatic classification into metal and non-metal.

The shape of color variation has also been used for recognition tasks. In 1990 Swain used color histograms for indexing into a large database of multi-colored objects [Swain, 1990]. The distribution of colors on each object is invariant with rotations and translations in the image plane, and changes slowly with changes in view angle and scale, allowing the histogram in an image to be matched to the appropriate model in the database. Funt and Finalyson have proposed histogramming color *ratios* instead of color values to allow for changes in incident illumination color [Funt *et al.*, 1991]. Although these histogram techniques allows real-time identification and location of complex patterned objects as cereal boxes and striped clothing, they do not make any other analysis of scene characteristics.

There has also been some work in analyzing histogram shape when interreflection is present. Funt and Drew have observed that for interreflection between diffuse surfaces, the pixels for each surface will lie on a plane defined by the color vector of the object's color and the color vector defined by the interreflection color [Funt and Drew, 1991]. Under some circumstances, the two objects will have different planes. If

the intersection of the two planes can be found, the interreflection can automatically be removed from the pictures. The method does not require shape information about the surfaces, other than the assumption that they are convex. Since the objects are assumed to be diffuse surfaces, the method does not consider highlights or interreflection involving highlights.

Bajcsy and Lee have described an algorithm that segments images that contain both highlights and interreflection [Bajcsy *et al.*, 1989]. The analysis is done on two-dimensional histograms in a color space similar to chromaticity. First color constancy is performed using a white reference object. After this step, a highlight on an object will have the same hue as the body color, so when the image is segmented by hue, highlight and body regions will be kept together. Since the highlight is a whitened region (after color constancy) it will have a different saturation value than the body region, so any highlighted areas may be identified by a further segmentation by saturation. Interreflection is recognized when it causes small deviations from the object's hue. The algorithm has been applied to actual images; however in these scenes the hue did not vary significantly in interreflection areas. Other types of images can show dramatic hue shifts due to interreflection.

Color histograms have been used for many years in image understanding research, and have proved very useful in a variety of tasks. Several researchers have used physics-based models to explain the relationship between scene characteristics and histogram shape. However, until now, physics-based models have not been fully exploited to explain the detailed structure of the histogram.

1.2 Physics-Based Models for Image Understanding

Physics-based models are appealing since they describe a systematic relationship between scene parameters and image values. This approach makes it possible to analyze image features, such as highlights, which cannot be modeled well with statistics. Physics-based vision has become increasingly popular in the last few years and has been applied to a variety of tasks.

A physics-based model used extensively in color vision research is the "dichromatic reflection model", first proposed in [Shafer, 1984]. The model states that the light reflected from a dielectric object is the sum of two components: a surface component and a body component. The model further states that each component in turn may be separated into a color term that depends only on wavelength, and a geometric

term that depends only upon viewing geometry. This equation gives rise to the planar distribution predicted by Shafer [Shafer, 1984] and observed by both Klinker [Klinker *et al.*, 1987] and Gershon [Gershon, 1987].

The spectral component of the surface reflection is often modeled as being the same as that of the incident light. This assumption is sometimes called the Neutral Interface Reflectance (NIR) model [Lee, 1988]. The spectral component of the body reflection is modeled as being the product of the object color and the illumination color. This means that if the illumination color can be identified, then its influence may be divided out of the body color to give the object's reflectance. This provides a method for color constancy under changing illumination conditions [Klinker, 1988], [D'Zmura and Lennie, 1986], [Healey and Binford, 1987a], [Tominaga and Wandell, 1989].

Recently Sato and Ikeuchi have analyzed image sequences of shiny surfaces taken while the light source is moving [Sato and Ikeuchi, 1992]. Each pixel is individually separated into body reflection and surface reflection components using singular value decomposition, and the surface normal at that point is recovered. The method analyzes the rank of the matrix formed from the sequence of color values; since each pixel from a dielectric is composed of two reflection components, the rank will be two. The method also works on metal objects, when the rank of the matrix is one (since metals only exhibit surface reflection). However the method requires a sequence of images with lights at different, known positions.

The roughness of an object also has an important effect on its appearance. When something appears shiny, it is usually assumed that the surface is optically very smooth. The amount of shininess is usually associated with the brightness and sharpness of highlights. Although several people have worked on analyzing surface roughness from images, no work has previously been done on how roughness affects the appearance of color histograms. This is not surprising since the previous work in roughness has mainly focussed on black and white images.

In order to analyze surface roughness from images, a particular model of surface roughness and its affect on appearance is needed. In [Mundy and Porter, 1980] and [Porter and Mundy, 1981] the roughness of metal surfaces is estimated using the Beckmann-Spizzichino model [Beckmann and Spizzichino, 1963]. The work of [Healey and Binford, 1987b] used the Torrance-Sparrow model to analyze highlights and determine local shape from highlights [Torrance and Sparrow, 1967]. However, this method did not provide an estimate of surface roughness.

In [Nayar *et al.*, 1991], the Beckmann-Spizzichino model is compared with the Torrance-Sparrow; it is

shown that they predict essentially the same surface reflection when surfaces are reasonably rough. This work also proposes a hybrid model that incorporates both these roughness models. In [Ikeuchi and Sato, 1990] and [Solomon and Ikeuchi, 1992] roughness parameters are estimated using the Torrance-Sparrow model. Shape information is provided by a range finder [Ikeuchi and Sato, 1990] or by 4-light photometric stereo [Solomon and Ikeuchi, 1992], and the model parameters are determined using the least squares fit to the image brightness and shape data. These techniques all consider object shape in recovering roughness estimates.

Interreflection presents a very difficult problem for machine vision, and it is only recently that it has been studied in actual images. However, it too can be explained with physical models. The theoretical nature of the problem was examined a considerable time ago by Horn. In [Horn, 1977] a derivation is given for the interreflection of two semi-infinite planes that intersect at right angles. These requirements are relaxed to allow the planes to meet at angles other than 90 degrees and to allow the planes to have a finite length from the intersection (although still infinite in the direction along the intersection). The calculation is said to closely approximate the actual behavior close to a corner for real intersections of flat objects. However the usefulness of this derivation is limited by the assumption that the surfaces are Lambertian.

Another theoretical application of physical models to interreflection is a variation of the dichromatic reflection model [Brill, 1989]. The idea is that where the number of sensors is one greater than the number of different types of reflection functions, the range of colors seen on each object will span a (hyper)plane in sensor space. Therefore the normal to this plane may be used to segment a scene. Although the main focus of the paper is upon matte objects with two light sources, it mentions that the theory is equally applicable to the situation where one light source illuminates a dielectric object with a highlight (the dichromatic reflection model), and to the situation where a matte object is illuminated by a single light source plus interreflection from a nearby (matte) object. This work is theoretical in nature and has not been applied to any actual images containing interreflection.

The observation has been made that outdoors where the sun is yellow and the sky is blue, shadows tend to look bluish [Sloan, 1977]. Although this is a very qualitative statement, it implicitly contains the idea that shadow regions are lit by interreflection from the atmosphere and that this effect is more important in areas not lit by direct illumination from the sun.

The idea that interreflection plays an important part in the appearance of shadow regions is also given in [Forsyth and Zisserman, 1989]. In this paper, an equation for interreflection is derived and the simulated

results are compared with the results of actual images. The authors also show that interreflection is more pronounced for objects with high albedo (white) than for objects with low albedo (black). However this work looked at Lambertian reflection only. Shiny black objects will usually show considerable interreflection of other objects in the scene.

In 1990, Nayar demonstrated the usefulness of physical models for recovering scene information from real images with interreflection [Nayar *et al.*, 1990]. This paper presents an algorithm that determines shape and albedo of objects, and that explicitly predicts and accounts for interreflection. An initial estimate of surface shape and reflectance is made without accounting for interreflection. That initial estimate is used to calculate the amount of expected interreflection by using the radiosity method [Goral *et al.*, 1984]. This in turn allows a better estimate of shape and reflectance. The process is repeated until the solution converges. The algorithm uses black and white pictures rather than color, and uses multiple exposures with the light source in different positions to obtain shape information. It also relies upon the objects being matte, without any highlights to cause interreflection.

As mentioned in section 1.1, Funt and Drew have considered interreflection between matte surfaces. Their work uses a "one-bounce" model, considering light that is interreflected at most once between surfaces. In their terminology, light that comes from the illumination and is reflected by an object directly into the camera is called "no-bounce". They analyzed interreflection using color-space analysis of histogram [Funt and Drew, 1991] and by using the calculus of variations [Drew and Funt, 1992]. Jang also analyzed interreflection between matte objects, using a model that considers at most once bounce of interreflection [Jang, 1991]. Both this method and those of Funt and Drew rely upon having a pre-segmented image and identification of the colors involved in interreflection ahead of time.

Only the work of Lee considers interreflection among shiny objects that exhibit highlights [Lee, 1991]. The model used is an extension of the dichromatic reflection model, that considers the four types of interreflection that are inherent in a one-bounce model between shiny surfaces. Although all four types are listed, their relative strengths are not addressed. It also assumes that the hue change caused by interreflection is very small, which is not true in many images. This work did not consider the effect of roughness on interreflection and provided no quantitative analysis of interreflection.

Physics-based models have proved very useful in machine vision research, for tasks ranging from analyzing color to estimating roughness. However, physics-based models may be used in a more complete way to make more detailed predictions about color appearance, and to recover more information about the

scene.

1.3 A More Complete Analysis of Color Histograms

One problem with previous work in analyzing color has been that it did not fully consider the effect of non-color parameters on color appearance. It is immediately obvious that the color of an object has an effect upon the color observed in an image of that object. A red object is expected to appear red, a yellow object is expected to appear yellow, and so forth. The study of color constancy in machine vision has emphasized the fact that the color of illumination also has an effect on color appearance. People outside the color and vision communities are often unaware of this, since human color vision is advanced enough to compensate for almost all changes in illumination. Nevertheless, the impact of illumination color on image appearance is obvious to those who work with color images, and is a logical consequence of the physics of color reflection.

This thesis addresses the way other scene parameters affect the color distribution across objects. Since optical roughness determines whether a surface looks matte or shiny, it is obvious that it affects scene appearance in some way. However previous researchers that have looked at roughness have only considered its effect on image intensity. By considering the combination of both body and surface reflection components in color images, this thesis shows that smooth surfaces exhibit colors that are not seen anywhere on rough surfaces and vice versa.

Similarly, it has long been known that highlights move when imaging geometry changes. This has posed substantial problems for stereo algorithms since the two cameras are viewing objects in the scene from somewhat different angles. This makes it difficult to find correspondences between the two images since highlights are not a stable feature. This thesis will show that a change in imaging geometry not only changes the position of highlights, but their actual colors as well. The colors observed in a highlight viewed from straight on are actually not present when the same highlight is seen from an oblique angle and vice versa.

Since the observed colors are affected by both surface roughness and imaging geometry, the obvious hope is that the color can be analyzed to recover those scene parameters. Physics-based reflection models show that the color varies in a systematic way as roughness and imaging geometry are varied. This thesis will show how this systematic variation can be exploited to recover these scene parameters.

An understanding of the effects of roughness and imaging geometry upon color distribution also show

how to estimate the illumination color from the image of a single object with highlights. Although it has long been recognized that the pixels from this object will lie in a plane defined by the body reflection and surface reflection colors, the problem has been in identifying these two vectors from the infinite set of vectors that lie in the dichromatic plane.

This problem can be solved if there is more than one object with highlights in the scene: then the surface reflection vector can be identified by the intersection of the dichromatic planes. If there is only one object, a vector may be obtained by the best fit to the highlight cluster in the color histogram. This vector was used by Klinker to approximate the surface reflection cluster [Klinker, 1988]. The degree to which this approximation is accurate depends upon the surface roughness and imaging geometry. Varying these parameters has a systematic effect upon the cluster appearance, which may be analyzed to recover an improved estimate of the illumination color.

Research in color images has traditionally only been able to recover a *relative* color measurement for objects and illumination. Estimates of object color considered only chromaticity, not absolute reflectance, since it was considered impossible to tell a bright source illuminating a dark object from a dim source illuminating a light object. However by considering the appearance of highlights, it is possible to show that there is a definite difference in appearance for these two scenarios. This thesis will show how the difference may be exploited to distinguish between object albedo and illumination brightness.

Whenever there is more than one object in the scene, there is the possibility for interreflection between the two objects. The usual assumption made in image understanding is that a scene is illuminated by a single light source which emits light onto the scene. This light is reflected by individual objects in the scene, and some of that light enters the camera. In this model the energy reflected by an object does not interact with the light source nor with other objects in the scene. However, even simple pictures with a very small number of objects often exhibit interreflection.

Most of the work on the problem of interreflection has only considered light which is interreflected once between matte objects. This thesis will refer to this as the "secondary reflection" model, since it considers light that is reflected from two surfaces. (By contrast Funt refers to this as "one bounce" since it is interreflected only once [Funt and Drew, 1991].) More than two reflections are obviously possible, but the effect of additional reflections upon the image are shown here to be small for inhomogeneous dielectrics.

The matte assumption is appealing since researchers only have to consider two color components:

primary body reflection that comes directly from the object, and secondary body reflection that came first from the body reflection of a nearby object. As Funt points out, these two components mean that pixels form a plane, somewhat analogous to the dichromatic plane [Funt and Drew, 1991].

However many objects that exhibit interreflection have some amount of shininess. This shininess can have a noticeable effect upon the observed interreflection. Therefore this thesis extends the Dichromatic Reflection Model to the case of secondary interreflection. By considering the body and surface reflection components separately, secondary reflection adds four additional color components to the two that come from primary reflection. The relative strengths of these components determine the degree to which interreflection is noticeable in images.

Once highlights are considered in interreflection, it becomes necessary to consider the amount of shininess in order estimate the relative strengths of the interreflection components. Therefore the surface roughness is an important parameter in analyzing interreflection. The "hue shifts" that are observed in interreflection will depend upon objects' surface smoothness as well as their colors. This thesis will show how considering the components of interreflection due to surface reflection leads to a better understanding of object appearance, and may give some additional information that cannot be recovered from simpler scenes without interreflection.

1.4 Thesis Outline

This thesis shows how the structure of color histograms may be exploited to recover valuable scene information. It begins in chapter 2 with an explanation of the reflection models used in this work. Chapter 3 describes how the histogram for a single object is related to various scene parameters. Chapter 4 extends this to show how the histogram is affected by interreflection from a nearby object. These two chapters show how scene parameters affect the color distribution in images. This type of description is very valuable, but does not necessarily mean that it can be applied to vision algorithms. In chapter 5, this thesis begins describing how the color distribution can be analyzed to recover scene parameters from color variation. Chapter 5 shows how the color histogram may be analyzed if an idealized sensor could be employed. Chapter 6 extends the analysis for a more realistic camera that has characteristics similar to those used by vision researchers today. Chapter 7 shows results in estimating scene properties when the algorithm is applied to real images. Chapter 8 summarizes the work of this thesis, discusses the contributions made, and points out directions of

future research.

Chapter 2

Physics-Based Models

Much of machine vision has evolved from classical signal processing. This approach treats image data as an arbitrary 2-dimensional function $I(x,y)$. When the image data is viewed this way, it is natural to take derivatives, apply Fourier transforms, or use statistical models. This idea has had a lot of success in areas such as edge detection.

The problem with this approach is that for real images the function is not arbitrary, but depends upon physical laws. The goal of an image understanding system is not really to find edges in the image, but to find edges in the world. A vision model based on the physics of how highlights are formed will have a much better chance interpreting them.

This chapter explains the physics-based models used in this thesis. The first section describes a reflection model for objects that exhibit both matte reflection and highlights as a result of two separate reflection processes. Then the following sections describe reflection models for each of these two types of reflection.

2.1 Dichromatic Reflection Model

A physics-based model used extensively in color vision research is the "Dichromatic Reflection Model", first proposed in [Shafer, 1984]. The model states that the light L reflected from a dielectric object is the

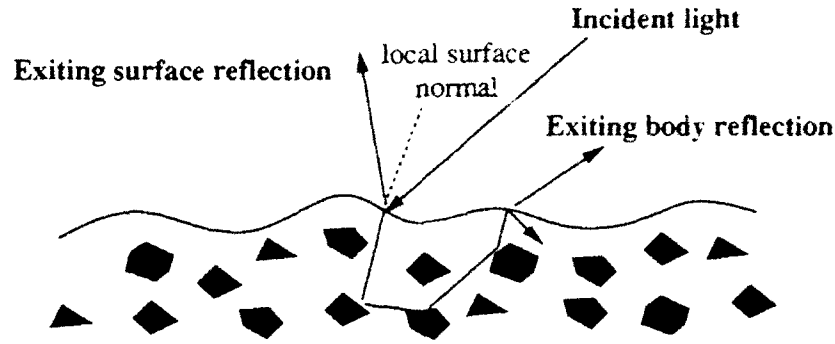


Figure 2.1: Process of dichromatic reflection

sum of two components: a surface component L_s and a body component L_b .

$$L(\lambda, \theta_i, \theta_r, \theta_p) = L_b(\lambda, \theta_i, \theta_r, \theta_p) + L_s(\lambda, \theta_i, \theta_r, \theta_p)$$

This corresponds to the two different ways in which light is reflected by inhomogeneous dielectrics. These surfaces are typically composed colorant particles dispersed throughout a clear medium. This model applies to plastics, painted surfaces, ceramics, paper, cloth, etc. The model does not apply to homogeneous surfaces such as metals, which exhibit only surface reflection. When light passes from the air into the medium of an inhomogeneous dielectric, a change in the index of refraction causes some of the light to be immediately reflected back into the air. The rest of the light is refracted into the surface where it is bounced around between colorant particles. These particles selectively absorb certain wavelengths of light, giving rise to the object's characteristic color. A diagram of this process is shown in figure 2.1.

Each of the two components is a function of the wavelength of light (λ) and the angles of incidence (θ_i), reflection (θ_r) and phase angle (θ_p). These angles are pictured in the diagram in figure 2.2. θ_s is the angle between the surface normal and the perfect specular direction (the bisector of the incidence and reflectance directions).

The Dichromatic Reflection Model further states that each component in turn may be separated into a color term c that depends only on λ , and a geometric term m that depends only upon θ_i , θ_r , and θ_p .

$$L(\lambda, \theta_i, \theta_r, \theta_p) = m_b(\theta_i, \theta_r, \theta_p)c_b(\lambda) + m_s(\theta_i, \theta_r, \theta_p)c_s(\lambda)$$

This equation gives rise to a planar distribution. Each pixel may be viewed as the sum of some amount of body reflection plus some amount of surface reflection. This idea is illustrated in figure 2.3.

For the purposes of this thesis, the surface color term $c_s(\lambda)$ and body color term $c_b(\lambda)$ will only capture

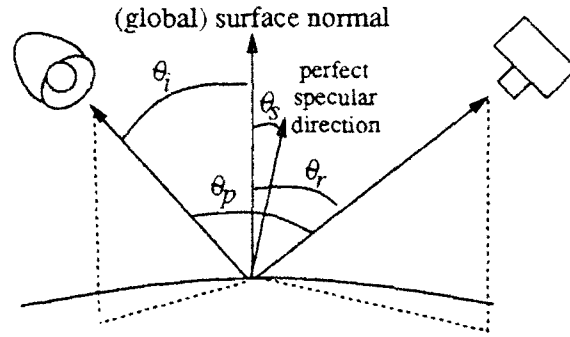


Figure 2.2: Angles that define imaging geometry

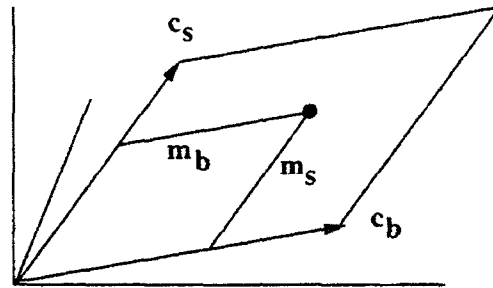


Figure 2.3: Dichromatic reflection components in a single pixel

relative reflectance, which includes hue and saturation but not brightness. The brightness will be captured in the geometry terms m_b and m_s . These functions of wavelength can be converted into RGB triples by multiplying the sensor responsivities for red, green, and blue and integrating over the visible spectrum.

$$\vec{c}_b = \begin{bmatrix} m_b \int c_b(\lambda) S_R(\lambda) d\lambda \\ m_b \int c_b(\lambda) S_G(\lambda) d\lambda \\ m_b \int c_b(\lambda) S_B(\lambda) d\lambda \end{bmatrix}$$

$$\vec{c}_s = \begin{bmatrix} m_s \int c_s(\lambda) S_R(\lambda) d\lambda \\ m_s \int c_s(\lambda) S_G(\lambda) d\lambda \\ m_s \int c_s(\lambda) S_B(\lambda) d\lambda \end{bmatrix}$$

The vector quantities $\vec{c}_b$ and $\vec{c}_s$ are then normalized to have length one.

The dichromatic reflection model shows how certain types of objects exhibit two types of reflection, and that these two types of reflection will typically have different hues. This hue difference means the histogram will fail in a plane in the color space. Furthermore, the two reflection processes also have different geometrical distributions associated with the physical reflection models.

2.2 Body Reflection Model

The body reflection is usually assumed to be Lambertian. Under this assumption, a point on an object appears equally bright from any reflectance angle. The only dependence upon imaging geometry is that the reflectance is proportional to the cosine of the incidence angle. The Lambertian model gives the body reflection m_b component as

$$m_b(\theta_i, \theta_r, \theta_p) = \gamma B_b \cos(\theta_i) \quad (2.1)$$

where θ_i is the angle of illumination incidence. The gain of the camera in converting photons measured by the CCD array into pixel values is represented by γ . The brightness of the body reflection is represented by the term B_b . This factors in both the reflectance of the object (albedo) and the intensity of the light. If the intensity of the illumination is known, then these factors may be separated and the albedo alone may be recovered.

Recent work by Wolff has shown that that the Lambertian assumption is very good for a large range of incidence and reflectance angles; however, there is a considerable error at extreme angles such as are seen at occluding edges [Wolff, 1992]. Points that are viewed from an oblique angle (θ_r close to 90°) are much less bright than would be predicted from the Lambertian model. Although the Wolff model has a slightly different dependence upon incidence angle than the Lambertian model, both models predict that the intensity falls off monotonically as the incidence angle is increased.

2.3 Surface Reflection Model

Unlike the body reflection, the surface reflection component is very dependent upon the reflectance angle. This is easily noticed by watching a highlight move as the view moves. The surface reflection also falls off much more quickly than body reflection. It has its maximum approximately at the perfect specular angle where the angle of incidence equals the angle of reflectance.

The spectral component of the surface reflection is modeled in this thesis as being the same as that of the incident light. This is an approximation, since the Fresnel equation that governs surface reflection shows that the reflection is dependent upon the wavelength of the incident light. However this dependence is weak across the visible spectrum. The assumption that surface reflection color is identical to illumination color is

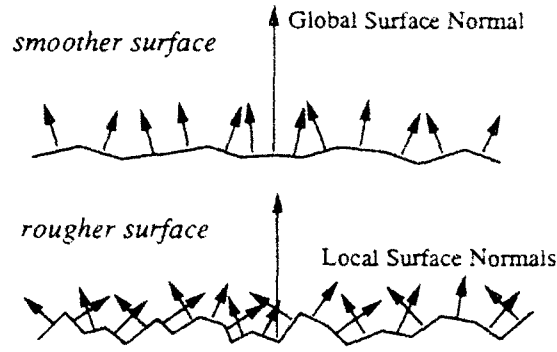


Figure 2.4: The facet model of surface reflection

a common one. It follows from this assumption that once the illumination vector is identified, its influence may be divided out of the body color to give the object's reflectance. This provides a method for color constancy under changing illumination conditions [Klinker, 1988], [D'Zmura and Lennie, 1986], [Healey and Binford, 1987a], [Tominaga and Wandell, 1989].

This thesis uses the Torrance-Sparrow model of scattering [Torrance and Sparrow, 1967]. This model views surfaces as a collection of tiny facets, each of which may have a local surface normal that is different from the global surface normal. For smoother surfaces, most of the facets have an orientation very close to the global surface normal. As a surface becomes rougher, it has a greater number of facets that are tilted with respect to the global normal. This is illustrated in figure 2.4. The distribution of facet normals is typically modeled as Gaussian, with σ describing the standard deviation. The distribution is also assumed to be isotropic, with rotational symmetry about the surface normal, giving a probability distribution p of

$$p(\theta) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{\theta^2}{2\sigma^2}\right) \quad (2.2)$$

where θ is the tilt of the facet with respect to the global surface normal.

A small value of σ , such as 1° , means that the standard deviation of facet angles is small, and that the surface is very smooth. A large value of σ , such as 10° , means a rough surface. The facets are assumed to be an order of magnitude larger than the wavelength of visible light, but this optical roughness is still much finer than the roughness that could be felt with fingertips or that would be used for texture analysis. For the remainder of this thesis, it will be assumed that the facet size is a constant for the surfaces to be analyzed.

The Torrance-Sparrow used for scattering gives the amount of surface reflection m_s as

$$m_s(\theta_i, \theta_r, \theta_p) = \gamma B_s \frac{F G(\theta_i, \theta_r, \theta_p) \alpha}{\sigma \cos(\theta_r)} \exp\left(-\frac{\theta_s^2}{2\sigma^2}\right) \quad (2.3)$$

where B_s is the intensity of the illumination; γ is the camera gain; α is a constant that includes $1/\sqrt{2\pi}$ from equation 2.2 and the facet size (a variable in the original Torrance-Sparrow model); and F is the Fresnel coefficient that describes what percentage of the light is reflected at the interface. F is a function of geometry, wavelength, polarization state, and index of refraction of the material in question; however, this thesis will follow the "Neutral Interface Reflection" model and assume that F is a constant for a given type of material.

G is an attenuation factor that also depends upon geometry and comes into play at grazing angles. G is a function of incidence angle, reflectance angle, and phase angle and is derived in [Torrance and Sparrow, 1967].

$$G = \text{MIN} \left\{ 1, \frac{2 \cos(\theta_s) \cos(\theta_r)}{\cos(\theta_p)}, \frac{2 \cos(\theta_s) \cos(\theta_t)}{\cos(\theta_p)} \right\} \quad (2.4)$$

The phase angle, θ_p , is the angle between the light source and the camera with respect to the object. Since there is a $1/\cos(\theta_r)$ term in the denominator of equation 2.3, that would imply that the surface reflectance would go to infinity as the reflectance angle approaches 90 degrees. However, as θ_r increases to 90 degrees, the attenuation term G goes to zero faster than the $1/\cos$ term goes to infinity, so that equation 2.3 does not make the improbable prediction of surface reflection becoming infinite. The G term corrects for the fact that facets will tend to mask or shadow other facets for grazing incidence and reflectance angles. The $1/\cos(\theta_r)$ term predicts that the maximum of the surface reflection will not always be exactly at the perfect specular angle. This models the "off-specular peaks" observed in real images [Torrance and Sparrow, 1967].

A different model for surface roughness is given in [Beckmann and Spizzichino, 1963]. In [Nayar *et al.*, 1991] this model is compared with the Torrance-Sparrow model and it is demonstrated that the two models make very similar predictions when the surface are not extremely smooth. A recent computer graphics model also makes similar predictions about the dependence of surface reflectance upon surface roughness [He *et al.*, 1991].

The Dichromatic Reflection Model shows how the different colors of body reflection and surface reflection imply a planar distribution in color space. The next chapter will show how the different geometric reflection processes of the two components give rise to a particular distribution of pixels within that plane.

Chapter 3

Structure of the Color Histogram

When talking about the color histogram, this thesis will mean a distribution of colors in the three-dimensional RGB space. Increasingly, vision algorithms are designed with color images in mind since color cameras are becoming cheaper. More importantly, color is a rich phenomenon that has the potential to provide much information about the scene. Traditionally, histograms are associated with a number of "bins" that cover the range of expected values and a frequency count to record the number of data points that fall within each bin. For a typical imaging system with 8 bits for each color band, there are 256^3 bins into which a pixel may fall. This work only considers whether a bin is full or empty in the three-dimensional color space. A fourth dimension could be used to record the number of pixels which have a particular RGB value. Of course, a fourth dimension would be difficult to visualize. More importantly, this fourth dimension would be dependent on such things as image magnification. If the camera were to zoom in on a particular object, the number of pixels it occupies in the image would grow, causing a change in the frequency count of the histogram.

Figure 3.1 contains a sketch of a typical color histogram for a dielectric surface without interreflection. This simple case will be studied first to show what can be learned from it before proceeding to the more complex analysis of interreflection. As labeled, the histogram has two linear clusters of pixels: the *body reflection cluster* and the *highlight cluster*. The first of those clusters extends from the black corner of the cube (at point a) to the point of maximum body reflection (point b). The other cluster starts somewhere along the body reflection cluster (point c) and extends to the highlight maximum (point d).

If the object has a completely black body color, there will be no body reflection component and hence

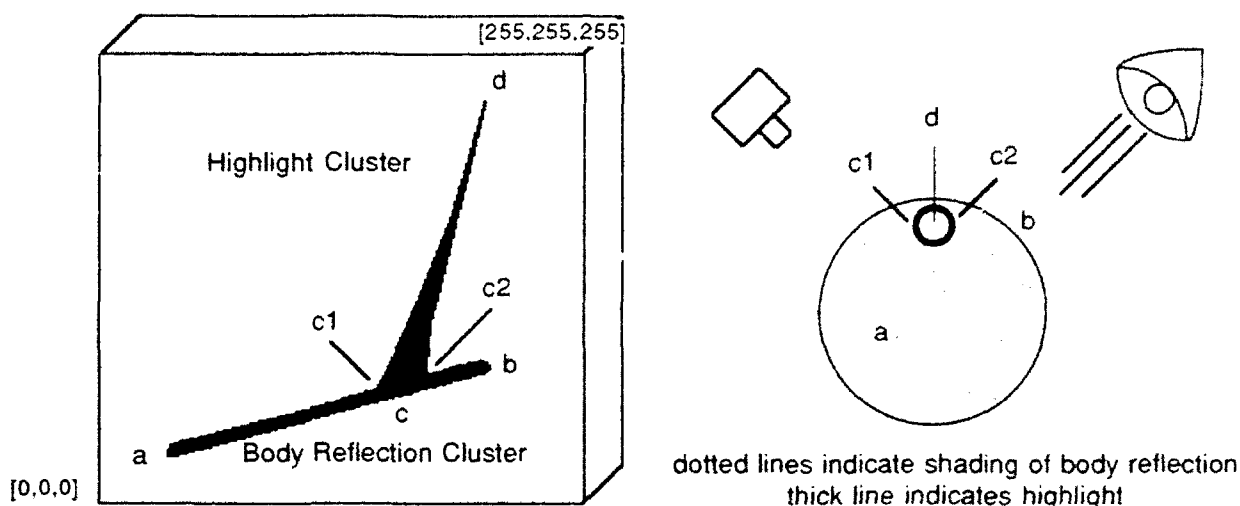


Figure 3.1: Histogram of a single object

no body reflection cluster. If the body reflection color is the same hue as the surface reflection color, the body reflection cluster and highlight cluster will be collinear. This is the case of white and gray objects since neither their body color nor surface color imparts any hue to reflected light. Therefore, objects of these colors cannot be analyzed by this type of method. Similarly, metals by their nature do not exhibit body reflection, and so will not have the color distribution shown in figure 3.1.

This chapter examines the structure of the color histogram, describing in detail the shape, orientation, and location of both histogram clusters. The highlighted points in the diagram shown in figure 3.1 (a, b, etc.) correspond to significant locations in both the histogram and the image and may be used to measure the dimensions of the histogram. The dimensions of the clusters relate in a specific, quantitative way to key scene parameters including object color, illumination color, surface roughness, and phase angle.

3.1 The Body Reflection Cluster

The linear cluster referred to as the body reflection cluster corresponds to pixels that exhibit mostly body reflection with very little surface reflection. If there is no ambient illumination in the scene, this cluster begins at the black point in the cube (point a), corresponding to points on the surface whose normal is 90 degrees or more away from the direction of the illumination. Such points fall beyond the self-shadow line and so are completely dark. The point at the other extreme of the body reflection cluster (point b), corresponds to the largest amount of body reflection seen anywhere on the object. For Lambertian objects,

pixels at this point in the histogram correspond to points on the surface with normals pointing directly at the light source. Assuming that the body reflection component is Lambertian, and recalling equation 2.1, the body reflection m_b obeys the relation

$$m_b(\theta_i, \theta_r, \theta_p) = \gamma B_b \cos(\theta_i)$$

Thus pixels located half-way along the body reflection cluster would correspond to surface points with normals $\cos^{-1}(1/2)$ or 60 degrees away from the illumination direction.

If the object exhibits all possible surface normals, the body reflection cluster will be full length and densely filled with pixels. If the object is composed of a small number of flat surfaces, there will be gaps in the body reflection cluster. It is possible that such an object will not show any highlight at all even though it is very shiny. The sketches of histograms in this thesis assume that the objects being observed exhibit a broad, continuous distribution of surface normals. However, small gaps in the body reflection cluster will not harm this analysis. The only absolute requirement is that surface normals corresponding to points **a** and **b** must be visible to the camera.

The sketch shows that the body reflection cluster is long and narrow. This agrees with the simulations used for this thesis and also with real, high-quality images of clean surfaces. Even for somewhat rough surfaces, the amount of surface reflection is very low over all but a very limited number of reflectance angles. Slightly thicker clusters have been observed for images with greater amounts of noise and for images of less uniform surfaces, but even so the cluster is highly linear.

A vector fitted to this cluster (from point **a** to **b**) will point in the direction of the body reflection color which is the product of the object color and the illumination color. Automatic fitting of this vector has been successfully demonstrated [Klinker, 1988]. Once the illumination color has been determined from analysis of the highlight, the object reflectance may be calculated by dividing out the influence of the illumination, as proposed in some color constancy methods [D'Zmura and Lennie, 1986], [Healey and Binford, 1987a], [Tominaga and Wandell, 1989].

Assuming that there is some point on the object which has a surface normal pointing directly at the light source, and which is visible to the camera, then at that point the cosine of the incidence angle will be one. This means that the length of the fitted vector (the magnitude $|\vec{ab}|$) corresponds to the gain times the object's apparent brightness B_b . If the intensity of the illumination could also be recovered from highlight analysis, then the albedo of the object could be separated out (assuming that the gain of the camera had been

calibrated). Otherwise it will be impossible to tell a bright light shining on a dark surface from a dim light shining on a light surface. Fortunately highlights provide an invaluable clue to distinguishing these cases.

To summarize this section:

- The vector fitted to the body reflection cluster (from point a to b) will point in the direction of the body reflection color, which is the product of the object color and the illumination color.
- The length of the fitted vector (the magnitude $|\vec{ab}|$) corresponds to the gain times the object's apparent brightness B_b . If the intensity of the illumination is recovered from highlight analysis, then the albedo of the object B_o can be separated out (assuming that the gain of the camera had been calibrated).

3.2 The Highlight Cluster

The cluster of pixels called the highlight cluster corresponds to those pixels that show a non-negligible amount of surface reflection. This corresponds exactly to the area of the image that we would call the highlight area. In the histogram, the highlight cluster starts where it intersects with the body reflection cluster (point c), usually in the upper half of the body reflection cluster. The highlight cluster extends upwards from there to the brightest point of the highlight (point d). For many shiny objects, the highlight is so bright that the highlight cluster suffers color clipping and touches the white point of the color cube (an RGB triple of [255, 255, 255] for an 8 bit camera) where the highlight has saturated the camera [Klinker, 1988].

Recall from equation 2.3 that the amount of surface reflection m_s is given as

$$m_s(\theta_i, \theta_r, \theta_p) = \gamma B_s \frac{F G(\theta_i, \theta_r, \theta_s) c}{\sigma \cos(\theta_r)} \exp\left(-\frac{\theta_s^2}{2\sigma^2}\right)$$

Obviously the magnitude of the highlight is dependent on a great many things, including illumination intensity, imaging geometry, and surface roughness. The next few sections will describe how the physical model expressed by this equation leads to a definite direction, shape and position for the highlight cluster.

3.2.1 Direction of Highlight Cluster

Surface reflection occurs at the interface, so it is not affected by the body color of inhomogeneous objects. The Fresnel coefficient F is very weakly dependent upon incidence angle and wavelength of illumination [Shafer, 1984]. Thus it is often assumed to be a constant for a given type of material, as in the Neutral Interface Reflectance (NIR) model. Following this assumption, all the terms in equation 2.3 relate to the magnitude of surface reflection. Thus the color of the surface reflection component will be equal to the illumination color.

It is natural to assume that the direction of the highlight cluster is related in some way to the surface reflection color. The highlight cluster is usually long and narrow in shape and a vector can be fitted to it (from point *c* to *d*). Klinker argued that the direction of this vector will usually correspond closely to the surface reflection color [Klinker, 1989]. This assumption is valid for fairly smooth objects where the highlight has a small area, and for imaging geometries where the body reflection changes slowly over that area. In these situations, the amount of body reflection at the base of the highlight cluster and the amount at the top of the cluster varies by only a small amount.

For example, if the camera and light source are separated by 0 degrees with respect to the object (so they are located right on top of each other in space), the perfect specular angle will be at $\theta_i = \theta_r = 0$. The highlight will be spread around this angle. Right around 0 degrees, the cosine of the incidence angle changes slowly. If the surface is optically smooth, the highlight will be confined to angles close to 0 degrees, and the amount of body reflection varies only a tiny amount over the entire area of the highlight. Since these highlight pixels will all contain roughly the same amount of body reflection, any color change within the highlight can be assumed to be due to surface reflection alone. Therefore, for scenes where the highlight occurs over a very small range of body reflection values, the assumption that the surface reflection color is equal to the highlight cluster direction is a good one.

On the other hand, if the object is optically rough and the highlight occurs on part of the object where the cosine of the incidence angle changes more rapidly, then the amount of body reflection can not be assumed to be constant over the area of the highlight. The body reflection at the base of the highlight cluster (point *c*) may vary significantly from the amount at the tip (point *d*). This means that a vector fit to the highlight cluster will represent some change in body reflection as well as increasing amounts of surface reflection, causing the vector to be skewed away from the direction of the illumination color. This skew means that the

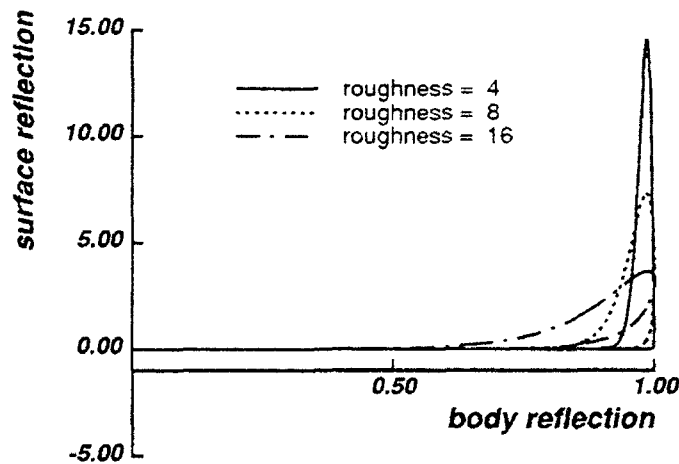


Figure 3.2: Dependence of highlight cluster direction on object roughness

estimate of the illumination color made from fitting a vector to the cluster will be somewhat inaccurate.

This phenomena can be visualized by projecting the histogram into the plane defined by the body reflection color and surface reflection color. Dichromatic reflection is simulated for objects with the same body color but with different roughness values. The light source and camera are 20 degrees apart with respect to the object. Figure 3.2 shows a cross section of the histograms that result. The horizontal direction is the direction of increasing amounts of body reflection; all three histograms fall right on this line for pixels where there is no surface reflection. The vertical direction is defined by increasing amounts of surface reflection. For a surface where the standard deviation of facet angles (roughness) is 4° , a vector fitted to the highlight cluster will point exactly in this direction. A vector fitted to the highlight cluster when σ equals 8° will deviate slightly from the vertical direction. In a more extreme case, where σ equals 16° , the vector will deviate dramatically.

Figure 3.3 shows a sketch of this histogram for the roughest surface, with key points labeled. The estimate of illumination color from this highlight cluster will give the direction of the light color as $\vec{cd}$, whereas the correct illumination color is defined by the the direction $\vec{c'd}$.

In the color histogram, the vectors describing the body reflection color and the illumination color are not generally perpendicular. The angle between them depends upon the colors of the two reflection components, which are not generally known in advance. If the vector fitted to the highlight cluster does not point exactly in the direction of increasing amounts of surface reflection, the estimate of illumination color will be off by some amount.

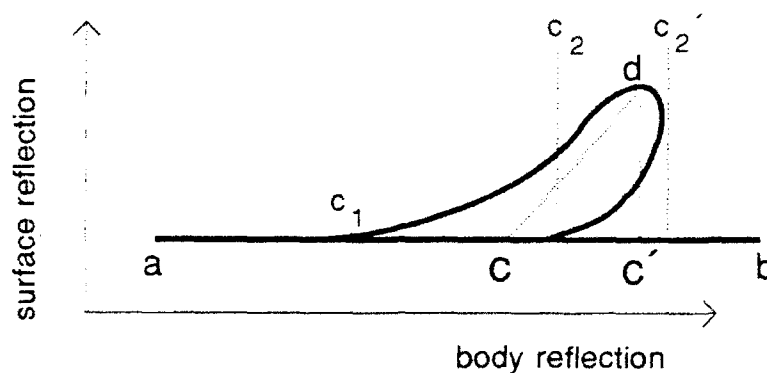


Figure 3.3: A skewed highlight

Furthermore, the direction of that vector is used to calculate the amount of body reflection at each point. In [Shafer, 1984] it was proposed that an "intrinsic image" of body reflection could be generated by determining the surface reflection vector and projecting all pixels along this vector onto the body reflection vector. If the fitted vector were skewed from the correct direction, pixels would be projected onto the wrong part of the body reflection vector. In figure 3.2, all highlight pixels should be projected straight down to calculate the amount of body reflection present. However, for the roughest surface ($\sigma = 16$), pixels would end up being projected at an angle, so that the amount of body reflection present would be underestimated.

[Klinker, 1988] demonstrated automatic fitting of a vector to highlight clusters to estimate illumination color. The vector fitted to the highlight cluster (from point *c* to point *d*) is a good first estimate of the illumination color, but this thesis has shown that it may be skewed. If the surface roughness and the imaging geometry are known, then the amount of skewing can be calculated and compensated for.

In summary:

- The vector fitted to the highlight cluster from point *c* to *d* is an estimate of the illumination color.
- That vector may be skewed away from the illumination color, depending upon the roughness and imaging geometry.

The next few sections will show how the roughness and imaging geometry may be estimated directly from the histogram itself.

3.2.2 Length of Highlight Cluster

When looking at highlights on a variety of surfaces, it is quickly observed that highlights are brighter and sharper on some surfaces, while they are dimmer and more diffused on other surfaces. (The fact that highlights may be diffused is a good reason to avoid the term "diffuse reflection" and use "body reflection" instead.) Very shiny surfaces exhibit only a tiny amount of scattering of the surface reflection, whereas very matte surfaces have a great deal of scattering. This scattering of surface reflection is a result of the optical roughness of the surface.

Color figure A.1 shows an image of two billiard balls and the associated histogram for the yellow ball on the left. (The yellow line shows where the ball was hand segmented to remove the number spot.) The billiard balls have a very smooth finish, giving a shiny appearance. Color figure A.2 shows the same balls with the same viewing geometry, but here the balls have been sprayed with an aerosol. This deposits tiny droplets on the surface, making it optically rough. Comparing the two histograms, it is immediately obvious that the highlight cluster for the shiny one is considerably longer than the cluster for the matte ball.

It is obvious from equation 2.3 that the sharpness of the peak is determined by the standard deviation σ , and that the height of the peak is inversely proportional to σ . Intuitively this makes sense, since surface reflection that is scattered over a very small area will be more "concentrated." A smooth object will have a small standard deviation of facet slopes, σ , resulting in a long, narrow highlight cluster. A rough object will have a large σ , and so will exhibit a shorter and wider cluster. Figure 3.4 shows a plot of the length of the highlight cluster vs. the object's roughness for simulated images where all other factors have been held constant.

Suppose that the task is to estimate roughness values for a variety of different color objects. The body reflection will be different for each of the surfaces, but the amounts of surface reflection are not affected by this. The length of the highlight cluster (the distance from c to d) indicates the degree of roughness, regardless of body color.

According to equation 2.3, it will be necessary to know the Fresnel coefficient for the type of material being observed in order to relate roughness to highlight cluster length. However, it turns out that for a wide range of plastics and paints, the indices of refraction are very nearly identical. For these types of materials the index of refraction is typically between 1.45 and 1.56 with the majority between 1.52 and 1.55. Fresnel's law for surface reflection at normal incidence is $F = ((n - 1)/(n + 1))^2$ where n is the index of refraction. The

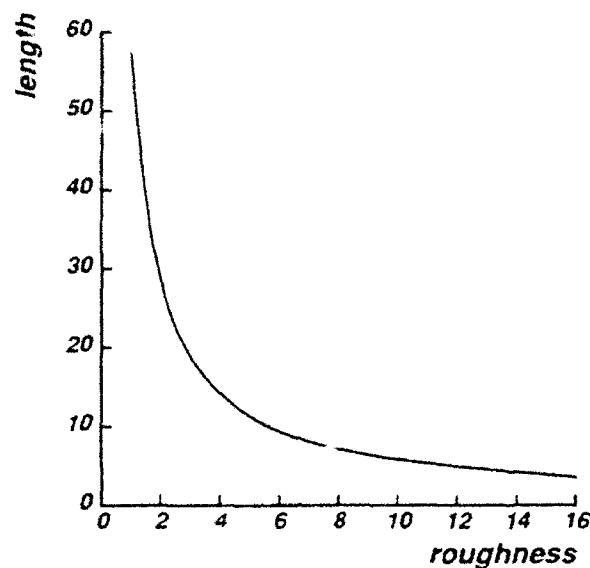


Figure 3.4: Dependence of highlight cluster length on object roughness

index of refraction is a function of wavelength, but for these types of materials, it is nearly constant across the visible range. Figure 3.5 shows the dependence of the length of the highlight cluster upon roughness for different indices of refraction. The variation is small enough to allow an estimate without knowing the material, although for an accurate measure the exact index of refraction associated with the material would be needed. Henceforth this thesis will assume that materials have an index of refraction of 1.5, corresponding to 4.0% Fresnel reflectance.

Figure 3.2 showed that if the surface is very rough ($\sigma = 16^\circ$), the highlight cluster will be skewed from the direction of the illumination color. This makes the cluster longer than it would otherwise be. The amount of surface reflection at the brightest point is the vertical displacement in the graph (the distance from point c' to point d in graph 3.5) rather than the distance along the highlight cluster (distance from point c to point d). For the smoother surfaces shown in the graph, the length of the highlight cluster is virtually identical to the amount of surface reflection. For very rough surfaces, it will be necessary to account for the highlight cluster's skew when calculating the length of the highlight cluster.

Equation 2.3 also indicates that the intensity of the light source affects the magnitude of the surface reflection, and thus the length of the highlight cluster. This is captured by the term B_s . If all objects are imaged under the same light source, a *relative* measure of roughness (indicating how much rougher one surface is than another) could be obtained without knowing the strength of the light. However, in order to calculate an absolute roughness value, the light source brightness must also be analyzed. It is obvious from

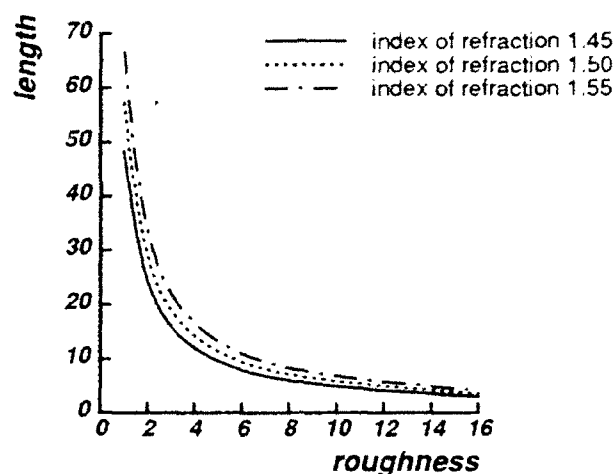


Figure 3.5: Dependence of length on roughness for different indices of refraction n

equation 2.3 that the length is directly proportional to this brightness. It will be assumed that the gain of the camera γ has been calibrated and that the facet size α is known. If not, then only the overall gain $\gamma\alpha B_b$ could be recovered from the histogram.

The graphs in figures 3.4 and 3.5 were calculated for an imaging geometry in which the light source and camera are separated by zero degrees. However, equation 2.3 predicts that the imaging geometry will have an effect upon highlight magnitude, as indicated by the $\cos(\theta_r)$ term in the denominator and the attenuation term G in the numerator. Figure 3.6 shows how the length of the highlight cluster changes as the camera and light source are separated by different phase angles with respect to the object. It demonstrates that the length changes slowly as the imaging geometry changes. Figure 3.7 shows how the length varies with roughness for a variety of phase angles. Each curve represents a different phase angle. If the imaging geometry is not known, it would be impossible to make fine distinctions of roughness. It would still be possible to tell a very smooth object ($\sigma = 1^\circ$) from a fairly rough one ($\sigma = 10^\circ$).

The result of this section is that the roughness of surfaces made from similar materials can be directly compared by examining their color histograms. This feature of the histogram does not depend upon the shape of the object, and so does not require the fitting of surface orientations to image data. Once the length of the highlight cluster has been measured, it can be used to make an estimate of surface roughness. Unfortunately, the true length cannot always be obtained. For example, the highlight in figure A.1 is so shiny that it saturates the camera. This means that the highlight cluster is clipped, so that its true length is unknown. As can be seen from the images in figures A.1 and A.2, rough objects can be distinguished from smooth ones, but it would not be possible to distinguish *fairly* smooth surfaces from *very* smooth ones by

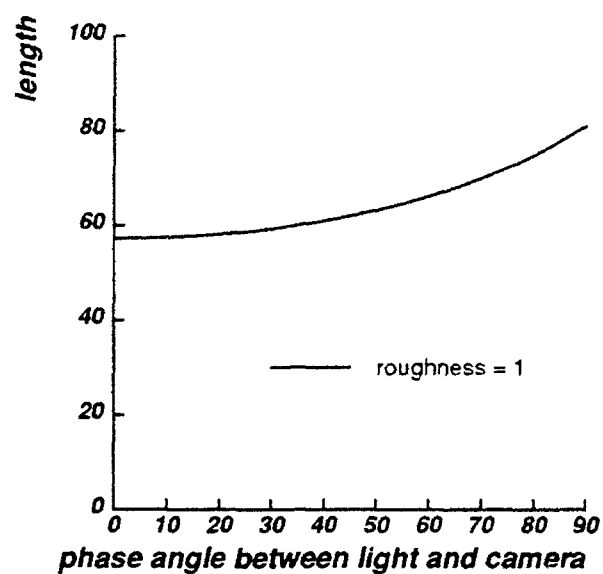


Figure 3.6: Dependence of length on imaging geometry

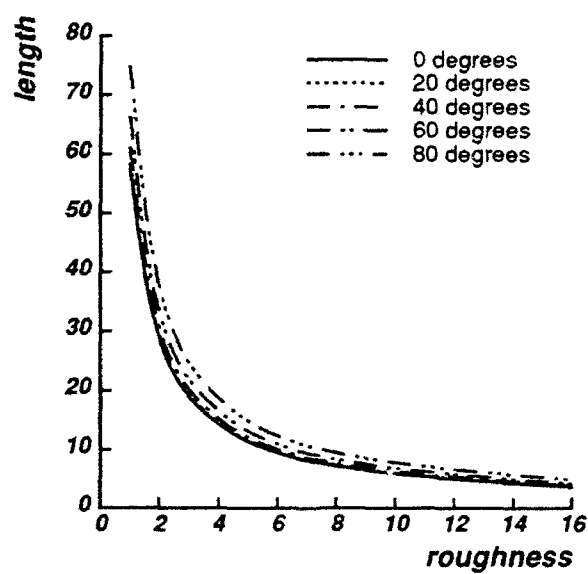


Figure 3.7: Dependence of length on roughness for different imaging geometries

this feature of the histogram.

To summarize this section:

- The length of the highlight cluster is inversely proportional to the roughness of the surface.
- The length of the highlight cluster is directly proportional to the intensity of the illumination.
- The length of the highlight cluster is weakly dependent upon the imaging geometry.
- The measurement of the highlight cluster length will be affected if the estimate of illumination color is skewed.

3.2.3 Width of Highlight Cluster

Another difference between the histograms for the shiny and matte balls in figures A.1 and A.2 is the width of the histogram at the meeting point between the highlight cluster and the body reflection cluster (the distance from point c1 to point c2 in the diagram of figure 3.1). The histogram of the matte ball is wider where the two clusters meet than the histogram for the shiny ball. Equation 2.3 shows that for rougher objects, with a larger standard deviation of facet angles σ , the surface reflection is scattered more widely.

The degree to which the surface reflection is scattered may be recovered from the width of the highlight cluster. Regardless of object shape, rougher surfaces show a non-negligible amount of surface reflection over a larger number of reflectance angles. In the color histogram, a noticeable amount of surface reflection results in pixels that are displaced from the body cluster in the direction of the illumination color. By taking any highlight pixel and projecting along the surface color vector onto the body reflection vector, the amount of body reflection present in that pixel can be calculated. By considering all the pixels in the highlight area of the image and looking at how much body reflection is in each of them, some range of body reflection magnitudes is obtained. If the surface is very smooth with a sharp, bright highlight, that range will be small. However for a rougher object with a more diffused highlight, the range of body reflection magnitudes will be larger since the highlight is spread over a larger number of surface normals.

This property is independent of object size or shape. It simply shows the variation in surface normals over the area of the highlight. It is not necessary to fit a surface shape to the image to know how much scattering the object exhibits. It is encoded right there in the histogram. The only requirement of the surface

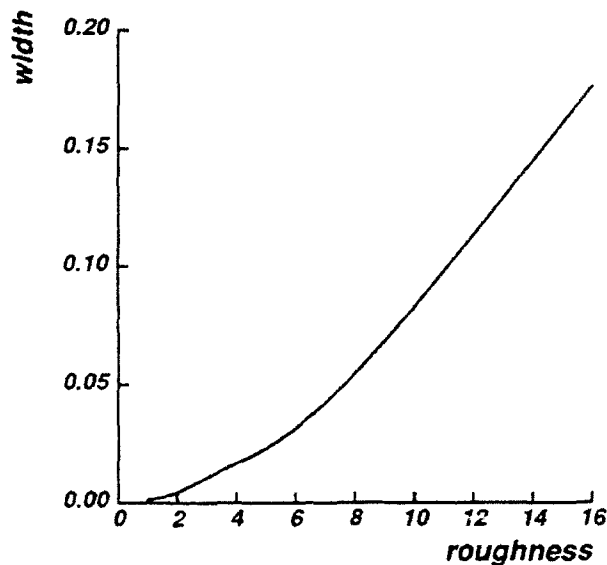


Figure 3.8: Dependence of highlight cluster width on object roughness

shape is that the object has surface normals corresponding to points c_1 and c_2 (see figure 3.1) visible to the camera.

Objects with different roughness values were simulated; then those pixels that showed a noticeable amount of surface reflection were identified. ("Noticeable" means that the pixel was further from the body reflection cluster than the noise threshold.) The body reflection for each of these points was calculated and the variation in value recorded. This variation was divided by the overall length of the body reflection vector to yield a fraction (the length of $\overline{c_1c_2}$ divided by the length of $\overline{ab}$ in figure 3.1). A fraction of 0.5 would mean that the highlight cluster's base extended across half the length of the body reflection cluster. Figure 3.8 shows how the highlight cluster width varies with the surface roughness if a point light source is used.

Moreover this measurement is not hampered by camera saturation since body reflection generally does not saturate the camera. It does not depend on how *much* surface reflection is at any given point, only whether there is *any* at all. It is, however, disturbed by camera blooming. Camera blooming means that when one pixel becomes saturated, excess charge spills over into neighboring pixels, making their values unreliable. In cameras that are prone to blooming, a highlight bright enough to saturate the camera will also spread out beyond its true width. If at all possible, a camera that is resistant to blooming should be used to take pictures of objects with highlights.

In the case of very rough surfaces, the highlight cluster may be skewed from the direction of the illumination color, as shown in figure 3.3. The width of this cluster where it meets the body reflection

cluster *underestimates* the variation of body reflectance values. The amount of body reflectance is correctly determined by projecting all pixels along the illumination color vector (straight down for this graph). The correct measure of highlight width is the variation in body reflectance (the distance from c_1 to c'_2 rather than the distance from c_1 to c_2). However, the highlight and body reflection colors are not generally perpendicular in most histograms, so just measuring the horizontal extent of the highlight cluster will not be sufficient. For the smoother surfaces shown in figure 3.2 (when $\sigma \leq 8^\circ$), the highlight cluster is centered above its base, so the width of the base and the horizontal extent of the cluster are virtually identical. For these types of surfaces, the width at the base is a very good approximation of the overall width. However, for very rough surfaces, it will be necessary to account for the skewing when determining the width of the cluster.

The brightness of the illumination B_s will also have some effect on the the width of the highlight cluster. As the light intensity is increased, points on the surface that had amounts of surface reflection just under the threshold may now have amounts that exceed the threshold. Clearly the width will grow as the light intensity grows. However the growth is not linear, since width measures the variation of body reflection values over the area of the highlight, not the intensity of the highlight directly. Figure 3.9 shows the relationship between incident light intensity and highlight cluster width for a fixed value of roughness and phase angle. The horizontal axis shows increasing light intensity, from 50% to 100% of a theoretical light's maximum output. The width grows very slowly over this entire range.

The conclusion is that the width is also affected by increasing illumination intensity, but in a different way from the effect of increasing intensity upon highlight length (where the growth is linear). This indicates that there is a difference between the histogram that results from a bright source illuminating a rough object, and the histogram from a dim source illuminating a smooth object. Since the length grows faster than the width as illumination intensity is increased, this means that the ratio between the length and the width width will increase as the illumination intensity is increased.

Although the width of the highlight cluster at its base does not depend upon the object's size and shape, it does depend upon the imaging geometry. This is due to the fact that the width measurement indicates the *projection* of highlight angles onto the body reflection vector. This effect is illustrated in figure 3.10.

The diagram on the left of figure 3.10 shows a surface where the highlight spreads 15 degrees in every direction from its maximum. When the camera and light source are separated by a phase angle of 30 degrees with respect to the object, the perfect specular angle will be at 15 degrees with respect to the illumination vector. The highlight will spread over points with surface normals ranging from 0 degrees to 30 degrees.

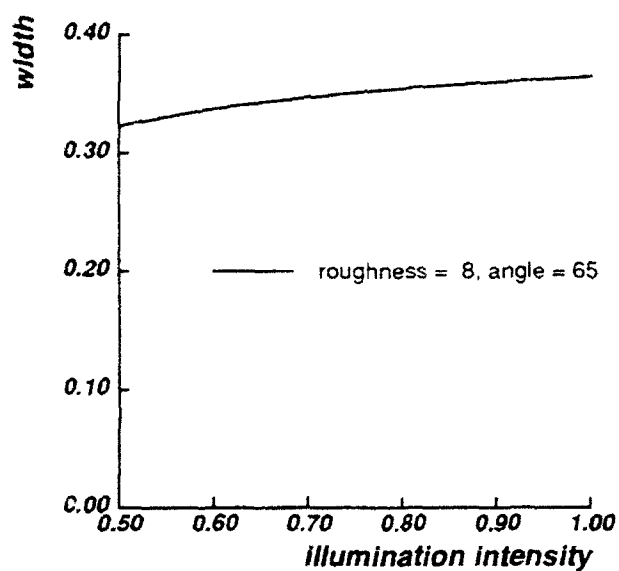


Figure 3.9: Dependence of width on incident light intensity

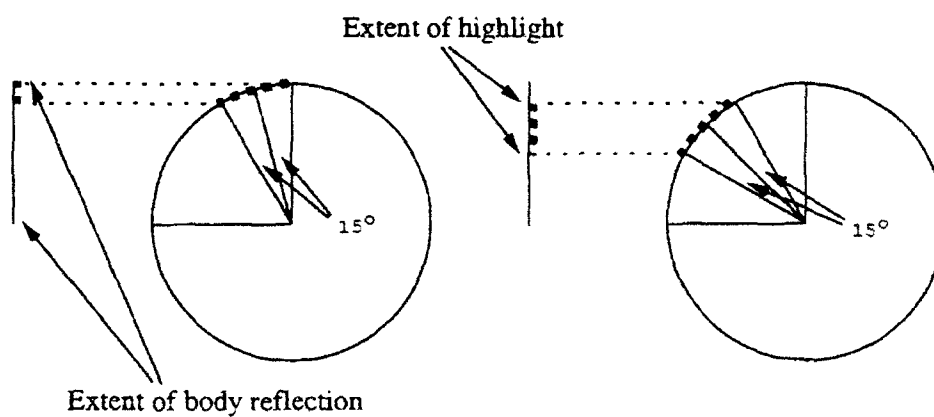


Figure 3.10: Illustration of why width varies with imaging geometry

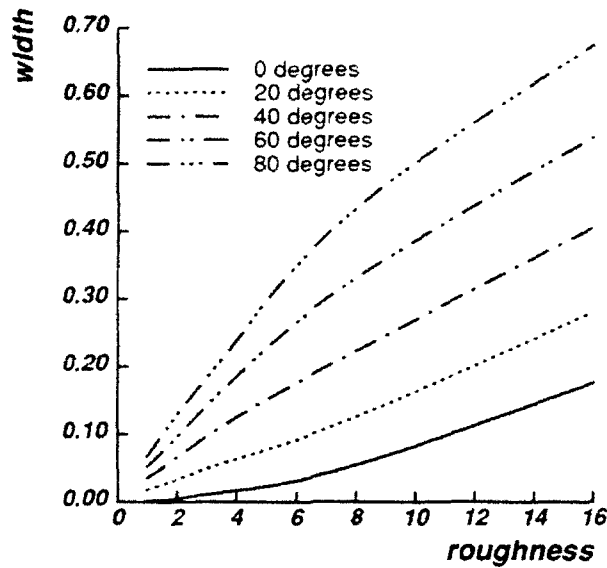


Figure 3.11: Dependence of width on roughness for different imaging geometries

(For the purpose of this explanation, ignore the influence of the $1/\cos(\theta_r)$ term.) The amount of body reflection at these points will vary from $\cos(0) = 1.0$ to $\cos(30) = .87$, a width of 0.13.

If the camera and light source are separated by 90 degrees, the perfect specular angle will be at 45 degrees with respect to the light source. This is shown in the diagram on the right of figure 3.10. One side of the highlight will spread 15 degrees closer to the direction of the illumination, so that it reaches points with surface normals of 15 degrees with respect to the light. The opposite edge of the highlight will spread 15 degrees in the other direction, to those points with surface normals of 45 degrees. Thus the amount of body reflection will vary from $\cos(30) = .87$ to $\cos(60) = .50$, a width of 0.37.

Figure 3.11 shows how the width of the highlight cluster varies with roughness for a variety of imaging geometries. The angle label is a measure of the phase angle (in degrees) that separates the light source and camera (with respect to the object). It is assumed that the object is small enough and far away enough from the camera that this angle is the same for all points on the surface.

For the case of highlight cluster width, the measurement is very sensitive to different viewing geometries. At the same time, the width measurement is very *insensitive* to changes in illumination intensity. The width of the highlight cluster also depends upon the degree of scattering of surface reflection, which is an indication of surface roughness.

In summary:

- The width is strongly dependent upon the roughness of the object.
- The width is also strongly dependent upon the imaging geometry.
- The width grows very slowly as illumination intensity is increased.
- The measurement of the highlight cluster width will depend upon the estimate of illumination color.

3.2.4 Intersection Point of Highlight and Body Reflection Clusters

When the diagram in figure 3.1 was first introduced, the highlight cluster was described as beginning "somewhere" along the body reflection cluster. This is rather vague and unsatisfying so a natural question to ask would be, "Just where along the body reflection cluster does it begin?"

In attempting to automatically extract the body reflection and highlight clusters, Klinker sought to set a limit upon where the program should search for the start of the highlight cluster. She derived the "50% heuristic" which stated that for a large range of viewing geometries, the highlight cluster would start somewhere in the upper 50% of the body reflection cluster [Klinker, 1988].

The distance along the body reflection cluster where the two clusters meet (the length of $\overline{ac}$ divided by the length of $\overline{ab}$ in figure 3.1) shows the amount of body reflectance at those points on the surface that are highlighted. Assuming that the body reflection is Lambertian, then equation 2.1 says the amount of body reflection is proportional to the cosine of the incidence angle θ_i . If the two clusters meet at the maximum point on the body reflection cluster, it means the highlight occurs at those points that have the maximum amount of body reflection, which is precisely those points with surface normals pointing directly at the light source. If the two clusters meet halfway along the body reflection cluster, the highlight must occur at points with surface normal pointing $\cos^{-1}(1/2)$ or 60 degrees away from the illumination direction.

Since it is assumed that the body reflection is Lambertian, it does not depend in any way upon the angle from which it is viewed. Thus the body reflection does not provide any information about the camera direction. However, the surface reflection is dependent upon both the illumination and camera directions. Neglecting for the moment the $1/\cos(\theta_r)$ term in equation 2.3, it is easy to see that the maximum amount of surface reflection will occur approximately at those points on the surface where the angle of incidence equals the angle of reflection. Thus if the highlight occurs at a point where the surface normal faces 10 degrees

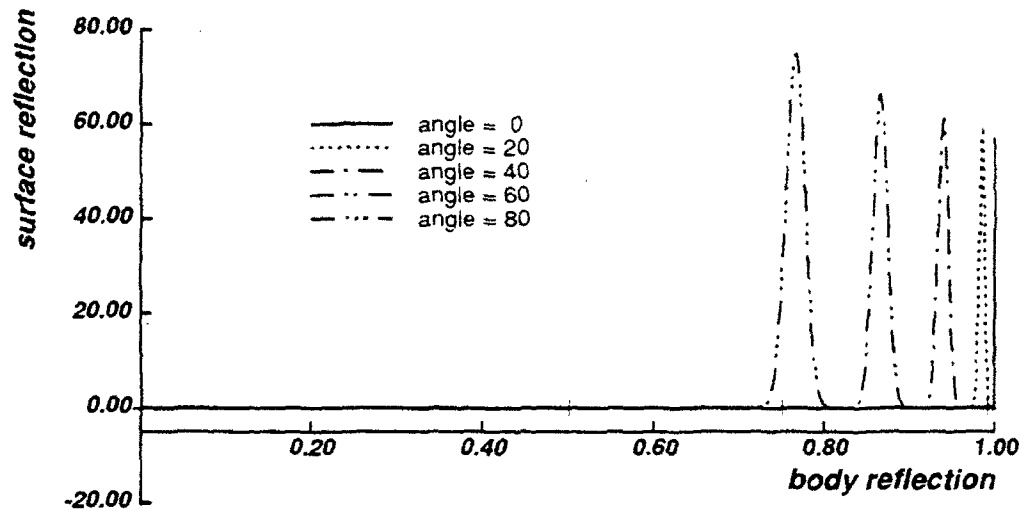


Figure 3.12: Illustration of how highlight cluster varies with changing imaging geometry

away from the light source direction, the light source and camera must be 20 degrees apart with respect to that point on the surface.

Figure 3.12 graphically illustrates this phenomenon. Once again the histograms have been projected into the plane defined by the body reflection and surface reflection colors, so that the horizontal direction indicates increasing amounts of body reflection and the vertical direction indicates increasing amounts of surface reflection. This time, the amount of roughness has been held constant at $\sigma \approx 1^\circ$ while the phase angle between the light source and camera has been varied from 0 to 80 degrees. This graph shows how the intersection point ($|\overline{ac}|/|\overline{ab}|$) decreases as the phase angle between the camera and light source increases. Incidentally it also shows how the length and width of the cluster are affected by imaging geometry as described in sections 3.2.2 and 3.2.3.

It does not matter whether the object has one highlight or many. The highlight or highlights will always occur at points with the same surface normal for a given imaging geometry, and that is what determines the intersection point of the clusters. Figure 3.13 graphs intersection point vs. imaging geometry. The horizontal axis shows the phase angle between the light source and the camera with respect to the object. A negative number means that the light source is to the left of the camera, whereas a positive number means the light source is to the right of the camera. The vertical axis is the ratio $|\overline{ac}|/|\overline{ab}|$. As can be seen from the graph, the meeting point is never less than 1/2 of the way along the body reflection vector, showing that the 50% heuristic is a sound one.

The previous few paragraphs assumed that the intersection point of the two clusters shows the amount of

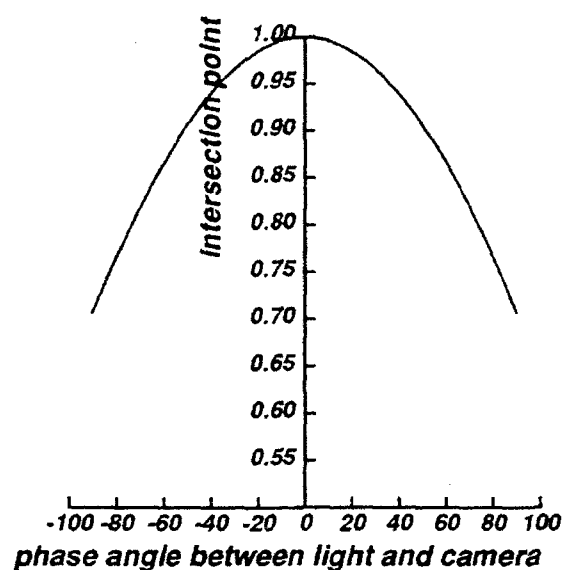


Figure 3.13: Dependence of intersection point on imaging geometry

body reflectance at the brightest point in the highlight. This assumption is valid for cases where the highlight cluster extends in approximately the same direction as the illumination color. However if the highlight cluster is significantly skewed away from the direction of the illumination color, the intersection point will give an incorrect estimate of the imaging geometry. This is the case for the roughest surface ($\sigma = 16^\circ$) shown in figure 3.2. All three histograms correspond to a phase angle of 20 degrees, and indeed the highest point of each highlight cluster occurs at the same horizontal coordinate (which indicates the amount of body reflection). However, for the rough surface shown in figure 3.3, the correct intersection ratio is given by $|\overline{ac'}|/|\overline{ab}|$ rather than $|\overline{ac}|/|\overline{ab}|$. This skewing of the highlight cluster means that the corrected estimate of illumination color is needed to correctly calculate imaging geometry for very rough surfaces.

The $1/\cos(\theta_r)$ term in the surface reflection equation means that the maximum amount of surface reflection will not always occur precisely at the perfect specular angle. This is particularly true of rougher surfaces where the highlight is spread over a wide range of reflectance angles so that $1/\cos(\theta_r)$ varies enough to make a difference. This gives rise to "off-specular peaks" described in [Torranee and Sparrow, 1967] and [Nayar *et al.*, 1991]. Figure 3.14 shows how the intersection graph changes for different roughness values. The graphs shows that the roughness does indeed affect the intersection point of the two clusters, but only by a small amount for the most common imaging geometries (light source and camera separated by less than 45 degrees).

The relationship described here between the intersection of the two clusters and the position of the light

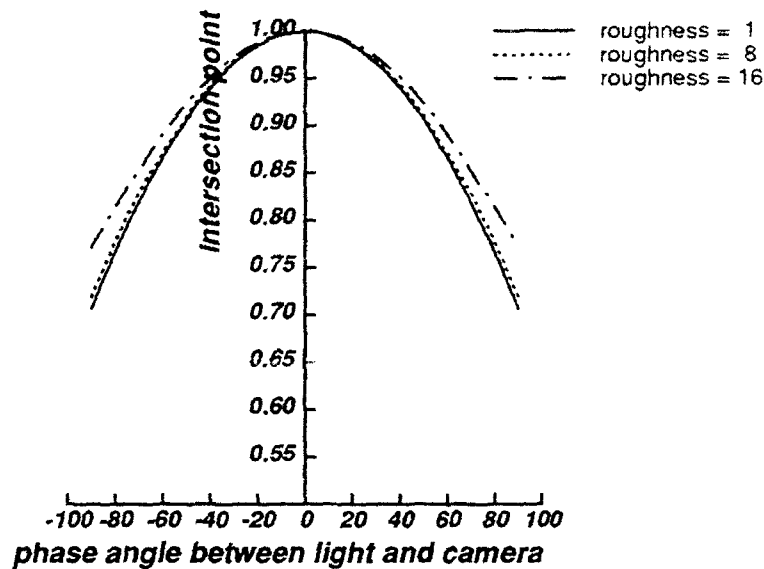


Figure 3.14: Dependence of intersection on geometry for different roughness values

source assumes that surface normals corresponding to key points in the histogram (a, b, and d in figure 3.1) are visible to the camera. In particular, the requirement that point b is visible means that the phase angle must be less than or equal to ninety degrees. Otherwise the point of maximum highlight would occur at surface normals out of view of the camera. While situations where the light is further than 90° from the camera can occur, a large number of the surface normals that would otherwise be visible to the camera will be dark, since they will be beyond the self shadow line. This imaging geometry is not common in scenes where the object is deliberately illuminated for the picture, since it is customary to illuminate scenes so a large number of surface normals will be visible.

This method for determining the phase angle between the light and camera by examining the highlight cluster's intersection with the body reflection cluster is implicitly photometric, since it relies on the amount of body reflection at certain points. However, it does not require the fitting of surface orientations to image data, nor does it require knowledge of the albedo of the surface. Moreover, it is done from a single image.

To summarize this section:

- The intersection of the two clusters is strongly dependent upon the imaging geometry.
- The intersection measurement is weakly dependent upon imaging geometry.
- The intersection measurement will be decreased if the estimate of illumination color is skewed.

	Roughness	Phase Angle	Illumination Intensity
Length	strong	weak	strong
Width	strong	strong	weak
Intersection	weak	strong	none

Figure 3.15: Dependence of histogram features upon scene parameters

Recently Wolff has shown that the Lambertian model is somewhat inaccurate for body reflection from many real surfaces, particularly at extreme reflectance angles; he has proposed a different model to replace it [Wolff, 1992]. Also, some researchers prefer to use the Beckmann-Spizzichino model [Beckmann and Spizzichino, 1963], or a combination of the Beckmann-Spizzichino and the Torrance-Sparrow models [Nayar *et al.*, 1991], to describe surface scattering. While a choice of different models would change the exact positions of the graphs in this section, the general relationships would remain the same. These dependencies of the color histogram upon scene parameters are a consequence of the different geometric distributions as well as the different hues of the two components of reflection.

Sections 3.2.2 and 3.2.3 showed how the roughness of the object affects the from the length and width of the highlight cluster; however they also showed that there is some dependence on imaging geometry. Now this section shows how the imaging geometry affects from the intersection point of the two clusters, but that there is some dependence upon roughness. The degree of dependence of each of these histogram measurements upon the scene parameters is characterized in figure 3.15.

Furthermore, in section 3.2.1 it was noted that the direction of the highlight cluster could be skewed away from the direction of the illumination color, depending upon the amount of roughness and the imaging geometry. Yet the estimates of cluster length, width, and intersection point depend to a certain extent upon knowing the true direction of the illumination color. Obviously these factors are all interdependent. One solution might be to first estimate those things that are weakly dependent on unknowns and then iterate. However this thesis has chosen to solve for all values simultaneously. That solution will be developed in chapter 5.

Chapter 4

Color Histogram Structure when Interreflection is Present

In many models of image formation, the illumination is considered to have a very simple form: either the illumination is invariant with geometry (modeling "ambient" lighting) or is non-zero in only one direction (modeling a "point source"). In both these cases the illumination is considered to have only one color so that algorithms like color constancy may be applied. The usual model allows only "primary reflection," in that the light is assumed to hit objects in the scene once before entering the camera.

Unfortunately this type of model is not sufficient to explain interreflection, which can occur even in simple images. Therefore, it would be advantageous to use a more general model which allows for reflection of light between objects. In order to do this, this thesis will use an expanded version of the dichromatic reflection model.

4.1 Modeling Interreflection

As is done with the traditional dichromatic reflection model, the reflection from an object is separated into body reflection and surface reflection. Consider what happens when each of the two components from one surface (A) is reflected by another surface (B) in two different ways. This gives rise to "secondary reflection" with four components whose magnitudes are:

$m_b^A m_b^B$ body reflection of A times the body reflection of B

$m_b^A m_s^B$ body of A times surface of B

$m_s^A m_b^B$ surface of A times body of B

$m_s^A m_s^B$ surface of A times surface of B

If these are taken together with the primary reflection components, they make a total of six reflection components [Shafer *et al.*, 1990], [Bajcsy *et al.*, 1990].

In the “tertiary reflection” case, the light makes three bounces before entering the camera, giving 2^3 or 8 possible reflection components. Taken together with the primary and secondary reflection components, there are a total of 14 reflection components. In general, for n -ary reflection, there are 2^n reflection components with **exactly** n bounces, for a total of $2^{n+1} - 2$ reflection components with **up to** n bounces. Fortunately, objects usually reflect a fraction of the light impinging upon them, so that in most cases the high-order reflection components will approach zero for large n .

Color figure A.3 shows the results of simulating interreflection for two cylinders that are identical except for hue. The Torrance-Sparrow model was used to simulate the surface reflectance, with a roughness of $\sigma = 1^\circ$. In this simulation the light source and camera are at the same position in space and are infinitely far from the objects. The top portion of the picture shows the results when only primary reflection is allowed, so there is no interreflection. The middle section of the picture shows simulated primary and secondary reflection only. The bottom section shows a simulation of primary, secondary, and tertiary reflections. The difference caused by the addition of tertiary reflection is small and not noticeable to the human eye. Henceforth this thesis will only consider primary and secondary reflection for this type of surface.

Figure 4.1 shows a graph of the magnitudes of the 6 reflection components for the simulated cylinder A shown in figure A.3. Each of the components represents a distinct path from the light source to the camera. The horizontal axis is the angular position of points on the cylinder, with zero corresponding to the points with surface normals pointing at the camera, and negative angles corresponding to the side facing the other cylinder. The vertical axis indicates the relative strength of the component at that position on the cylinder. Notice that although there is no body reflection from cylinder A past -90 degrees (this is beyond the self-shadow line), some of the interreflection components are positive at -100 degrees since these points can “see” light reflected from cylinder B.

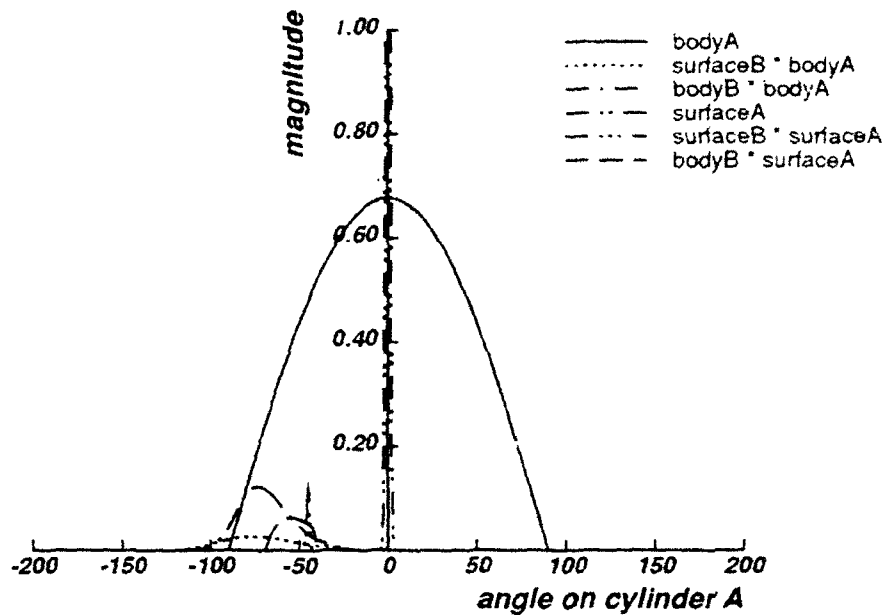


Figure 4.1: Interreflection components, by path

Although this model considers 6 different paths by which light may reach the camera, the components will show at most four different hues. This is because the Neutral Interface Reflection model predicts that light that bounces from the surface of A will have the same hue as the illumination color; thus $m_s^A m_b^B$ will have the same hue as m_b^A , while $m_s^A m_s^B$ will have the same hue as m_s^A . Figure 4.2 shows a graph of the magnitudes of the 4 distinct hues obtained when components with the same hue are combined.

Figure 4.3 shows a sketch of the color histogram that arises from a single object with interreflection. The relevant parts are labeled. The body reflection cluster and highlight cluster are in the same positions as they were for the simpler histogram. However the highlight cluster has been renamed the *primary highlight cluster*. This is because there is now a distinct highlight cluster called the *secondary highlight cluster*. The *interreflection cluster* refers to the remaining clump of pixels that deviate from the body reflection near its base. The next few sections will examine in detail these new histogram features.

4.2 The Secondary Highlight Cluster

The secondary highlight cluster results from interreflection, specifically the surface-to-surface component. It corresponds to those pixels that form a secondary highlight, a mirror image of the highlight on the other object. Although the secondary highlight is quite small, it is visible in the photo of the shiny billiard balls

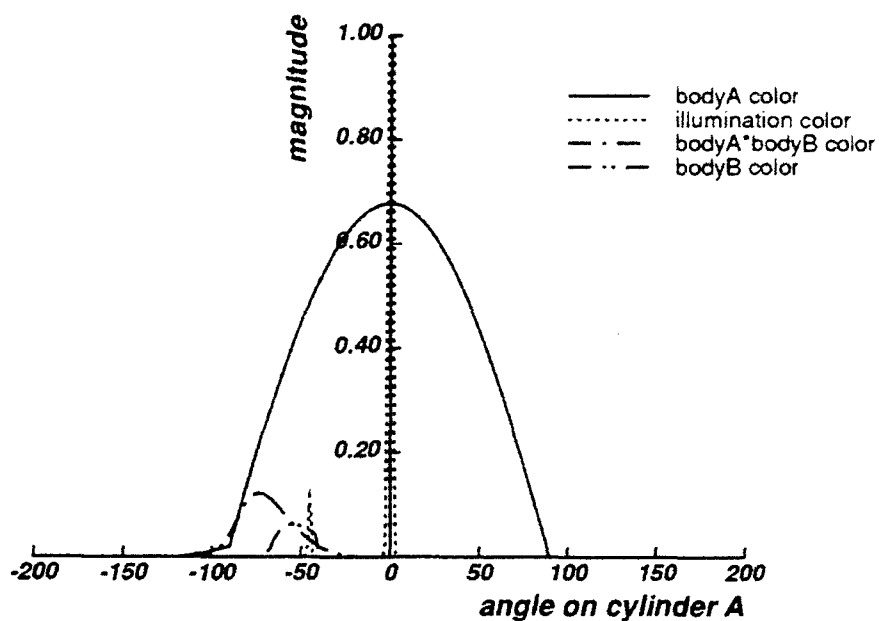


Figure 4.2: Interreflection components, by hue

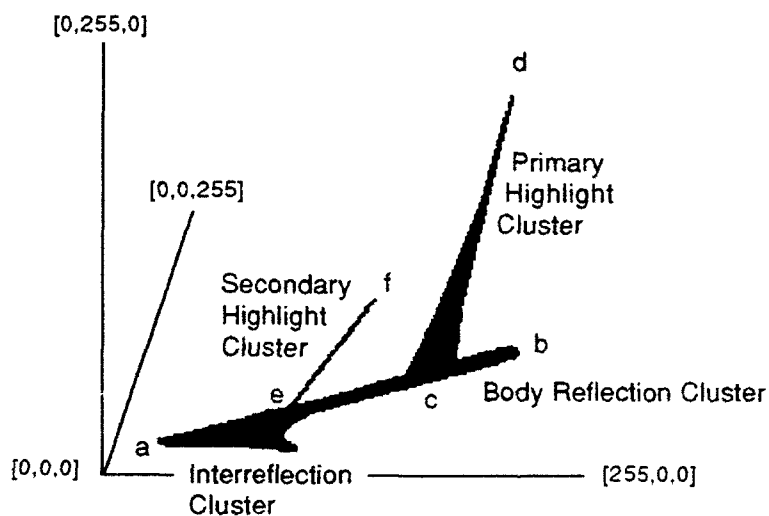


Figure 4.3: Histogram for interreflection

in figure A.1.

However, the secondary highlight cluster is absent from the histogram corresponding to the matte balls in figure A.2. Accordingly, a secondary highlight is not discernible in the original image.

4.2.1 Length of the Secondary Highlight Cluster

The billiard balls in figures A.1 and A.2 demonstrate that the secondary highlight depends upon the surface roughness. A primary highlight on surface *A* has a brightness related to its roughness parameter σ as explained in section 3.2.2. When a second surface *B* reflects a mirror image of this primary highlight, the brightness of the resulting secondary highlight must vary as the brightness of the primary highlight. If *A*'s primary highlight is dimmer, then *B*'s secondary highlight must be dimmer also. Furthermore, the brightness of the secondary highlight on *B* also depends upon the reflectance characteristics of *B*. If the surface of *B* is optically rough, then it will scatter light that impinges upon it. This applies to light that comes directly from the light source, but also to light that is reflected from other objects. As *B* gets rougher, the secondary highlight will become more scattered and hence less bright.

The brightness of the secondary highlight determines the length of the secondary highlight cluster. This is because larger amounts of surface-to-surface reflection result in pixels displaced further from the body reflection cluster. This length (the distance from points *e* to *f* in figure 4.3) depends upon the roughness of both surfaces. If both surfaces are optically rough, the secondary highlight may be completely missing, resulting in a length of 0. If both surfaces are smooth as in figure A.1, the secondary highlight cluster will have a substantial length.

Notice however that the secondary highlight cluster in figure A.1 does not reach the white point of the cube, indicating that the secondary highlight does not saturate the camera. In this picture the primary highlight did saturate the camera, so its true length cannot be measured from the histogram. But when the surface reflection is reflected a second time, only a fraction of the incoming light is reflected. Even though the primary highlight is too strong for the camera, the secondary highlight is within its dynamic range. Thus the secondary highlight provides a second constraint upon the roughness of the object, with the benefit that it is much less likely to saturate the camera.

4.2.2 Direction of the Secondary Highlight Cluster

The secondary highlight cluster grows upwards from the body reflection cluster as the amount of surface-to-surface reflection increases. This is similar to the case for the primary highlight cluster, which grows upward as the amount of surface reflection increases. Assuming that reflection at the interface does not affect the color of reflected light, then surface-to-surface reflection should have the same color as surface reflection.

However, after examining the histogram in figure A.1, the direction of the secondary highlight cluster appears to be different from that of the primary highlight cluster. Although the direction of these clusters is mostly determined by the direction of increasing illumination color, it is also affected by the amount of color change in the other reflection components over the area for which the highlight occurs. Section 3.2.1 showed that for relatively smooth surfaces, the amount of body reflection changes slowly over the area of the primary highlight, so that the direction of the associated cluster is very nearly the direction of the illumination color.

The same claim cannot be made for the secondary highlight. The region of the object where it tends to occur is often experiencing a rapid change in body reflection. Furthermore it may also be experiencing a rapid change in the amount of other interreflection components. In figures 4.1 and 4.2, the secondary highlight is centered around -45 degrees. In this region of the cylinder, the body color of cylinder A is changing faster than it is in the area of the primary highlight (centered around 0 degrees). Furthermore, the body-to-body and surface-to-body components are dropping off in this region. The changes in these color components affect the direction of the secondary highlight cluster. Thus a vector fitted to the secondary highlight cluster may deviate even more from the illumination color than the vector fitted to the primary highlight cluster.

4.2.3 Intersection Point of the Secondary Highlight Cluster

The location of a secondary highlight on an object depends upon the relative positions of the objects and upon the positions of the light source and camera. For surfaces that exhibit only perfect mirror-like reflection, a secondary highlight on *B* occurs only where the angles of incidence and reflectance are equal on object A, the angles of incidence and reflectance are equal on *B* and the reflectance ray from A is coincident with the incidence ray on *B*. The diagram in figure 4.4 shows the angles for which this occurs when two cylinders are aligned so the line connecting them is perpendicular to the optical axis of the camera, and the camera and

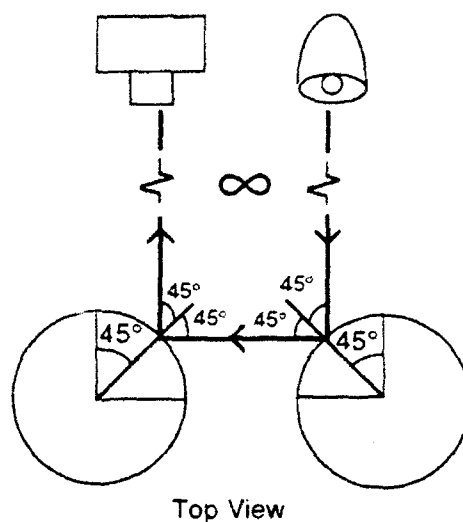


Figure 4.4: Location of a secondary highlight

light source are at infinity.

If the light source is moved so that it is separated from the camera by a non-zero phase angle relative to the objects, the angles for which perfect specular reflection occurs will change. This is also the case when the cylinders are moved relative to each other. Furthermore because of off-specular peaks, the brightest point of the secondary highlight will not always be at the perfect specular angle.

Thus, it is surprising to find that the position of the secondary highlight on the cylinder actually changes very little as the light source and camera are separated by increasing angles. Figure 4.5 shows the change in the location of the secondary highlight for the simulated cylinders as the light source is moved relative to the camera. The cylinders remain aligned such that the line connecting their center is perpendicular to the optical axis of the camera. The horizontal axis of the graphs shows the phase angle between the camera and light source, with negative numbers indicating that the light source is to the left of the camera when facing the objects. The vertical axis is the reflectance angle on the cylinder at which the secondary highlight has its peak. Notice that when the camera and light source are separated by zero degrees, the peak occurs at -45 degrees as shown in figure 4.4. Over the whole range of phase angles shown, the secondary highlight only moves from -41 degrees to -54 degrees.

However, examining the point in the histogram where the secondary highlight cluster meets the body reflection cluster shows that this point is highly variable. Figure 4.6 demonstrates how the intersection point varies with imaging geometry. The vertical axis is the distance along the body reflection vector as a fraction

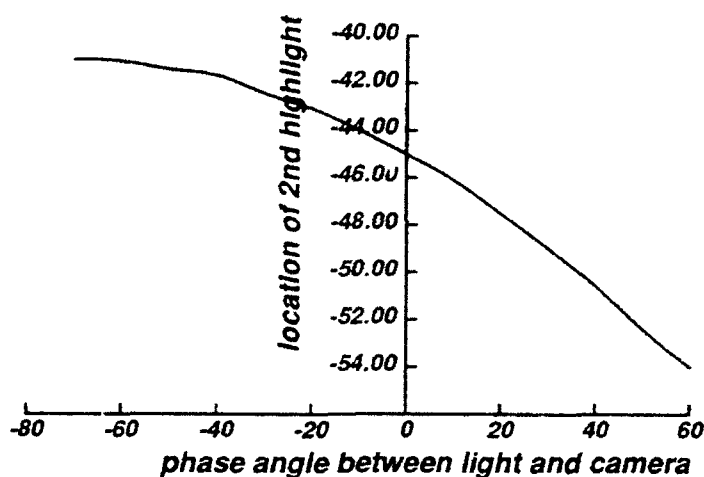


Figure 4.5: Angle at which secondary highlight occurs

of its total length (the length of $\overline{ae}$ divided by the length of $\overline{ab}$ in figure 4.3).

Even though the secondary highlight moves very little as the light is moved around, the body reflection is directly following the light source. When the light is 45 degrees to the left of the camera (at -45), the secondary highlight is actually at the maximum point of body reflection (100%). Looking back to figure 3.13, it can be seen that the primary highlight's intersection point for that same imaging geometry is at 92% of the maximum body reflection. This means that the primary and secondary highlight clusters have "crossed over" in the histogram, with the secondary highlight cluster actually intersecting the body reflection cluster *above* the primary highlight cluster. The histogram for a simulated red cylinder in this imaging situation is graphed in figure 4.7. (Notice that the primary highlight cluster is bent when it reaches the camera maximum of 255.)

As with the primary highlight cluster, the intersection point of the secondary highlight cluster also depends on the surface roughness. If the highlight is scattered over a wider range of reflectance angles, the $1/\cos(\theta_r)$ term varies more across the highlight, causing off-specular peaks. Just as off-specular peaks affect the intersection of the primary highlight cluster and body reflection cluster, they also change the intersection of the secondary highlight cluster. This dependence is extremely small, as shown in figure 4.8. For very rough surfaces, the secondary highlight will probably be absent altogether.

Assuming the secondary highlight cluster can be found in the histogram, its intersection point may give a more precise estimate of imaging geometry than the primary highlight cluster, since the variation with angle is greater. However, there is an ambiguity centered about -45 degrees. Moreover the graphs in figures 4.6

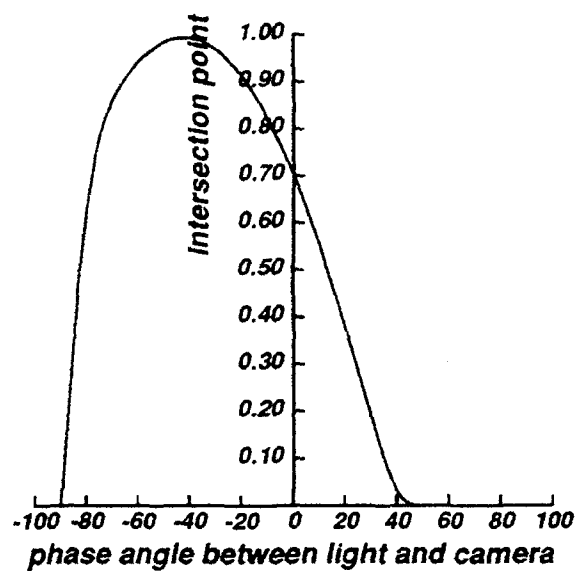


Figure 4.6: Dependence of secondary highlight intersection point upon geometry

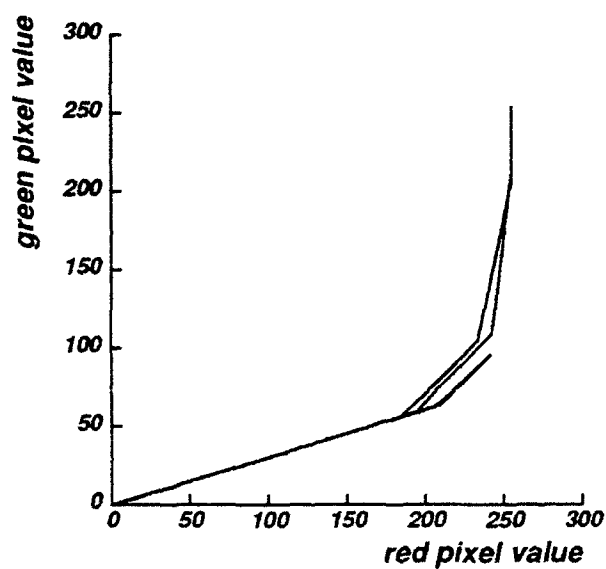


Figure 4.7: Histogram where primary and secondary highlight clusters have crossed over

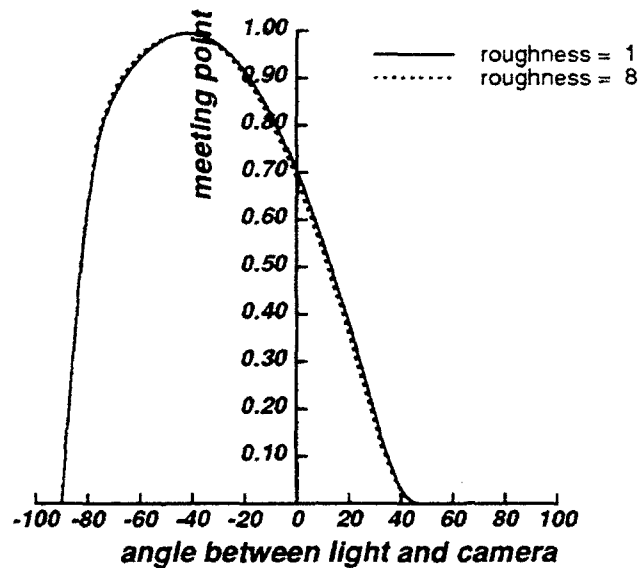


Figure 4.8: Dependence of intersection upon geometry for different roughness values

and 4.8 only apply to situations where the cylinders are aligned as previously described. For more general viewing situations and object shapes, the graphs may be different. Fortunately, it is already possible to make an estimate of light position from the graph in figure 3.14 for the primary highlight cluster. The position of the secondary highlight will allow a resolution of that graph's ambiguity about zero. The secondary highlight will occur on a side of object *A* that faces the other object *B*. If the secondary highlight cluster intersects the body reflection cluster in the upper region, that implies the light is also on that same side. Conversely, if the secondary highlight intersection occurs in the lower portion of the body reflection cluster, that indicates that the light is on the opposite side.

If the primary and secondary highlight clusters can change places in the histogram, it will be necessary to find a way to distinguish them. Since the secondary highlight cluster will be shorter than the primary one, the length can be used to tell them apart. Somewhere around -30 degrees the graphs show that both clusters have an intersection point at 97% of the length of the body reflection cluster. If the direction of these two clusters is also the same, they will be indistinguishable. Also if both surfaces are matte, the secondary highlight cluster will be completely absent.

4.3 The Interreflection Cluster

The interreflection cluster refers to those pixels that fall outside the dichromatic plane defined by the body reflection vector and illumination vector. These pixels are precisely those that experience the hue shift seen in interreflection. They cause considerable trouble for segmentation algorithms since they cannot be explained by the regular dichromatic reflection model. In figures A.1 and A.2, some of the pixels from the yellow ball have the same color as pixels from the red ball, so they could not be correctly segmented based on hue.

Still, some things can be said about the interreflection cluster. If the secondary highlight cluster is caused by surface-to-surface reflection, the interreflection cluster must result from some combination of the other three types of interreflection: body-to-surface, surface-to-body, and body-to-body reflection.

4.3.1 Surface to Body Reflection

The surface-to-body reflection component of interreflection causes few difficulties for computer vision algorithms. Assuming that surface reflection does not color the illumination, then the illumination from the surface of an object *A* that is reflected by the body of object *B* will have the same hue as light that comes straight from the light source to the body of *B*. The object *A* acts a duplicate light source next to *B*. Any pixels that have a significant surface-to-body component will merely be moved further up the body reflection cluster since they have the same hue but are brighter. For example, points beyond the self shadow line of *B* would normally be at the black point of the color cube. If they face a lit side of object *A* they will have a small amount of the body reflection color of *B*, moving them further up the body reflection cluster for *B*.

4.3.2 Body to Body Reflection

The body-to-body reflection component of interreflection is the one that has most often been explored by vision researchers [Horn, 1977], [Forsyth and Zisserman, 1989], [Nayar *et al.*, 1990], [Jang, 1991], [Funt and Drew, 1991]. They all make the assumption that the surfaces they are dealing with are completely matte and so ignore the other three components of interreflection. The work in [Bajcsy *et al.*, 1989] works on objects with highlights, but where the interreflection has a small enough effect on hue to be segmented along with non-interreflected pixels.

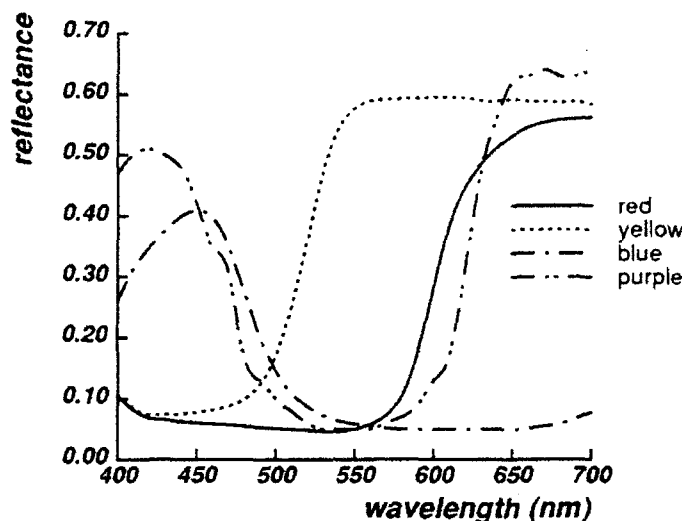


Figure 4.9: Spectral reflectances of four colors

The body-to-body reflection color is the product of three spectral functions: the illumination color, the body color of the first object, and the body color of the second object. This multiplication takes place at each wavelength within the visible spectrum. Suppose for discussion purposes that the color of the illumination is white so that its influence can be ignored. Then the spectral signal resulting from body-to-body reflection color will be high only for those portions of the spectrum where both body colors are high. The resulting spectral signal will be referred to here as the "overlap color" since it reflects light only where the spectra of the two original colors overlap.

To illustrate the effects of color overlap, this section will examine the effects for four different colors of objects: red, yellow, blue and purple. The spectral reflectance curves for these objects are shown in figure 4.9. The colors associated with these curves look strong, or saturated, with very little added white. The fact that each of the curves is near zero for some portion of the spectrum means that no white could be subtracted out to make the color look more pure. These spectral signals are turned into RGB values by multiplying them by the camera's spectral responsiveness in each of the three color bands and then integrating.

For the case of body-to-body reflection between a red object and a blue one, the color overlap is near zero. This is because the spectral reflectance of the red object is high in the longer wavelengths where the blue reflectance is low, and the blue reflectance is high in the shorter wavelengths where the red reflectance is low. There is no part of the spectrum for which both colors reflect appreciable amounts of light. Thus the body-to-body reflection component of interreflection between two objects of these colors will be practically

zero.

If matte blue and red objects are placed next to each other, they will not show a hue shift since the body-to-body reflection is insignificant. Their color histograms will show this fact. A hue shift would cause pixels to fall outside the dichromatic plane, causing the formation of the interreflection cluster. If there is no body-to-body reflection and no hue shift, then all pixels of the two objects will fall within the dichromatic plane. If it is assumed that the objects are completely matte, even those points on the objects that face each other will be located on the body reflection clusters in the histograms.

For the case of body-to-body reflection between a red object and a yellow one, the overlap color will be red. This is because both of these colors reflect a lot of light in the red region of the spectrum. Although the yellow color has a lot of reflectance in the green portion of the spectrum, the red has practically none here. Neither color shows any appreciable reflectance in the blue portion of the spectrum.

Consider what happens when the red body-to-body interreflection is added to the simple body reflection of each of the objects. For the red object, there actually is no noticeable hue change as a result of this type of interreflection. This is because both the body and the body-to-body components are red, and the result of adding red light plus some more red light merely yields red light. Pixels that are influenced by this color overlap will be *brighter* than they would have otherwise been, but the hue will be the same. Figure 4.10 shows the spectral product of the red and yellow colors and also shows what happens when this overlap color is added to the original red color. The resulting color is essentially just a scaled up version of the original red color.

On the other hand, if the red body-to-body interreflection is added to the simple body reflection of the yellow object, a hue change will be noticed. Figure 4.10 shows that when the red overlap color is added to the yellow color, the resulting spectral reflectance curve has a significantly different shape. Thus pixels on the yellow object that have significant amounts of body-to-body reflectance will have a noticeably different hue.

The conclusion from this is that although the body-to-body reflection *color* is symmetrical, the influence on object appearance is not. If matte red and yellow objects are placed next to each other, the resulting histogram will show an interreflection cluster for the yellow object, corresponding to the pixels that are hue-shifted. But the histogram for the red object will have no interreflection cluster. Any pixels that have body-to-body reflection will have the same red hue as pixels without interreflection. As a result these pixels

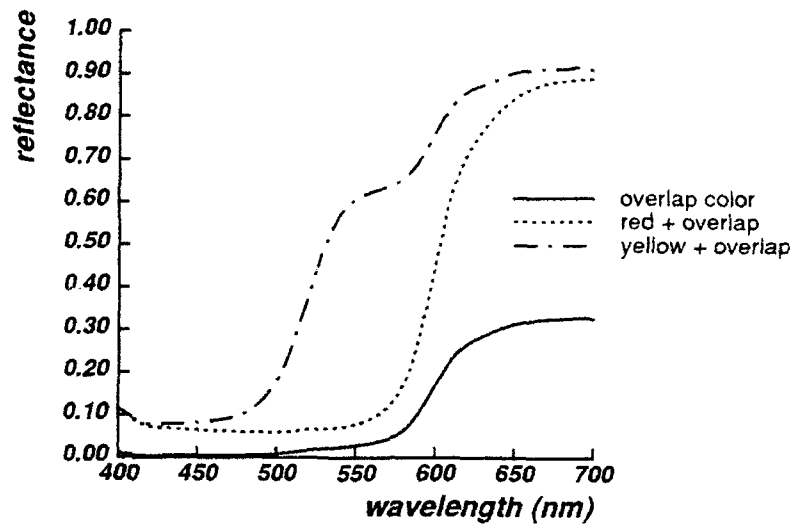


Figure 4.10: The color overlap and its effect

will fall on the body reflection cluster.

The interreflection cluster on the histogram for the yellow object will correspond to those points on the surface that receive increasing amounts of the body-to-body reflection color of red. As a point receives more body-to-body reflection, it will move further from the yellow body reflection cluster, in the direction of increasing amounts of red. As described in [Funt and Drew, 1991] and [Jang, 1991], these pixels will fall on a plane defined by the simple body reflection color (yellow) and the overlap color (red). Since the overlap color is essentially the same as the simple body reflection color of the red object, the pixels will fall on the plane between the two vectors fitted to the body reflection clusters.

Finally, consider the case where the two objects of interest are yellow and purple. Figure 4.9 shows that the yellow color has high reflectance in the red and green regions of the spectrum, while the purple color has high reflectance in the blue and red portions of the spectrum. This means that the color overlap will again have a red color. However this time, the overlap color is different from both the object colors.

For the case of interreflection between matte yellow and purple objects, both objects will have a body-to-body reflection component of red. Thus both objects will have a noticeable hue shift. This will be manifested in the histograms as interreflection clusters that deviate from the body reflection clusters. Both clusters will be in the direction of increasing amounts of red. For the yellow object, the pixels experiencing interreflection will fall on the plane defined by the vector fitted to the body reflection cluster and the vector that points in the direction of the red overlap color. Similarly for the purple object, the interreflection pixels

will fall on the plane defined by the vectors for the purple body color and the red overlap color

Taken together, the interreflection pixels from both surfaces form planes that intersect at the overlap color. This is the situation described in [Funt and Drew, 1991]. They use this fact to automatically detect the overlap color for matte objects; it can then be subtracted from pixels to form an image with the interreflection removed. However this technique only works if the overlap color is different from *both* body reflection colors, and it also assumes completely matte objects.

4.3.3 Body to Surface Reflection

Many objects in the real world are not matte. Shiny plastic or painted objects can mimic a mirror, giving a very clear reflection of nearby objects. This effect is a result of the body-to-surface component of reflection. It is especially noticeable for shiny objects that are dark, where the body reflection is low and does not distract the eye.

Surface reflection is always present in inhomogeneous objects. Any time light encounters an interface with a change in the index of refraction, some of the light will be immediately reflected back at the interface. Thus even the matte objects that have been modeled as Lambertian will experience surface reflection. These surfaces are optically rough enough that the surface reflection is scattered in every direction, making it indistinguishable from body reflection. This explains why a black shiny object looks darker than a black matte one. With a matte black object, some of the surface reflection is scattered in the direction of the eye, making the object look lighter. A shiny black object reflects the surface reflection in a particular direction; if the surface is viewed from some other direction, that surface reflection is not seen, making the object look darker.

If an object is shiny, it may reflect some of the color of a neighboring object directly into the camera. Consider a red object again, but this time a shiny one. If a yellow object is nearby, some of the yellow body reflection from it may be reflected at the surface of the red object, causing the appearance of a yellow reflected image on the red surface. This will cause a noticeable hue shift on the object, and hence affect the histogram.

The yellow body-to-surface component of interreflected pixels is added to the other reflection components at that point. It will be added to the red color from the simple body-reflection and also to the red color of

the body-to-body reflection. In the histogram, these pixels will form an interreflection cluster that deviates from the body reflection cluster, in the direction of increasing amounts of yellow.

If the surface of the red object is optically rough, it will still have a body-to-surface component from the yellow object. However it will scatter that reflection widely so that most of it misses the camera. In this case there will be no noticeable influence from the yellow object, and hence no interreflection cluster in the histogram. This means that the deviation of the interreflection cluster from the dichromatic plane depends not only upon the colors of the two objects, but also their degree of roughness.

If both the body-to-body reflection and the body-to-surface reflection are high, the resulting histogram depends upon the body color of the other object and the overlap color of the two objects. In the case of the shiny yellow object next to the red object, these are the same, since the body-to-surface color and overlap color are both red. Both components of interreflection will cause the interreflection cluster to lie on the plane defined by the yellow and red vectors.

However, for some objects these two components will have different colors. This is the situation when a shiny yellow object is next to a purple one. The body-to-body color will have a red hue whereas the body-to-surface color will have a purple hue. In this case the interreflection cluster may not lie neatly in a plane, but fall between the plane defined by the yellow and red vectors and the plane defined by the yellow and purple vectors.

4.3.4 Analysis of the Interreflection Cluster

After describing the influence of each of the components of interreflection, it is now obvious that the interreflection cluster in the histogram is affected by several things. If it is known from scene geometry that interreflection may occur, and the body colors and roughness values of the objects involved are also known, one could reason forward and say how the colors will be distributed in the histogram. This is essentially a graphics problem, where the objects are modeled first, and then the image is created. In vision problems researchers start with the image and want to generate a model of the objects.

However, after examining the interreflection cluster of a histogram, it is possible to make some predictions about the object. Section 4.3.2 showed that the interreflection cluster will be absent if the objects are matte and there is no color overlap between the colors, or if the color overlap has the same color as the body color.

This thesis has already shown how to estimate object roughness from other histogram features. If the object has been determined to be optically rough, then the absence of an interreflection cluster puts constraints upon what color objects could be close by.

If on the other hand, analyzing other features in the histogram has indicated that the object is very smooth, then it is likely to have body-to-surface reflection. The plane formed by the interreflection cluster will indicate directly the color of the other object.

Alternatively if the colors of both objects are known, and the interreflection cluster has been identified from the image, it might be possible to break down the image into interreflection components. This analysis of the histogram would enable one to remove not only the influence of body-to-body reflection as is shown in [Funt and Drew, 1991], but also the other components that are present when surface reflection is considered.

The end result would be additional "intrinsic images." Shafer originally proposed breaking down an image with highlights into an intrinsic image of body reflection and an intrinsic image of surface reflection [Shafer, 1984]. Klinker showed that these images may be automatically generated from the original image, showing what objects would look like if they were completely matte (body reflection only) and how they would appear if made of a shiny black material (surface reflection only) [Klinker, 1988]. However, the method assumed no interreflection. An analysis of the components of interreflection could be used to show how an object with highlights will appear if nearby objects are removed from the scene. It might also be used to show how the interreflection would appear if the objects were completely matte, or how it would look between objects that only exhibit surface reflectance. Automatic analysis of images with interreflection is a subject for future work.

Chapter 5

Analysis of the Color Histogram for an Ideal Sensor

Chapters 3 and 4 described the relationship between the shape of the color histogram and various interesting scene properties. Understanding the relationship is the first step in analyzing color histograms. The next step is figuring out how to actually exploit the histogram to recover quantitative measures of the scene properties.

This chapter will assume an ideal sensor that has highly desirable properties. This kind of camera, which is unobtainable in practice, shows the limit of what could be obtained under the best possible imaging conditions. Chapter 6 will show how the requirements may be relaxed for the case of a more realistic camera.

5.1 Properties of an Ideal Sensor

For the purposes of this thesis, an ideal sensor is one that obeys the imaging equation

$$p_x = \gamma \int L(\lambda, \theta_i, \theta_r, \theta_p) S_x(\lambda) d\lambda \quad (5.1)$$

where p_x is the pixel measurement in each band and x is one of red, green, or blue. $S_x(\lambda)$ is the responsivity of the sensor in each of the bands; the integration is performed over those wavelengths where S_x is greater than zero.

Typically the non-zero wavelengths are assumed to be those wavelengths of light that are visible to

humans, but that need not be the case here. Many of the CCD cameras used in image understanding are very sensitive to infrared wavelengths and may result in images that look somewhat surprising to humans [Novak *et al.*, 1990]. However that does not affect the dichromatic reflection model, as long as the camera is able to sense wavelengths where there is a difference in the amount of light caused by body reflection as opposed to surface reflection.

Equation 5.1 implies that the camera behaves linearly since

$$\int n L(\lambda, \theta_i, \theta_r, \theta_p) S_x(\lambda) d\lambda = n \int L(\lambda, \theta_i, \theta_r, \theta_p) S_x(\lambda) d\lambda \quad (5.2)$$

for all constants n . This means that if the light intensity or surface reflectance is increased, the pixel value will be proportionally larger. In practice, many cameras are reasonably linear in their response. Even when they are not, a linear correction is easy to apply once the non-linearity has been measured [Novak *et al.*, 1990].

Equation 5.2 also implies that the camera has an infinite dynamic range since n may equal 10, 100 or even a million. Obviously real cameras do not have this property; they have some maximum value that can be measured. After this maximum has been reached, increasing the brightness of a point will not cause an increase in the corresponding pixel value. The result is that the pixel value is "clipped" at the maximum value of the camera. In the ideal model, clipping does not occur.

When taking real pictures, researchers can usually adjust the aperture, light level, or exposure time so that the brightest point in the image is within the dynamic range of the camera. However this presents problems in scenes that contain highlights. Highlights are much, much brighter than surrounding non-highlight areas, so that if the camera is adjusted to measure highlights properly, the non-highlight areas will be so dim as to be lost in the camera noise.

Since it is precisely pictures with highlights that this thesis is concerned with, the assumption of unlimited dynamic range is much less realistic than the linearity assumption. This assumption will be removed in the next chapter.

A final assumption made about the ideal sensor is that it does not suffer from any noise problems. Shot noise, dark current noise, and pre-amplifier noise are three sources of noise that occur in real CCD cameras [Photometrics, 1989], [Healey and Kondepudy, 1991]. Equation 5.1 does not take these effects into consideration. In the next chapter these effects will be modeled and taken into consideration in the analysis

of histograms.

5.2 Knowns and Unknowns

The known image parameters which can be measured from the color histogram of a single object are the body reflection cluster's length and direction; the highlight cluster's length, width, and direction; and the intersection point of the two clusters. This gives four scalar values and two vector quantities. They will be referred to by the following variables:

- l - length of highlight cluster
- w - width of highlight cluster
- i - intersection of two clusters
- b - length of body reflection cluster
- $\vec{d}_h$ - direction of highlight cluster
- $\vec{d}_b$ - direction of body reflection cluster

The unknown scene parameters which are to be recovered from the histogram can also be divided into scalar values and vector quantities. The scalar values are the surface roughness, the phase angle between the light source direction and camera direction, the light intensity, and the surface albedo. The vector quantities are the chromaticity of the illumination and that of the object. These variables are:

- σ - optical roughness
- θ_p - phase angle
- B_s - illumination intensity
- B_o - object albedo
- $\vec{c}_s$ - chromaticity of light source
- $\vec{c}_o$ - chromaticity of object under "white" light

For convenience, the illumination and object reflectance have been separated into intensity components and chromatic components. As described in section 3.1 the object's chromaticity $\vec{c}_o$ may be recovered in a straightforward way from the direction of the body reflection cluster $\vec{d}_b$ and the color of the light source $\vec{c}_s$. The red component of the body reflection cluster is divided by the red component of the light source color, the green component divided by the green component, and the blue component divided by the blue component.

$$\vec{c}_o = [d_b^R / c_s^R, d_b^G / c_s^G, d_b^B / c_s^B]$$

The result is then normalized to length 1. This recovers only two parameters of the object's reflectance (a third parameter is albedo), but this is a useful approximation in most cases, since most objects have smoothly varying reflectance functions.

The albedo of the object B_o is also recovered in a straightforward manner from the length of the body reflection cluster b and the illumination intensity B_s . From equation 2.1 the body reflection magnitude is

$$m_b = \gamma B_b \cos(\theta_i)$$

where B_b is the object's apparent brightness. The body cluster length b is equal to the maximum value of m_b , which will occur when $\cos(\theta_i) = 1$. Therefore $b = \gamma B_b$. Since the apparent brightness of the object is the product of the object albedo B_o and the illumination intensity B_s , the albedo is

$$B_o = B_b / B_s = b / (\gamma B_s)$$

For the remaining unknowns the situation is not so simple. In chapter 3 it was shown that the length l is related to surface roughness and illumination intensity, but is also dependent upon imaging geometry and even the illumination color. The remaining knowns

- l - length of highlight cluster
- w - width of highlight cluster
- i - intersection of two clusters
- $\vec{d}_s$ - direction of highlight cluster

and unknowns

- σ - optical roughness
- θ_p - phase angle
- B_s - illumination intensity
- $\bar{c}_s$ - chromaticity of light source

will be examined in detail in the next few sections.

5.3 Analytic Solution

Equations 2.1 and 2.3 describe the amounts of body and surface reflection m_b and m_s as a function of imaging geometry and light intensity B_s . Equation 2.3 also shows how the amount of surface reflection at a given point varies with the roughness of the surface σ . The equations again are:

$$\begin{aligned} m_b &= \gamma B_b \cos(\theta_i) \\ m_s &= \gamma B_s \frac{F G(\theta_i, \theta_r, \theta_p) \alpha}{\sigma \cos(\theta_r)} \exp\left(-\frac{\theta_s^2}{2\sigma^2}\right) \end{aligned}$$

As mentioned in section 3.2.2, the Fresnel coefficient will be assumed to be 1.5. The Fresnel coefficient is only weakly dependent upon wavelength and incidence angle, and is very close to 1.5 for nearly all paints and plastics. The constant α includes the facet size and some constants of integration. The facets of the surfaces in question will be assumed to be of uniform size.

Unfortunately it is not possible to directly solve for the scene parameters from the histogram measurements. The length of the highlight cluster indicates the maximum amount of surface reflection seen anywhere on the object. For given values of σ , θ_p , and B_s the length l may be calculated

$$l = \text{MAX} \left[\gamma B_s \frac{F G(\theta_i, \theta_r, \theta_p) \alpha}{\sigma \cos(\theta_r)} \exp\left(-\frac{\theta_s^2}{2\sigma^2}\right) \right] \text{ over all values of } \theta_i, \theta_r, \theta_s \quad (5.3)$$

If it were not for the presence of G in the numerator and $\cos(\theta_r)$ in the denominator, the length would be easily derived. If they disappeared from the equation, the maximum value of the surface reflectance would occur at the perfect specular angle when θ_s is zero.

In order to derive the roughness from the length measurement with any accuracy, it is necessary to

consider the influence of G and $1/\cos(\theta_r)$. Recall from equation 2.4 that the formula for G is given as

$$G = \text{MIN} \left\{ 1, \frac{2 \cos(\theta_s) \cos(\theta_r)}{\cos(\theta_p)}, \frac{2 \cos(\theta_s) \cos(\theta_i)}{\cos(\theta_p)} \right\} \quad (5.4)$$

However G is not the sole source of difficulty for a direct solution. It is significant that the roughness parameter σ occurs both in the denominator and in the exponential of equation 5.3. Even if all terms besides σ were assumed to be constant and equal to 1, the equation would reduce to

$$l = \frac{1}{\sigma} \exp\left(-\frac{1}{\sigma^2}\right)$$

which has no analytic solution for σ . Clearly a direct solution of the complete surface reflection equation cannot be obtained.

5.4 Ideal vs. Measured

Although equation 5.3 has no analytic solution, it might be possible to solve it iteratively, through some sort of search (for example by gradient descent). However that assumes the length l can be accurately measured from the histogram. In section 3.2.2, the length is measured from the tip of the highlight cluster to its base *along the direction of the highlight color*. Unfortunately the highlight color is not typically known in advance. The highlight color may be estimated by fitting a vector to the pixels that form the highlight cluster, but as described in section 3.2.1, that estimate may be skewed.

This is shown in figure 5.1, which is the histogram of a simulated rough object illuminated with white light. The dotted "measured" line shows the *direction* calculated for the best fit vector to the highlight cluster. In this case the best fit line to the cluster will not pass through the brightest point in the highlight cluster. The *position* of the dotted line in figure 5.1 shows the projection of the brightest highlight pixel onto the body reflection vector *along the best fit vector*, since the length is calculated from the brightest pixel. The "ideal" line indicates the direction of the actual illumination color. The skewing causes the measured length to be longer than it would have been if the correct illumination color had been known.

This contrasts the concept of "ideal" length that one would like to obtain from the histogram, with the "measured" length that can actually be recovered without a priori knowledge about the illumination color. All the graphs presented in chapter 3 that relate highlight cluster length to roughness or imaging geometry used the ideal length, obtained by knowing the illumination color.

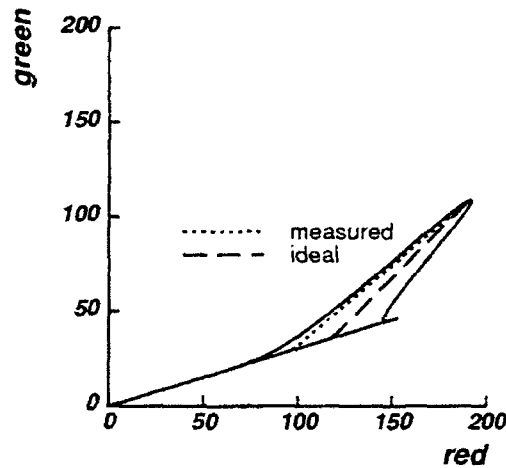


Figure 5.1: Skewed length measurement

Figure 5.2 shows how both the ideal length and measured length vary with roughness for the case where the camera and light source are separated by 20 degrees. For smooth surfaces both values show an inverse relationship between length and roughness. As the roughness increases, the deviation between the two curves grows; eventually the measured length begins *increasing* as the amount of roughness is increased. If only the measured length value is available, it would be difficult to tell a very smooth surface from a very rough one.

The variation of the "ideal" length with roughness *and* imaging geometry is shown in the surface presented in figure 5.3. The length is shown as the height of the surface, while incremental changes in roughness are shown along the x-axis and changes in phase angle are shown along the y-axis.

Figure 5.4 shows the variation in the "measured" length with step changes in surface roughness and phase angle. It looks somewhat different than the graph shown in figure 5.3.

A similar situation occurs with the measurement of width and intersection point. When the highlight is skewed, the values measured from the histogram will be different from the ideal values shown in the graphs in chapter 3. The intersection point is measured by taking the brightest pixel in the highlight cluster and projecting it onto the body reflection vector. This projection is done *along the vector* that was assumed to be the highlight cluster. In the absence of *a priori* information about the illumination color, the best-fit vector to the cluster will have to be used.

The point of intersection with the body reflection cluster is supposed to represent the amount of body reflection at the maximum highlight point. However, if the highlight cluster is skewed towards the body

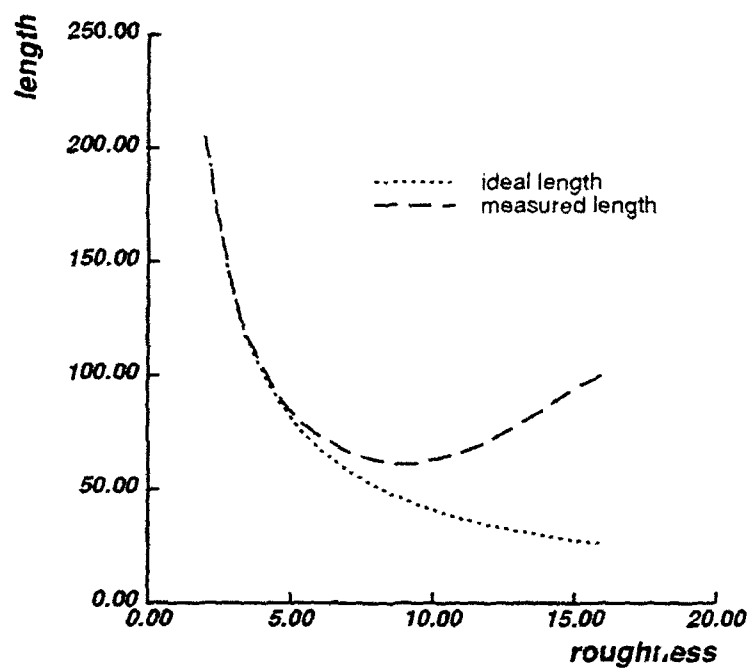


Figure 5.2: Ideal vs. measured length

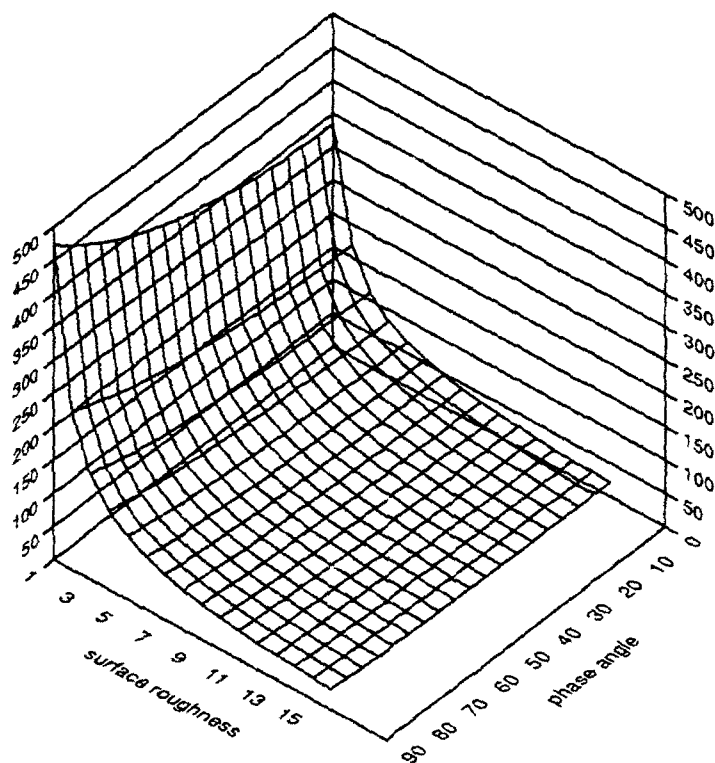


Figure 5.3: Variation in ideal length with changes in roughness and phase angle

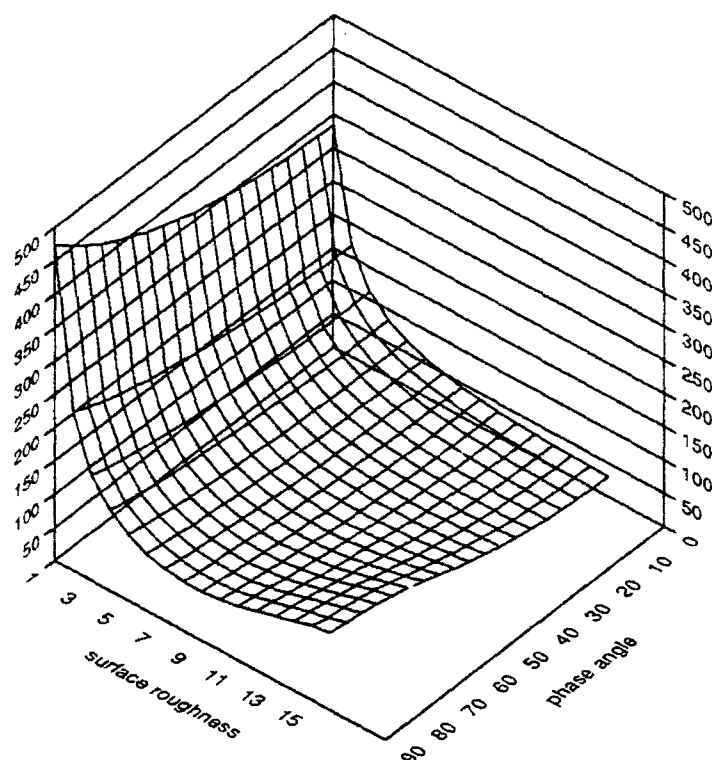


Figure 5.4: Variation in measured length with changes in roughness and phase angle

reflection cluster, the highlight pixel will be projected too far down, yielding an amount of body reflection that is too small. In some cases, the amount of predicted body reflection will be negative. This is shown in figure 5.5 where a histogram is shown for a simulated rough surface with $\sigma = 16^\circ$ and $\theta_p = 90^\circ$. The “ideal” line shows the direction of the actual illumination color and where it intersects the body reflection cluster when projected from the brightest point in the highlight. The “measured” line shows the direction of the vector fitted to the highlight cluster, as well as the projection from the brightest point along that vector.

Chapter 3 argued that intersection represents the amount of body reflection seen at the highlight peak as a fraction of the maximum amount of body reflection seen anywhere on the object. Logically speaking, it must fall between 0 and 1, since the amount of body reflection will never be negative. However the method used for measuring the amount of body reflection relies on the vector fitted to the highlight cluster. If this vector deviates from the correct highlight color, the intersection measurement will be off, and may even yield a negative value.

Figures 5.6 and 5.7 show how the ideal and measured intersection values vary with changes in the surface roughness and imaging geometry. Figure 5.7 shows that there is a range of roughness values and phase

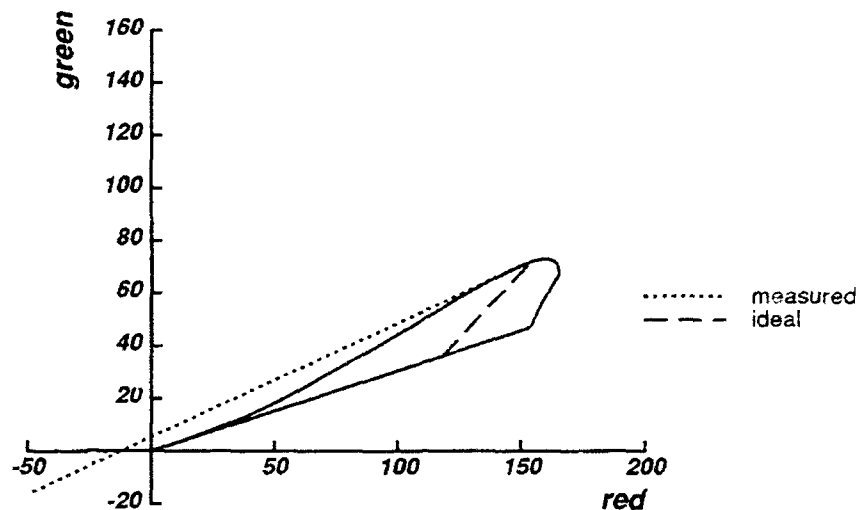


Figure 5.5: Histogram with a negative intersection

angles for which the measured intersection will be negative.

By the same token, the width measurement is also affected by skewing of the highlight cluster. The width is supposed to represent the range of body reflectance values which show a highlight compared to the range of body reflectance values seen anywhere on the object. So it too should be a fraction between 0 and 1. However if some parts of the highlight are calculated as having a negative amount of body reflectance, while others are calculated as having the maximum amount, the total range will be greater than the range of body reflectance values, giving a width measurement greater than 1. An example of this is shown in figure 5.8 which shows the same histogram from figure 5.5. Projecting all highlight pixels down along the measured highlight vector gives a width measurement greater than 1.

Figures 5.9 and 5.10 show how the ideal and measured width values vary with changes in the surface roughness and imaging geometry. Figure 5.10 shows the range of roughness values and phase angles for which the measured width will exceed 1.

In the absence of *a priori* information about the illumination color, a system can only measure what is available in the histogram. How can the ideal values be derived from the ones that are measured? Once that is done how can the image parameters of interest be recovered?

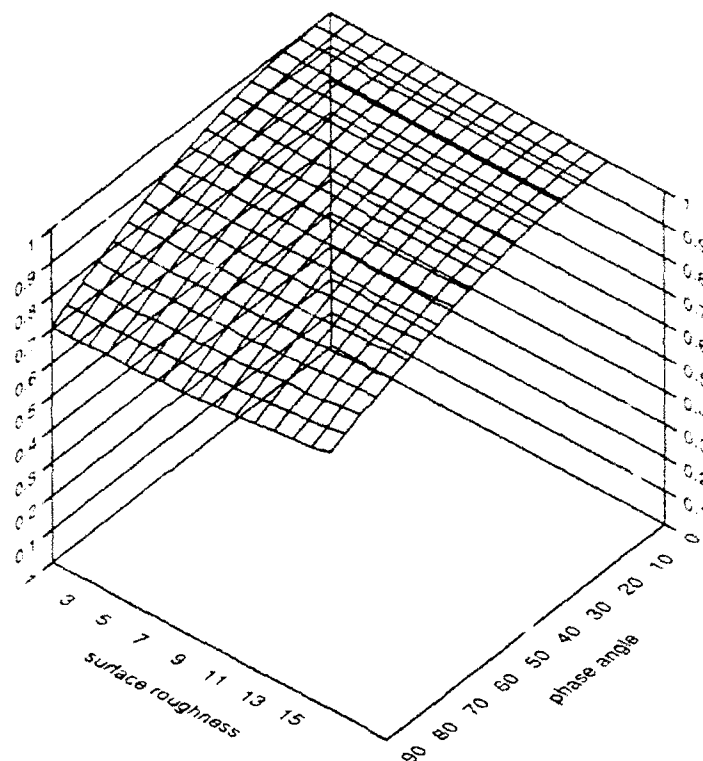


Figure 5.6: Variation of ideal intersection values

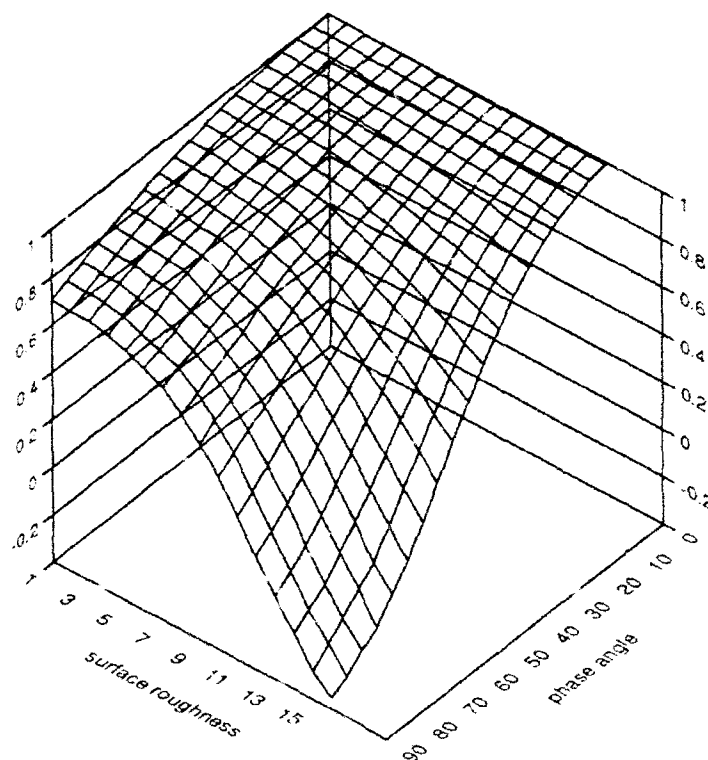


Figure 5.7: Variation of measured intersection values

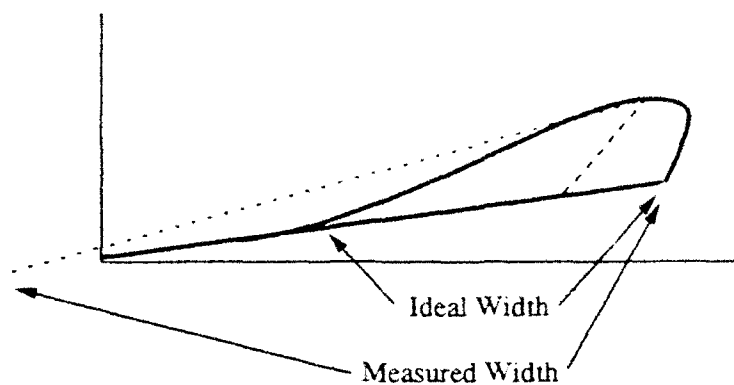


Figure 5.8: Histogram with a width greater than one

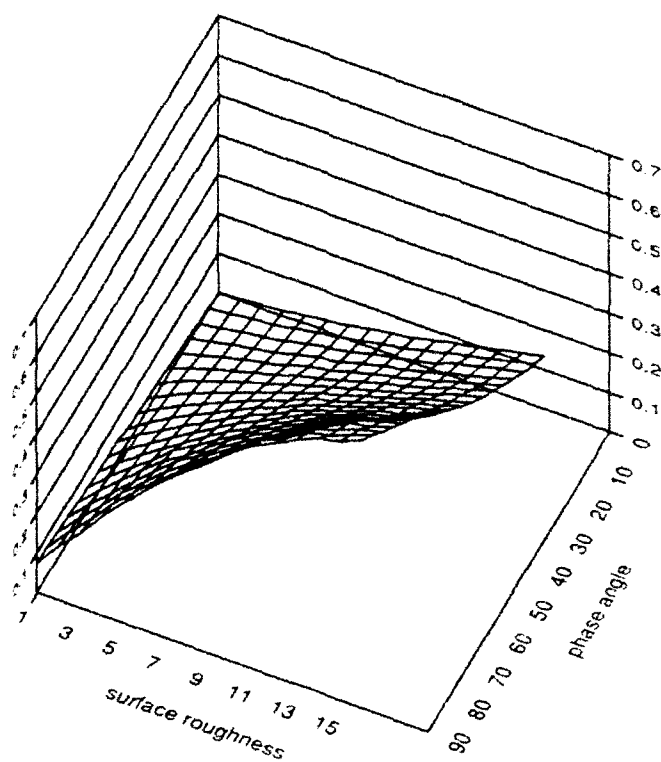


Figure 5.9: Variation of ideal width values

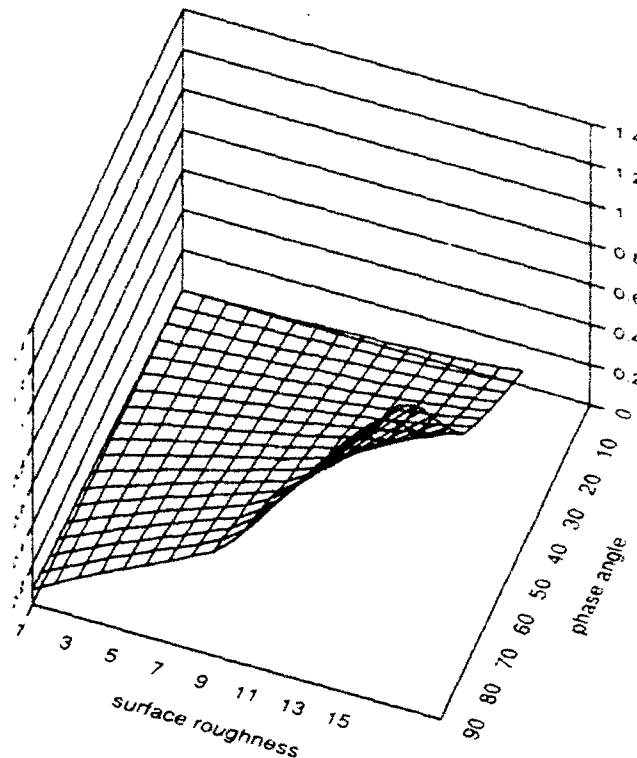


Figure 5.10: Variation of measured width values

5.5 Approximate Solution

The approach taken by this thesis is to recover scene parameters by an approximate method, *directly* from the initial histogram measurements. Therefore, it is not necessary to recover "ideal" histogram values from the "measured" ones. The ideal values of the histogram are a useful abstraction since their relationship to scene parameters is easy to explain. However, they cannot be obtained from the histogram without knowledge of the illumination color.

The surfaces shown in figures 5.4, 5.7, and 5.10 show the variation in measured length, intersection, and width as surface roughness and phase angle are varied. A given value of length describes a contour within the space of roughness and phase angles. The values for these scene parameters must lie somewhere on that contour. The additional measurements of width and intersection also have contours in the roughness-phase angle space. For a given histogram with its particular length, width, and intersection measurements, the contours may have a unique intersection in the space defined by roughness and phase angle.

The preceding discussion left out the additional scene parameter of light intensity. If that is considered

in addition to surface roughness and phase angle, these three scene parameters form a three dimensional space. A length measurement would then describe a two dimensional surface within that space, showing the possible roughness, phase angle, and light intensity values that could give rise to a histogram with that length highlight cluster. The intersection and width measurements would also describe a set of surfaces within that space. The hope is that these three surfaces intersect at a single point in that space, making it possible to recover a unique value for surface roughness, phase angle, and illumination intensity. So obvious questions are: how can these contours of equal-length, equal-width, and equal-intersection be obtained; and do these contours intersect to give unique solutions?

Section 5.3 pointed out that there is no analytic solution to generate the contours. However, the graph in figure 5.4 shows how the highlight cluster length varies with roughness and phase angle at discrete points. These values came from simulating an object with those parameters and then measuring the length of the highlight cluster in the the resulting color histogram. Obviously a detailed lookup table can be created by simulating a large range of roughness values, phase angles, and illumination intensities, and then storing the corresponding histogram measurements of length, width, and intersection in the table. Given a histogram associated with unknown scene parameters, the lookup table is searched to find the corresponding scene parameters.

The more interesting question is whether a unique solution exists for a given triple (l, w, i) . If some triple had more than one solution, that would mean that different combinations of scene parameters could give rise to identical histogram measurements. It would also mean that a search through the contours in parameter space could not be guaranteed to converge.

To visualize the distribution of possible (l, w, i) triples, figure 5.11 shows the space of all possible values for length, width and intersection. The x-axis encodes the range of highlight cluster lengths, the y-axis shows the range of widths, and the z-axis shows intersections. The surface within it is defined by lines of equal roughness in one direction and lines of equal phase angle in another direction. (In this figure, the illumination intensity is constant.) This surface shows which combinations of length-width-intersection are possible for histograms within the simulated range. Points in the space not falling on the surface correspond to histograms that do not make sense and thus cannot arise from real images.

Although the surface curves around in the L-W-I space, it does not intersect itself anywhere. This means that any triple of length-width-intersection that falls on the surface is associated with a unique set of surface roughness and phase angle values. The only remaining problem is to determine from these histogram

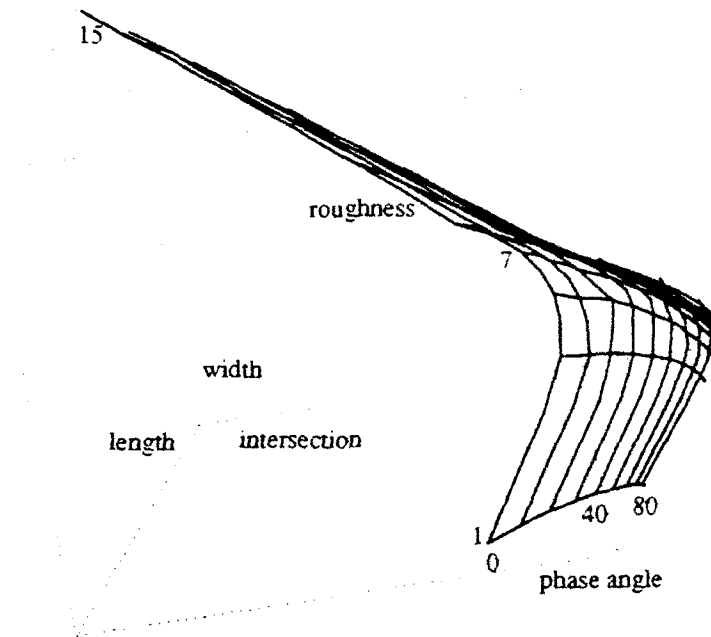


Figure 5.11: Surface of roughness and phase angles within L-W-I space

measurements where on the surface they will lie.

The surface in figure 5.11 shows only a range of roughness values and phase angles. If the range of illumination intensities were also shown, that would give rise to a three dimensional curved solid within the L-W-I space.

5.6 Generating Lookup Tables

A large range of roughness values, phase angles, and illumination intensities are used to create the lookup tables. The ranges are shown in the table in figure 5.12. A roughness value is the standard deviation of facet angles with respect to the global surface normal, in degrees. The phase angle is the angle between the camera and light source with respect to the object, also in degrees. The light intensity is a percentage of a hypothetical light's maximum output.

For each set of roughness, phase angle, and light intensity values, a simulated object is generated, and the

	Minimum	Maximum	Increment	Total Used
Roughness	1°	15°	2°	8
Phase Angle	0°	90°	10°	10
Intensity	50 %	100 %	10 %	6
Overall				480

Figure 5.12: Range of scene parameters used to generate lookup table

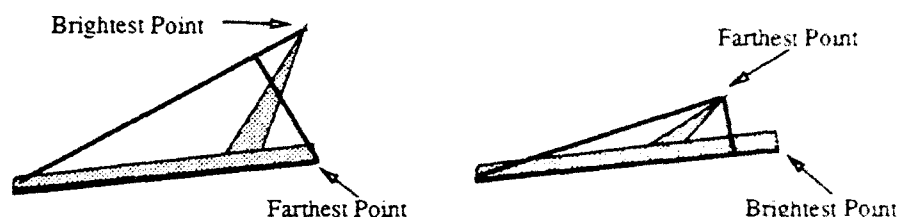


Figure 5.13: Finding the body cluster

histogram from the resulting image is measured to obtain the dimensions to be entered into the lookup table. The simulation employs the noise free camera equation 5.1, and the histogram is automatically separated into body reflection and highlight clusters. The technique is similar to that described by Klinker [Klinker, 1988].

First, a line is drawn from the black point of the color cube to the brightest point in the histogram. Then the point in the histogram farthest from that line is found. A second line is drawn from the black point to this new point. One of these two lines defines the body reflection cluster. Most often it is the second line, as shown in the histogram on the left in figure 5.13. However it could be the case that the brightest point of the histogram would be on the body reflection cluster, if the highlight were very dim. In this case the first line is the correct choice for defining the body reflection cluster. This situation is shown in the histogram on the right in figure 5.13.

The line for the body reflection cluster is chosen according to the "saturation rule". Surface reflection is assumed to have the color of the illumination. Body reflection is assumed to be the product of the illumination color and the body reflection color. Therefore the body reflection color must be narrower-band than the highlight color. For example, if the illumination does not have any blue in it, then the body reflection cannot either, even if the object has a color that we would call blue. If the illumination does have blue in it, the body reflection can contain as much blue as it does (if the object reflects 100% of the light in the blue range) or less. Therefore the line that represents a more saturated color is chosen as the body reflection line.

It is possible that the two lines are equally saturated by the conventional saturation calculation. For a normalized color $[r, g, b]$, the saturation S is defined as

$$S = 1 - 3\text{MIN}(r, g, b)$$

As an example, a yellow light and a red object both have a saturation of 1, since the blue component is zero. In the case where the saturations are equal, the "saturation rule" is extended to consider only the remaining non-zero components.

$$S = 1 - 2\text{MIN}(r, g)$$

Since red is more saturated than yellow by this definition, the red line would be chosen as the body reflection cluster. The vector that defines the chosen line becomes $\vec{d}_b$.

Once the body reflection line has been chosen, any pixels that are more than the noise threshold away from the body reflection line are assumed to be in the highlight cluster. The noise threshold may be chosen arbitrarily for the noise-free case, but some threshold must be chosen, for the highlight component never quite goes to zero with equation 2.3. In the more realistic case where there is camera noise, as will be discussed in chapter 6, the threshold will be based on the level of that noise. Eigenvector analysis is done on all pixels classified as highlight to determine the principle vector of the cluster. This vector is $\vec{d}_s$.

Because the vector $\vec{d}_s$ is determined by the best fit to all highlight pixels, it is not completely independent of the object shape. If an object has many surface normals corresponding to pixels on one side of the highlight cluster, and only a few corresponding to the other, that may bias the direction of the vector. This thesis assumes that such a systematic effect rarely occurs, and in any case is unlikely to cause a significant deviation in cluster direction. If this is not the case, an alternate method of determining vector direction, such as fitting to the convex hull of the highlight cluster, could be used. The convex hull of the highlight cluster depends only upon a broad distribution of surface normals; this is also true of the length, width, and intersection measurements.

Once the direction of the highlight cluster has been measured, the vector $\vec{d}_s$ is used to project all highlight pixels onto the body reflection vector $\vec{d}_b$. These projections determine the relative contributions of the vectors in each pixel. Each pixel p in the histogram can then be defined as

$$\vec{p} = m_s \vec{d}_s + m_b \vec{d}_b$$

This is essentially the dichromatic equation, but the highlight cluster direction $\vec{d}_s$ may differ slightly from

the actual highlight color. The scalar histogram measurements are then defined simply as

$$l = \text{MAX}(m_s) \text{ over all } \bar{p} \quad (5.5)$$

$$b = \text{MAX}(m_b) \text{ over all } \bar{p} \quad (5.6)$$

$$i = m_b / b \text{ for that } \bar{p} \text{ with maximum } m_s \quad (5.7)$$

$$w = [\text{MAX}(m_s) - \text{MIN}(m_b)] / b \text{ over all } \bar{p} \text{ for which } m_s > T \quad (5.8)$$

where the threshold is given by T .

The lookup-table approach assumes that there is a fundamental relationship between histogram dimensions of length, width and intersection, and the scene parameters of roughness, phase angle and illumination intensity. This relationship, described in chapter 3, relies upon physical models of body reflectance and surface reflectance, and how they interact to form color pixels. The model does not make assumptions about specific object shape, although it does assume a broad, continuous distribution of surface normals. At a minimum, the method requires that surface normals corresponding to key points in the histogram—**a**, **b**, **c1**, **c2**, and **d** from figure 3.1—are visible to the camera, since they are necessary for the calculations of length, width and intersection. The lookup table approach also expects that the relationship between histogram measurements and scene parameters is one-to-one, assuming that very different scene parameters cannot give rise to similar histogram shapes. Figure 5.11 demonstrates that the relationship is indeed unique.

5.7 Calculating Roughness, Phase Angle, and Illumination Intensity

The surface in figure 5.11 shows lines of equal roughness within the L-W-I space. Once length, width and intersection measurements have been obtained from a histogram, the problem is to determine which line they fall on. The lines shown in the figure come from the lookup table values and are fairly close together. However it is unlikely that the histogram from any given image will have measurements that fall exactly on these lines. Therefore some sort of interpolation is needed.

This thesis uses polynomial approximations to the surface in L-W-I space. It is assumed that the roughness σ can be approximated as an n th degree polynomial function of the length, width, and intersection measurements of the histogram.

$$\sigma \approx f_n(l, w, i) \quad (5.9)$$

$$\approx A + Bl + Cw + Di + El^2 + Fw^2 + Gi^2 + Hlw + Ili + Jwi + \dots$$

Function	R squared
$\sigma = f_n(l, w, i)$	0.877
$\theta_p = g_n(l, w, i)$	0.939
$B_s = h_n(l, w, i)$	0.712

Figure 5.14: Goodness-of-fit for polynomial functions

The lookup table provides the means for calculating the coefficients of the polynomial. It provides almost 500 sets of histogram measurements (see figure 5.12) and the associated roughness values. Least squares estimation is used to calculate the best fit n th degree polynomial to the data. A fourth degree polynomial was used in the experiments described here.

Similarly, the phase angle and illumination intensity are also approximated as polynomial functions of the histogram length, width, and intersection.

$$\theta_p \approx g_n(l, w, i) \quad (5.10)$$

$$B_s \approx h_n(l, w, i) \quad (5.11)$$

Least squares is used to calculate the coefficients for these polynomials also. Since the length measurement is a distance in RGB space, rather than a percentage of the length of the body reflection cluster, the length measurement was first divided by 255 so that it would fall in approximately the same range as the width and intersection values (roughly between 0 and 1). The fit of the least squares calculation to the data is shown for each function in the table in figure 5.14.

Generating the lookup table is obviously very time consuming (about 8 hours on a SPARC II) since it involves calculating almost 500 graphics simulations. However this and the coefficient calculation only need to be done once. Given the histogram of an object with unknown parameters, the system can automatically separate it into the two clusters and measure their dimensions as described in section 5.6. The polynomial calculation can then be applied to quickly estimate the roughness, phase angle, and illumination intensity. This algorithm is diagrammed in figure 5.15. The run-time portion is very quick, taking less than 3 seconds on a histogram containing about 3600 pixels. If the histogram has already been split into clusters in the process of segmenting the image [Klinker, 1988], the time to calculate the scene parameters is less than 1 second.

To test the polynomial approximations, one hundred test images were simulated and then analyzed by this method. These test images were noise-free, generated with the same idealized camera model used to

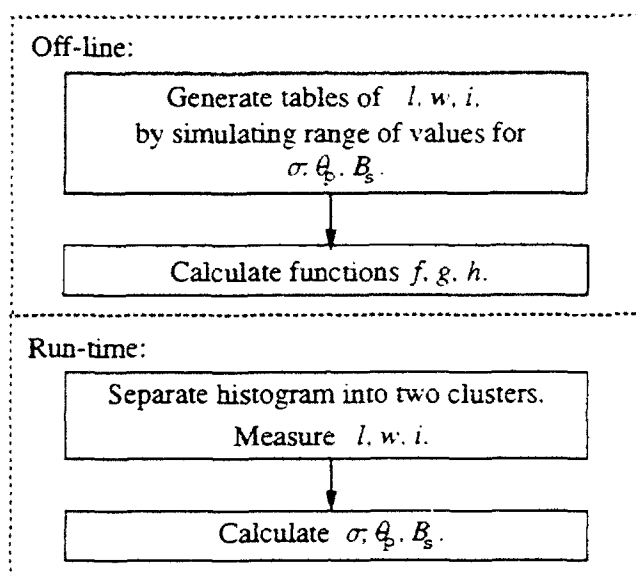


Figure 5.15: Algorithm for calculating roughness, phase angle, and illumination intensity

generate the lookup table. The surface roughness, phase angle, and illumination intensity values used in the test images were chosen by a pseudo-random number generator. The values were constrained to lie within the ranges used in the lookup table (see figure 5.12).

The calculated values of σ , θ_p , and B_s were compared with the original values used to generate the image. In almost all cases the calculated values were very close to the original ones. However, for 2 of the 100 test cases, the values were very obviously wrong. For example, a negative value of roughness or illumination intensity is clearly unreasonable. Fortunately bad values are easily detected by checking to see if recovered values are within the allowable range. Recovered values outside that range indicate that a different method should be used to recover the scene parameters.

For example, a lower degree polynomial approximation could be used for these cases. Although the fourth degree polynomial is better at approximating the function over its whole range, it may deviate somewhat at the extremes of the function. In fact the 2 problem cases occur when the roughness is very low (between 1 and 2 degrees) and the phase angle is very large (greater than 75 degrees). The problem disappears if a third degree polynomial is used instead, although the overall error on all cases is slightly higher.

The table in figure 5.16 shows the results for the remaining 98 cases where the fourth degree polynomial produced believable results. It shows the average error in recovering the parameters, and also reiterates the

Parameter	Average Error	Resolution of Table
Roughness	1.20°	2°
Phase Angle	4.40°	10°
Intensity	8.18 %	10 %
Cases Considered		98/100

Figure 5.16: Results on noise-free data

step sizes used in the table. The errors are lower than the table resolution, showing the interpolation method is fairly effective.

The results for calculating roughness and phase angle are very good. They show that these non-color parameters may be calculated with reasonably high accuracy just by considering the shape of the color histogram. The error in calculating illumination intensity result is a bit higher, although it still provides a useful estimate. This error is not too surprising, since the R-squared fit of the function to calculate illumination intensity is somewhat worse than for the other two functions (see figure 5.14).

5.8 Calculating Illumination Chromaticity

As has been discussed in section 3.2.1 the highlight cluster may be somewhat skewed from the direction of the highlight color. The skew is particularly pronounced for large phase angles and when the surface is rough. These two factors determine how much the body reflection changes over the area of the highlight. Figure 5.17 shows how the skew varies as a function of roughness and phase angle. The height of the graph is the angle (in degrees) between a vector fitted to the highlight cluster and the vector defined by the illumination color.

Therefore, if the analyzing program knows—or can calculate—the surface roughness and the imaging geometry, it can in turn calculate the amount of highlight skewing. Once the skew is known, its effect can be subtracted from the direction of the highlight cluster to give the true color of the illumination.

The previous section showed how to estimate the roughness and phase angle from the color histogram. These estimates will now be used to estimate the skewing. A similar lookup table approach is used. When the simulations are performed to calculate length, width, and intersection for the range of scene parameters given in figure 5.12, the skewing of the highlight is also calculated. In the graphics simulation the correct

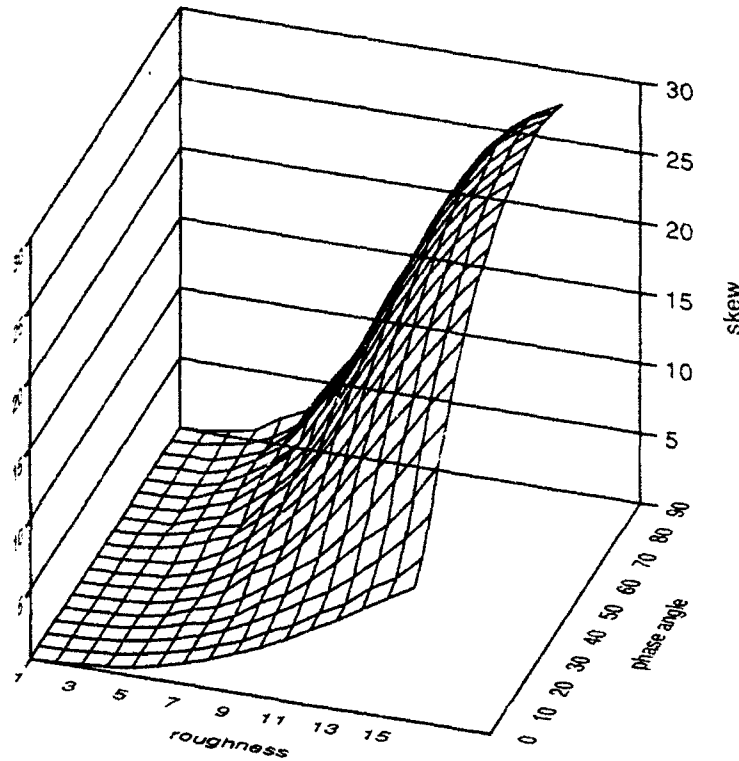


Figure 5.17: Variation in cluster skew with changes in roughness and phase angle

illumination color is obviously known in advance, so the angle between this color and the vector fitted to the highlight cluster is calculated and stored in the lookup table along with the values for length, width, and intersection. Then a polynomial function is used to calculate the skew angle as a function of roughness, phase angle, and intensity:

$$Skew \approx A_n(\sigma, \theta_p, B_s) \quad (5.12)$$

The coefficients of the n th degree polynomial function are calculated using a least squares fit to the data in the lookup table. A third degree polynomial was tested and gave an R-squared fit of 0.992.

Once the skew has been calculated, the normalized highlight color $\vec{c}_s$ is calculated from the measured direction $\vec{d}_s$ by the constraints

$$\vec{c}_s \cdot \vec{d}_s = \cos(Skew) \quad (5.13)$$

$$\vec{c}_s \cdot \vec{d}_b = \cos(Skew + \arccos(\vec{d}_b \cdot \vec{d}_s)) \quad (5.14)$$

$$\vec{c}_s \cdot (\vec{d}_s \times \vec{d}_b) = 0 \quad (5.15)$$

Equation 5.13 simply says that the angle between the highlight color $\vec{c}_s$ and the measured cluster direction $\vec{d}_s$ is *Skew*. Equation 5.14 indicates that the angle between the highlight color $\vec{c}_s$ and the body reflection

Average Error	1.73°
Average Skew	8.63°
Minimum Value	0.01°
Maximum Value	27.5°
Number of Skews > 1°	69
Cases Considered	98/100

Figure 5.18: Results for calculating skew

direction $\vec{d}_b$ is *Skew* plus the angle between the cluster direction and the body reflection direction (the skewed highlight cluster direction must fall **between** the true highlight color and the body reflection color). Equation 5.15 states that all three vectors must fall in a plane.

Obviously, if the polynomial functions described in section 5.7 produce unreasonable estimates of the roughness, phase angle, or intensity, there is little point in plugging them into the equation for calculating *Skew*. Therefore, for the 2 test cases that were rejected, the program did not attempt to calculate the illumination color. For the remaining 98 test cases, the skew angle was used to calculate the illumination color. The results are shown in the table in figure 5.18. The error in estimating skew is the difference—in degrees—between the correct skew angle and the calculated skew angle. The correct skew angle is easily calculated from the illumination color that was used to generate the simulation. The table shows the average error over the 98 cases considered. It also shows the minimum, maximum and average of the actual skew values. For 69 of the 100 test images, the scene parameters were such that the skew exceeded one degree.

As an example, one of the test simulations is shown in color figure A.4. This is a red cylinder under white light, so $\vec{c}_s = [0.58, 0.58, 0.58]$. The surface roughness has a 12.06° standard deviation of the facet angles; the phase angle separating the camera and light source with respect to the object is 63.30°; and the illumination intensity is 90% of maximum.

The histogram associated with this image is shown in figure 5.19. The graph shows the projection of the color histogram into the red-green plane. The program automatically divided the histogram into body reflection and surface reflection clusters. The vector fitted to the cluster is $\vec{d}_s = [0.81, 0.45, 0.37]$. The length of the highlight cluster was measured as 74.1; the width of the highlight cluster was measured as 0.47 (the highlight cluster extends over slightly less than half the body cluster); the intersection of the two clusters was measured as 0.51 (the brightest point in the highlight cluster was projected onto the body cluster just above the halfway point). These measurements were all obtained using the vector $\vec{d}_s$ to calculate the amount

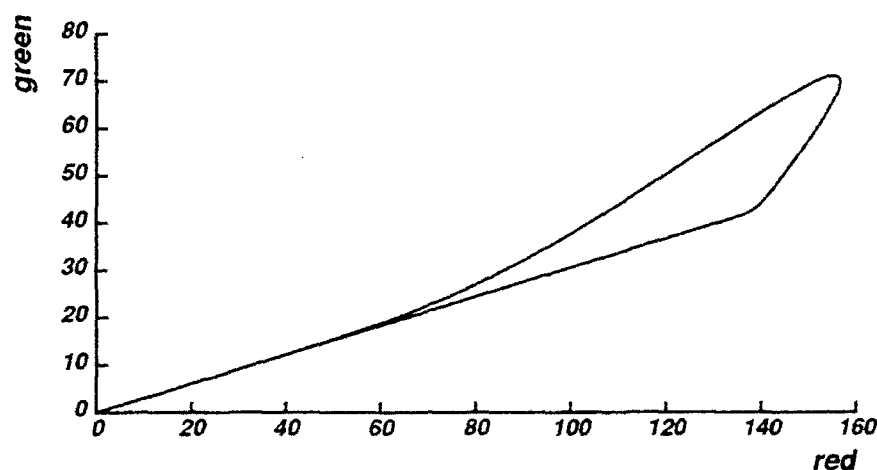


Figure 5.19: Histogram of simulated image

of body reflection and surface reflection.

The direction fitted to the highlight cluster is significantly skewed away from the direction of the actual illumination color. It is significantly reddened compared to the actual illumination color, and so would be a poor estimate of the illumination color. It would also yield an inaccurate estimate of the object color when the influence of the illumination color is divided out of the body reflection color.

Applying equations 5.9, 5.10 and 5.11 to the length, width and intersection measurements, the program estimated the roughness value as 11.99° , the phase angle as 67.05° , and the illumination intensity as 89% of maximum. Applying equation 5.12 to these values for roughness, phase angle, and illumination intensity, the skew between the highlight cluster direction and the actual illumination color was estimated to be 18.75° . Solving the constraint equations 5.13 through 5.15 produced an estimate of the light source chromaticity as $\bar{c}_s = [0.59, 0.57, 0.57]$. This is very close to the original white color.

The results for this example are summarized in the table in figure 5.20. The recovered parameters were also used to generate a new simulation picture, shown in color figure A.5. The difference between this image and the original image in figure A.4 is very subtle. The full algorithm, including the technique for estimating the illumination color, is diagrammed in figure 5.21.

The object reflectance used to simulate the cylinder $\bar{c}_o$ is $[0.94, 0.29, 0.16]$. If the highlight direction vector $\bar{d}_s$ is assumed to be the illumination color, then removing its influence from the body color yields an estimate of $[0.83, 0.47, 0.31]$ for the object reflectance. However, the corrected estimate of the illumination color obtained by calculating the highlight skew yields an estimate of $[0.94, 0.30, 0.17]$ for the object's

reflectance, much closer to the actual reflectance used to generate the image. Thus, accounting for the skew of the highlight leads to improved estimates of object reflectance for color constancy.

Calculating the illumination color also allows the system to generate intrinsic images of body reflection and surface reflection, as shown in [Klinker, 1988]. The intrinsic body reflection image is calculated by projecting all pixels along the direction of the highlight onto the body reflection vector. Similarly, the intrinsic surface reflection comes from projecting all pixels along the direction of the body color onto the highlight direction vector. Given that each pixel in the original image has been separated

$$\vec{p} = m_s \vec{d}_s + m_b \vec{d}_b$$

Then each pixel in the resulting intrinsic body reflection image is just

$$p_b = m_b \vec{d}_b$$

while the intrinsic surface reflection pixels are calculated

$$p_s = m_s \vec{d}_s$$

However if the highlight color has been calculated incorrectly by failing to account for the skew of the highlight cluster towards the body reflection color, those pixels in the highlight will be projected too far down the body reflection vector, making them appear too dark. This situation is shown in color figure A.6. The intrinsic images were calculated using the highlight direction $[[0.81, 0.45, 0.37]]$, as measured in the histogram before correcting for the skew. The cylinder on the left is the intrinsic body reflection while the cylinder on the right is the intrinsic surface reflection. Because this highlight cluster is skewed away from the white color of the original illumination, the intrinsic body reflection image has a dark area where the highlight was. The intrinsic surface reflection image is much redder than the actual illumination color.

However, when the skew of the highlight is accounted for, the corrected illumination estimate of $[0.59, 0.57, 0.57]$ can be used to calculate the intrinsic images. Now the pixels in the highlight will be projected correctly onto the body reflection vector. The result is shown in color figure A.7. The intrinsic body reflection image on the left appear correctly shaded, and the intrinsic surface reflection image on the right correctly shows that the highlight color is white.

This chapter has shown that the method developed works very well on perfect images generated without camera limitations. Any practical method will have to work on real images and must deal with noise and

Simulated Image	$\vec{c}_s = [0.58, 0.58, 0.58]$	$\sigma = 12.06^\circ$	$\theta_p = 63.30^\circ$	$B_s = 90\%$
Histogram Measurements	$\vec{d}_s = [0.81, 0.45, 0.37]$	$l = 74.1$	$w = 0.47$	$i = 0.51$
Recovered Parameters	$\vec{c}_s = [0.59, 0.57, 0.57]$	$\sigma = 11.99^\circ$	$\theta_p = 67.05^\circ$	$B_s = 89\%$

Figure 5.20: Results for test image

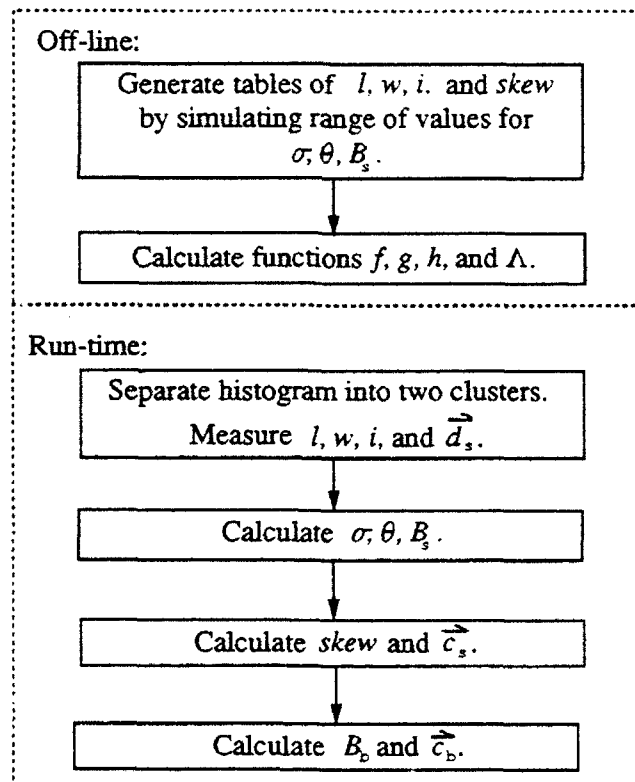


Figure 5.21: Algorithm for recovering scene parameters

other camera problems. The next chapter will show how the method developed here may be extended for a camera model that incorporates some limitations found in real cameras.

Chapter 6

Extending the Analysis to a Realistic Sensor Model

Chapter 5 described how to estimate scene parameters from a color histogram if an ideal sensor is used. Specifically, it assumed that the camera used to view objects is noise-free and linear over the range of all possible intensity values. The possibility of clipping at some maximum camera value was not considered in chapter 5.

However these are not very realistic assumptions. Some forms of noise exist in all real camera systems. And, while linearity of response over a limited range is common in many cameras, all real cameras and digitizers have some maximum response. Clipping at the maximum value is particularly common in pictures with highlights since they are so much brighter than other parts of the scene.

Since this thesis aims to analyze pictures with highlights, and since all real imaging systems have noise, it will be necessary to deal with these issues in order to make the algorithm work on real images. Section 6.1 will describe how the algorithm has been adapted to deal with noise, while section 6.2 will describe a method for dealing with the problem of clipping.

6.1 Modeling Camera Noise

Real cameras present many limitations for taking accurate images. CCD cameras, which are used by most vision researchers, present many problems such as dark current, shot noise, amplifier noise, line jitter, etc. [Healey and Kondepudy, 1991]. Some of these problems, such as shot noise, are an inherent property of CCD cameras and cannot be removed even with high-quality manufacturing.

This thesis models camera noise as having a Gaussian distribution with a mean of zero. The standard deviation of camera noise σ_c was modeled as 1.275. The camera is assumed to measure eight bits per pixel, giving a range from 0 to 255. Thus the ratio of maximum pixel value to standard deviation of noise is then 200.

Since the method described in this thesis estimates scene parameters by analyzing color histograms, it must consider the effect of noise on the histogram's appearance. Under the Gaussian noise model, each pixel may be viewed as occupying a probability cloud in color space. This cloud represents the pixel's ideal color value plus some uncertainty due to noise. The size of the cloud depends upon the standard deviation of the noise. If the population of pixels is large, the cloud will be filled according to the noise distribution. For a Gaussian distribution, the modal pixel will coincide with the ideal pixel value, with two-thirds of the pixels falling within one standard deviation of the ideal value.

One of the effects of noise on the histogram is that it will make the body reflection cluster thicker. A uniformly colored object imaged with an ideal camera would have a body reflection cluster that is a line only one pixel thick. Noise added to ideal pixels will correspond to some pixels that are displaced away from the ideal line. This will result in a thicker body reflection cluster. However, even in the ideal case the algorithm needed a threshold for distinguishing highlight pixels from pixels on the body reflection cluster, although this threshold could be arbitrary (see section 5.6). With an explicit noise model, the threshold should be based on the standard deviation of the camera noise. The threshold T was chosen to be $4\sigma_c$, so over 99% of the noisy body reflection pixels will fall within the body reflection cluster.

Noise added to body reflection pixels will also push some pixels up or down, along the direction of increasing amounts of body reflection. Noise added to those pixels which are at the tip of the body reflection cluster (point b in figure 3.1) will result in a cluster that is slightly longer. This will affect the measurement of the length of the body reflection cluster b and the subsequent calculation of the object's albedo B_o (see

section 5.2). Fortunately the effect is very small compared to the total length of the body reflection cluster, so it will be ignored.

Noise added to pixels in the highlight cluster will also displace those pixels, causing the highlight cluster to be slightly wider and longer than for the noise-free case. Unlike noise added to the body reflection cluster, this effect must be considered carefully, since the estimation method uses small differences in the width of the cluster to predict differing values of roughness. Therefore in dealing with noise in the highlight pixels a two pronged attack is used: the noise is explicitly accounted for when generating the lookup tables; and an attempt is made to remove some noise when analyzing each histogram.

When the lookup tables were generated for chapter 5, a computer program simulated noise-free images and measured the resulting histograms. Now, for the noisy camera model, images are simulated again, this time with pseudo-random Gaussian noise added. The resulting histograms are analyzed, and new values of length, width and intersection are recorded in the lookup table. These measurements will reflect the slight increases in the length and width of the two histogram clusters.

Furthermore, in measuring the histograms local smoothing is used. Each pixel from the simulated image is averaged with its four nearest neighbors before being plotted in the histogram. For pixels falling in the highlight or at its edge, this will have the effect of smoothing out the highlight somewhat, but will also tend to average out the noise. The effect of highlight smoothing on the length, width and intersection measurements will also be recorded in the lookup table.

The same techniques as described in section 5.7 are used to fit polynomials to the measurements in the new table. In analyzing noisy histograms, any pixels that fall outside the plane defined by the vectors fitted to the clusters, $\vec{d}_s$ and $\vec{d}_b$, are projected into that plane. Furthermore the body reflection vector is fitted slightly differently since it will now contain noise. The pixel furthest from the line drawn to the brightest pixel is still used, but it is an initial estimate. All pixels that are no further than the threshold T from that line are clumped together and the best fit line to that cluster becomes the new body reflection estimate. The process is repeated until the new vector estimate is within one tenth of a degree of the previous vector, or ten iterations have taken place.

The table in figure 6.1 presents the results when the method is tested on 100 simulated images with pseudo-random Gaussian noise added. The algorithm is the same as the one presented in figure 5.21. Only the lookup tables have been changed, to reflect the noise model of the camera. Comparing the noise-free

Parameter	Average Error	Resolution of Table
Roughness	1.08°	2°
Phase Angle	4.56°	10°
Intensity	6.85 %	10 %
Skew	1.39°	-
Cases Considered	99/100	

Figure 6.1: Results on noisy data ($\sigma_c = 1.275$)

Parameter	Average Error	Resolution of Table
Roughness	0.97°	2°
Phase Angle	7.07°	10°
Intensity	7.49 %	10 %
Skew	1.29°	-
Cases Considered	98/100	

Figure 6.2: Results on noisier data ($\sigma_c = 2.55$)

results in figures 5.16 and 5.18 with the results here, we see that the algorithm actually did slightly better with noisy data, except when calculating phase angle. This is a surprising result, but it is caused by the averaging that is done in on the noisy data, which has the effect of smoothing the histogram slightly.

A second experiment was performed to see how the system handles a larger amount of noise. Accordingly, the tables of length, width, and intersection were recomputed for a camera with twice the noise level ($\sigma_c = 2.55$). The results are shown in figure 6.2. They show that the estimates of phase angle are significantly worse than for the low noise case, although the error is still less than the resolution of the table. The errors in estimating roughness, illumination intensity, and skew angle are affected only slightly. This shows that the method does not suffer significantly if the noise of the camera is larger, *provided that noise has been modeled* when the lookup tables are calculated.

If the noise is not properly modeled when calculating the lookup tables, the method does not work well at all. For example, if the larger noise value ($\sigma_c = 2.55$) is used to simulate the test images, but the smaller noise value ($\sigma_c = 1.275$) is used to generate the lookup tables and compute the polynomial functions, the method produces implausible answers 26% of the time, and the errors in computing the scene parameters are much higher.

6.2 Modeling Camera Clipping

Like noise, clipping is an inherent property of real cameras since all real cameras have some maximum value that they can measure. Many real images with highlights suffer from clipping since highlights are often very bright. When taking real pictures, it is usually possible to adjust the aperture, light level, or exposure time to a certain extent. However, highlights on shiny objects are so much brighter than surrounding non-highlight areas, that if the camera is adjusted to measure highlights properly, the non-highlight areas will be too dark to rely upon their values.

Therefore any algorithm that hopes to work on real pictures with highlights should take clipping into account. The effect of clipping on the color histogram is easy to model. Any pixels that would have exceeded the maximum value (255 for an 8 bit camera) will be held to that value. As a result, the highlight cluster will not be as long as it otherwise would have been. Clipping usually occurs when the surface roughness has a standard deviation of facet angles of 3° or less.

In analyzing color histograms, clipping is easy to detect. If a pixel has the maximum value in any band, it was almost certainly clipped. However, dealing with the clipping presents a greater problem since the effect upon the length may be dramatic. Fortunately the width measurement is not affected at all, as long as there are some pixels at the edge of the highlight that do not saturate the camera. The intersection measurement can be affected, since it relies upon the projection of the brightest highlight pixel. In the case of clipping, the brightest *non-saturated* pixel is used to calculate the intersection point. Fortunately, the difference is usually very small. This is because the highlight cluster is usually very narrow at the point where it is clipped.

The length measurement is an important parameter in calculating the scene parameters, along with the width and intersection measurements. In order to calculate the three unknown values of surface roughness, phase angle and illumination intensity, three known histogram measurements are needed. If the length measurement cannot be used, it will have to be replaced with some other measurement.

Therefore the length measurement is replaced with another measurement of the cluster shape. The width measurement and the length measurement are a way of characterizing the shape of the highlight cluster. The algorithm needs to capture this shape since it is related to the relative distributions of body and surface reflection, which in turn are determined by the scene parameters we want to estimate.

When the length measurement is untrustworthy due to clipping, the method will then use a second width

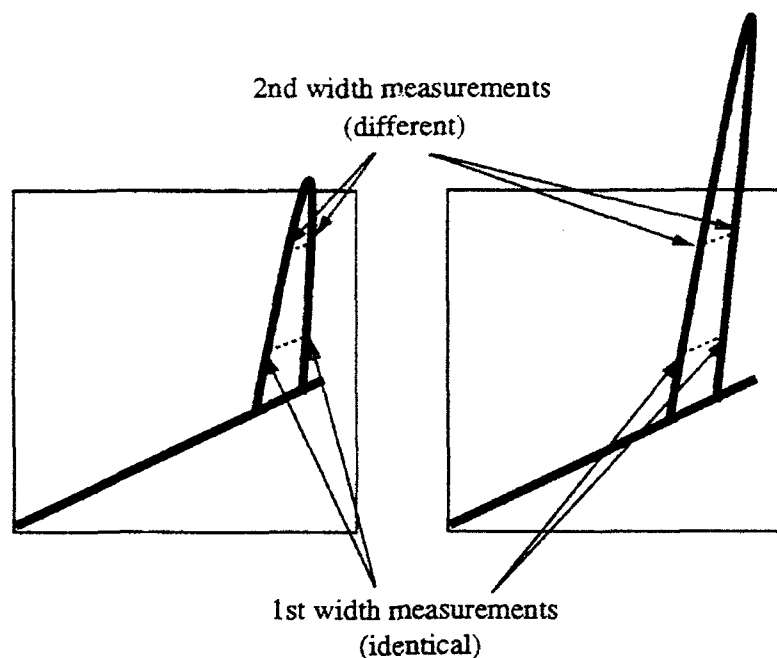


Figure 6.3: Purpose of second width measurement

measurement. The first width measurement tells how wide the highlight cluster is at the threshold distance T from the body reflection cluster. The second width measurement was chosen to show how wide the highlight cluster is at a distance of $3T$ from the body reflection cluster. This measurement will allow the system to distinguish between cases where the highlight cluster would have extended only slightly beyond the color cube boundary, and cases where the highlight cluster would have been much longer. These two cases are sketched in figure 6.3.

When the lookup tables are calculated, the length measurement is recorded, and also an indication as to whether or not it is clipped. At the same time, the second width measurement is recorded, as well as the first width and intersection measurements previously described. After the lookup table is generated *two* sets of polynomial functions are fitted: one for analyzing histograms that are not clipped, and one set for histograms that are. If the second width measurement is referred to as w_2 then the scene parameters may be calculated

Parameter	Average Error	Resolution of Table
Roughness	1.00°	2°
Phase Angle	6.07°	10°
Intensity	6.76 %	10 %
Skew	1.69°	-
Cases Considered	96/100	

Figure 6.4: Results on noisy data when clipping is present

Parameter	Noise-free	Noisy	Noisy and Clipped
Roughness	1.20°	1.08°	1.00°
Phase Angle	4.40°	4.56°	6.07°
Intensity	8.18 %	6.85 %	6.76 %
Skew	1.73°	1.39°	1.69°
Cases Considered	98	99	96

Figure 6.5: Errors for three camera models

by

$$\left. \begin{aligned} \sigma &= f_1(l, w, i) \\ \theta_p &= g_1(l, w, i) \\ B_s &= h_1(l, w, i) \end{aligned} \right\} \quad \text{if the histogram is not clipped} \quad (6.1)$$

$$\left. \begin{aligned} \sigma &= f_2(w_2, w, i) \\ \theta_p &= g_2(w_2, w, i) \\ B_s &= h_2(w_2, w, i) \end{aligned} \right\} \quad \text{if the histogram is clipped} \quad (6.2)$$

Since there are fewer data points in the lookup table for the clipped case, a third degree polynomial is fitted for functions f_2 , g_2 , and h_2 .

The results obtained by applying this technique to 100 simulated test images with Gaussian noise added ($\sigma_c = 1.275$) are shown in the table in figure 6.4. The results here are similar to those obtained without clipping, although there are more problem cases that had to be eliminated from consideration. Results from the noise-free, noisy, and noisy with clipping algorithms are all shown in figure 6.5.

Chapter 7

Application of the Algorithm to Real Images

Chapter 5 presented a method for analyzing color histograms assuming an ideal imaging system. Chapter 6 showed how that method may be modified in a systematic way for a more realistic camera system. This chapter will present some results from applying the algorithm developed in chapter 6 to real images.

The real images in this thesis were taken in the Calibrated Imaging Laboratory (CIL) at Carnegie Mellon University. The CIL provides a controlled setting for taking high-quality images. It also allows for a careful analysis of camera issues that may cause problems for vision algorithms. The equipment and capabilities are described in [Willson and Shafer, 1992].

7.1 Additional Problems with Real Cameras

Although chapter 6 modeled some fundamental limitations of real cameras, including noise and clipping, other problems that are commonly encountered in imaging were not modeled. These problems include nonlinearity, color imbalance, chromatic aberration, and blooming. Some problems such as linearity and color balance are easily fixed in software or with simple hardware. Others are more difficult to handle.

The degree of each of these problems is very specific to the particular camera system used. This makes them very difficult to model in a general way. Some systems may have been designed to reduce or even eliminate a specific problem, such as nonlinearity.

7.1.1 Nonlinearity and Color Imbalance

The methods for analyzing histograms developed in this thesis require that the camera have a linear response. This is because they measure specific distances in the color space and assume that these distances are proportional to the amounts of body reflection and surface reflection observed on the object.

For example, the intersection measurement described in section 3.2.4 is assumed to record an estimate of the amount of body reflection at the brightest point in the highlight. The ratio between the body reflection at that point and the maximum amount seen anywhere on the object is proportional to the cosine of the angle between the light source and the perfect specular direction. However if the camera is not linear in its response, that proportionality will no longer hold.

Fortunately, linear CCD cameras are commonly available. In fact CCD cameras are inherently linear since they produce a voltage that is proportional to the number of photons; the problem is that some models are designed to drive displays and have added circuitry that makes them nonlinear. In this case, the nonlinearity may be measured using a standard reference chart and corrected with a lookup table [Klinker, 1988], [Novak *et al.*, 1990]. The camera used to take the real images in this chapter already has a linear response, so no correction is necessary.

Another common problem with CCD cameras is color imbalance. CCDs are typically much less sensitive to blue wavelengths than they are to green or red wavelengths. In addition, incandescent lights are usually chosen for taking pictures in the CIL. This is for three main reasons: incandescent lighting is somewhat closer to the point source model than most fluorescent bulbs; the intensity of an incandescent light is easily controlled with a variable voltage supply; and fluorescent lights have a very spiky spectral power distribution that can cause interference patterns in the lens.

Incandescent lights are also strong in the long wavelengths and weak in the short wavelengths. Therefore when incandescent lights are used with CCD cameras, the blue response is extremely poor. If color pictures are taken without some form of color correction, blue objects will be extremely dark, and white objects will appear yellowish.

A simple solution would be to scale the blue band of all color images by an appropriate constant. The problem with this approach is that while scaling up the blue signal, it would also scale up the noise. Therefore it would be better to increase the signal in some other way, such that the values in the blue band are as

reliable as those in the red and green bands.

In the CIL, two different approaches have been used to increase the blue response. In both cases, color pictures are taken with a gray-scale camera and a filter wheel to produce the red, green, and blue bands. This allows adjustments to be made individually for each band. In the case of a camera with a fixed exposure time, the aperture is changed between images of the bands, so that the aperture can be much wider to admit more light for the blue band.

A white test card is used to select the appropriate aperture settings for each band. The aperture on the CIL cameras is under computer control with a motorized lens that has 2700 steps for aperture control. This aperture control method has the disadvantage that the depth of field changes slightly between bands, but this effect is not noticed in scenes with a small range of depths. Figures A.1 and A.2 shown in chapter 3 were taken with this technique.

A different technique is used for cameras that have a variable exposure time. Again, a white test card is used, although in this case the test images are used to select the appropriate exposure times. The exposure time can be as short as 0.0 seconds and as long as 6,553.6 seconds, with a resolution of 0.1 second. In the CIL, the exposure time for the blue band is typically ten times longer than the red or green exposures. The images shown in this chapter were taken with this latter technique.

7.1.2 Chromatic Aberration

Another common problem in real images is chromatic aberration. This is an inherent property of lenses, because the index of refraction of optical components varies as a function of wavelength. Chromatic aberration can be partially compensated for during manufacturing by using pairs of lens elements with offsetting dispersion functions. Unfortunately this type of compensation is usually only calculated for two wavelengths (red and blue) and for a small number of points in the image field.

Chromatic aberration has two main effects: the different color bands will have slightly different points of best focus; and the different bands will have slightly different magnifications. A third effect, which is observed in real lenses, is a lateral displacement of the images between bands, caused by slight misalignment of the lens components. These effects will cause noticeable shifts in the color of affected pixels. Experience in the CIL has shown that even high quality lenses suffer from chromatic aberration that is readily apparent

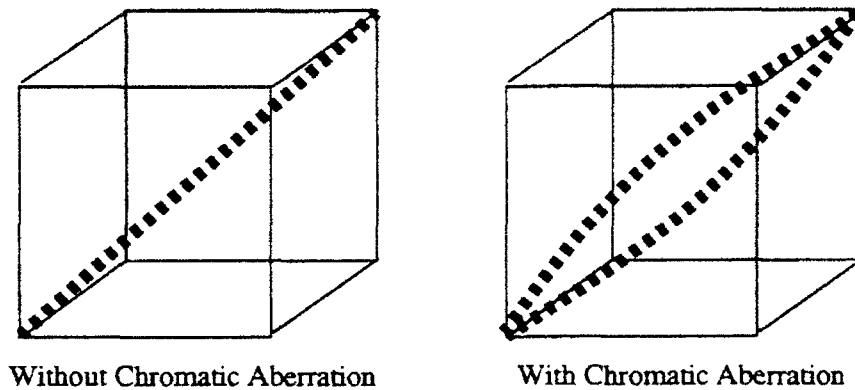


Figure 7.1: Sketch of the effects of chromatic aberration on a simple histogram, as described by Willson and Shafer, 1991

in histogram analysis. For example, the histogram of a black grid on a white background would be expected to show only white pixels, black pixels, and gray pixels where an edge in the image fell somewhere in the middle of a pixel. When chromatic aberration is present, some of the pixels at the edges will be shifted towards red whereas others will be shifted towards blue or green due to the different magnification and focus between the three bands.

This effect is not noticeable to the human eye at normal magnification, but is very obvious when the image is magnified. It is also obvious when looking at the histogram. The histogram of a black and white grid would be expected to fall on a straight line along the neutral axis of the color cube, as illustrated on the left of figure 7.1. Chromatic aberration causes significant deviations from the neutral line, as shown in the histogram on the right.

Chromatic aberration is most pronounced where there are sharp edges. And that causes problem not only for pictures of black lines on a white background, but also for pictures of highlights. For smoother surfaces, the transition between highlight and non-highlight areas will be more sudden, and thus the chromatic aberration will be more pronounced. As a result, the highlight cluster of histograms for smooth objects will be more widely scattered than would be expected. The degree of this scattering depends not only upon the smoothness of the object, but also upon the characteristics of the lens, the color of the object, and where in the image the highlight occurs. (Chromatic aberration is most pronounced in the periphery of the image.) This makes it very difficult to model in a general way.

This thesis takes a two-pronged attack upon the problem of chromatic aberration. The first step is to eliminate as much of the chromatic aberration as possible at the time the image is taken. This is done by

active lens compensation. Since the CIL allows fine control of the lens position, and since the color bands are imaged separately, the lens is refocused, rezoomed, and repositioned between the red, green and blue color bands. This approach typically reduces the chromatic aberration by an order of magnitude [Novak *et al.*, 1990], [Willson and Shafer, 1991].

Unfortunately, the method does not completely eliminate all chromatic aberration in images with highlights. Also it is time-consuming to calibrate and is not generally applicable to camera setups that do not have a finely controlled lens. Therefore the second approach is to take some chromatic aberration into account when measuring the dimensions of the color histogram.

Color figure A.8 shows a histogram for a shiny red ball. The original image is shown in the upper right hand corner of the image. The top of the highlight cluster is slightly fanned out, making it slightly wider than the bottom of the highlight cluster. Color figure A.9 shows the histogram associated with a very shiny red pail. Here the fan-out is even more pronounced. (The bottom few rows of the original image where blooming has occurred were manually eliminated from consideration.)

This fanning of the highlight cluster is particularly troublesome because it runs completely counter to the model of histogram behavior. The model says that there will be more variation of body reflection amounts and the base of the highlight cluster than at the upper end. A highlight cluster that is wider at the top implies that there are very bright highlight pixels occurring at several different reflection angles, but that for some reason pixels with lesser amounts of highlight occur at only a few reflection angles. If the histogram were to be believed, then the amount of highlight does not fall off as the off-specular angle is increased. This would mean a rather unusual distribution of facet angle. A more plausible explanation of the fan-out of the highlight cluster is that the chromatic aberration causes large hue shifts at points where the highlight is increasing very rapidly.

Since the chromatic aberration makes the highlight cluster wider in regions where the highlight increases rapidly, this would imply that the width measurement is not very trustworthy at the top end of the highlight cluster. A program analyzing the histogram in figure A.9 will measure a very wide highlight cluster, even though it corresponds to a very shiny object. This is because the original algorithm examine the amount of body reflection present in *all* highlight pixels to compute the width measurement. A pixel is classified as a highlight is the amount of surface reflection exceeds the threshold $4\sigma_c$. Although the model predicts that the widest part of the highlight cluster is at the base, this is not explicitly tested by the program.

Although the highlight cluster is plagued by chromatic aberration at the top end, the base of the cluster is still well behaved. Now that chromatic aberration is known to be a problem, the algorithm measures width over all highlight pixels *near* the threshold point, instead of over all pixels that *exceed* the threshold. Equation 5.8 that gives the width measurement in terms of the computed amounts of body reflection m_b and surface reflection m_s is modified to

$$w = [MAX(m_b) - MIN(m_b)]/b \text{ over all } \vec{p} \text{ for which } T < m_s < T + \sigma_c$$

Similarly, the second width measurement, used when the highlight cluster is clipped, becomes

$$w_2 = [MAX(m_b) - MIN(m_b)]/b \text{ over all } \vec{p} \text{ for which } 3T < m_s < 3T + \sigma_c$$

The process for measuring length and intersection remain unchanged.

7.1.3 Blooming

Color figure A.10 shows a 2x magnification of the lower section of the pail image in figure A.9. This picture clearly shows that the brightest section of the highlight has "bloomed" into neighboring pixels. The blooming is most pronounced in the red band of the image, creating a red halo around the highlight. This effect is not chromatic aberration since it is several pixels wide.

Blooming occurs when points on the CCD chip are so saturated by incoming photons that the change in one well on the chip spills over into neighboring charge wells. Some cameras are designed to be fairly resistant to blooming, an advantage when taking pictures of highlights. The camera used to take the pictures in this chapter is very prone to blooming. An interesting thing to note is that it blooms predominantly in the horizontal direction. The picture in figure A.10 is only a mild case of blooming. In some cases the highlight will smear all the way across the image.

Bloomed pixels that saturate the camera do not present a problem for the method developed in this thesis. Saturated pixels are easy to detect; any pixel with a value equal to the camera maximum is considered untrustworthy and disregarded. The fact that the object is too bright will be noted and the clipped method is used to calculate the scene parameters. However sometimes blooming will increase the values of nearby pixels *without* causing them to saturate. This means the pixel values may be increased by some arbitrary amount, to some new arbitrary value that is less than the camera maximum. Since the blooming may effect

very large areas of the image, it is difficult to handle in an automatic way. In the histogram for figure A.9, all pixels that occurred in the same rows as bloomed pixels were eliminated from consideration.

Blooming can be reduced a great deal by manufacturing a camera chip with charge wells further apart. A camera of this type would make it possible to consider all regions of the image when analyzing the histogram, instead of arbitrarily throwing away portions adjacent to saturated pixels. Saturated pixels do not present a problem for the method developed in this thesis, *provided* they do not spill charge into neighboring pixels.

This section has described some modifications made to the analysis method along with some techniques for taking high quality images. Taken together, they allow the algorithm to work on real images. The next few sections describe the performance of the algorithm in estimating phase angle, illumination intensity, and roughness from real images.

7.2 Estimating Phase Angle

An experiment was set up in the CIL to test the histogram analysis algorithm at estimating phase angle from real images. A series of images was taken with the camera and light source separated by an increasing phase angle. The angle was measured with a large protractor and strings to indicate the direction of the camera and light source. A diagram of the setup is shown in figure 7.2. The angles measured by this method were estimated to be accurate to within 5 degrees. The light was a 250 Watt spotlight; it was estimated to appear a few degrees across at the distance used in the experiment (about 5 feet). Therefore it is only a crude approximation of a point source.

The first image in the sequence is shown in color figure A.11. This picture was taken when the camera and light source were approximately 10 degrees apart. The phase angle was then increased by 10 degrees between each picture. The last image in the sequence is shown in color figure A.12. This picture was taken when the phase angle between the camera and light source was 90 degrees.

The program automatically split the color histograms of the objects into two clusters, fit lines to those vectors, and calculated the values of length, width, and intersection. The color histogram of the image in figure A.12 is shown in color figure A.13. The superimposed white lines show the lines fit by the program.

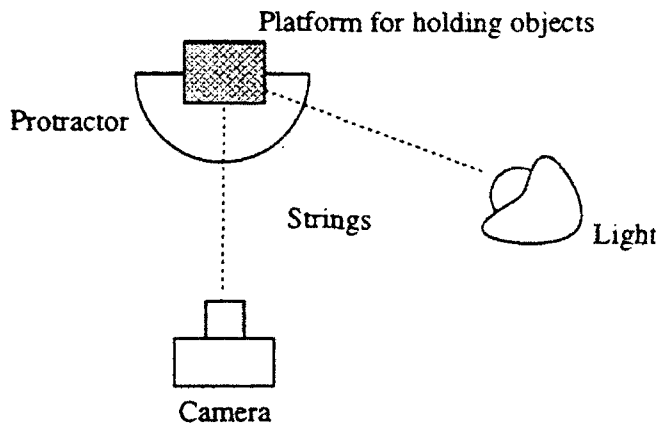


Figure 7.2: Experimental setup for estimating phase angle

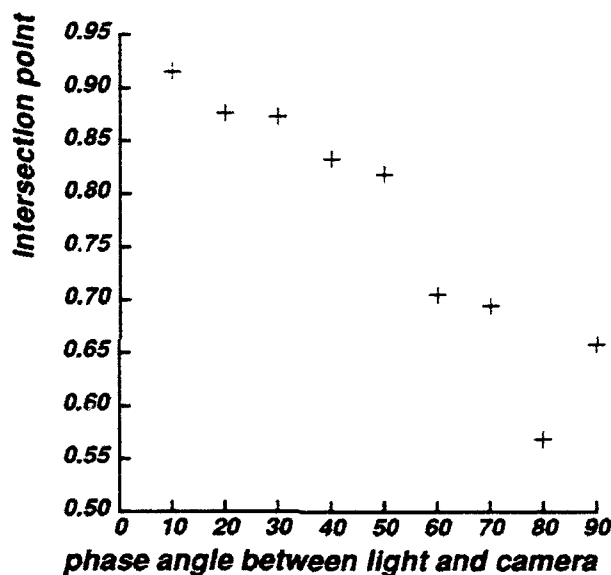


Figure 7.3: Plot of intersection point vs. phase angle

This process was repeated for each image in the sequence. Figure 7.3 shows the measured intersection value for each image, plotted against the phase angle that was in effect at the time the image was taken. A comparison of figure 7.3 with figure 3.13 in chapter 3 shows that, as predicted, the intersection decreases with increasing phase angle. The sole exception is at $\theta_p = 90^\circ$. A slight error in setting up the light could explain this problem: if the phase angle was actually slightly larger than 90° , the brightest point on the object would be out of view of the camera, throwing off the intersection ratio.

The same polynomial coefficients used in section 6.2 were used to calculate the phase angle from the length, width and intersection measurements. The lookup tables in that section assumed a camera noise $\sigma_c = 1.275$ and clipping at a maximum value of 255. The results are shown in figure 7.4. The dotted line

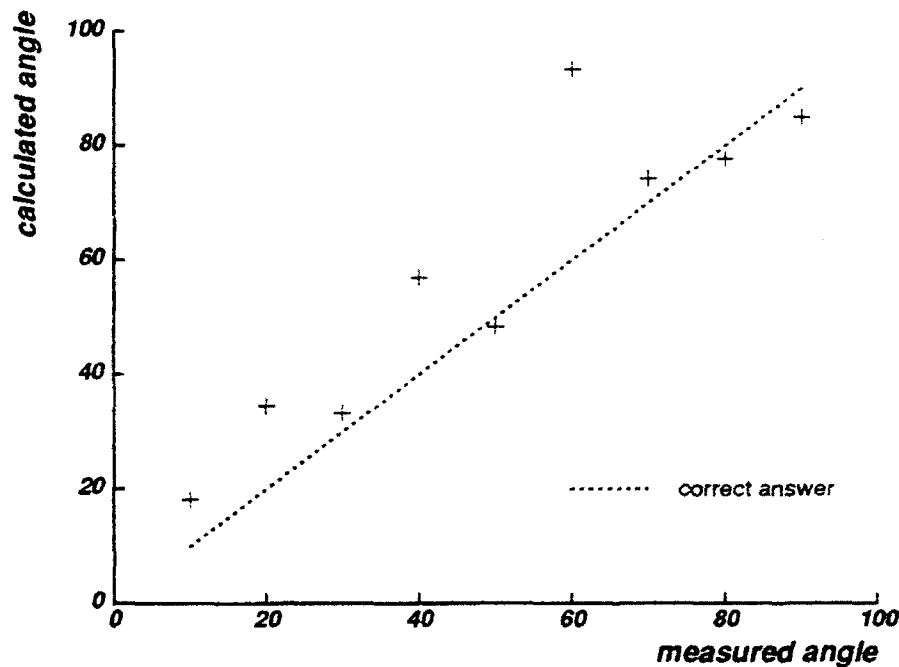


Figure 7.4: Results for experiments to calculate phase angle from real images

shows the correct answer, using the phase angle measured by the protractor as ground truth. The average error in estimating angle is 9.96° .

As the light was moved around, the distance from the object did not remain constant due to the layout of the lab. Therefore the intensity of the illumination was not assumed to be constant. For two of the images in the sequence, the illumination was close enough that the highlight saturated the camera and clipping occurred. This happened at $\theta_p = 60^\circ$ and $\theta_p = 80^\circ$. In these two cases the second set of equations (see equation 6.2) for clipping were used. Although this technique produced a reasonable answer for the case when $\theta_p = 80^\circ$, it did not do so well for $\theta_p = 60^\circ$. If this data point is eliminated from consideration, the average error in estimating the phase angle is 7.04° .

Overall the method developed for estimating phase angle from analyzing color histograms works fairly well, especially considering that the ground truth measurement of the phase angle is fairly crude. Also the lookup tables in section 6.2 were calculated without calibrating the simulated images to the conditions in the CIL. In particular, the noise of the camera was not measured precisely and the light source used in the experiments was not a point source as was used in the simulations. Also, the Torrance-Sparrow roughness model is an idealized model that assumes isotropic, Gaussian scattering; it may not describe very well the roughness of the real object used in these experiments.

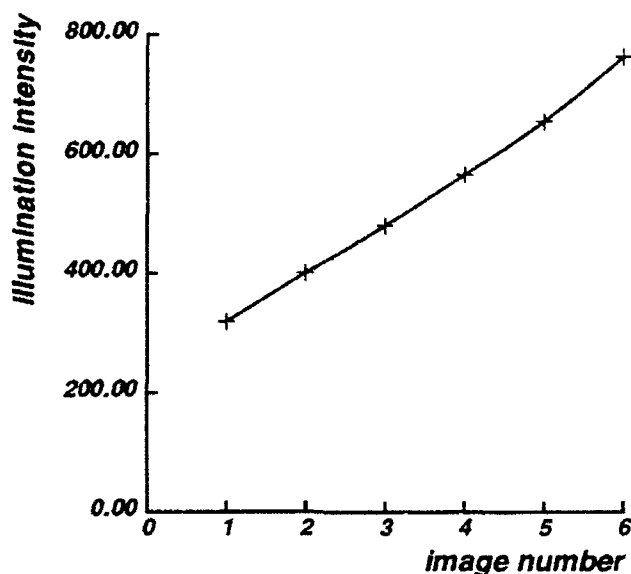


Figure 7.5: Measured illumination intensity for each image in the sequence

7.3 Estimating Illumination Intensity

A second experiment was performed in the CIL to test the performance of the algorithm at estimating illumination intensity. The spotlight was plugged into a variable voltage supply with a manually operated dial. A sequence of images was taken under increasing levels of illumination, while the imaging geometry and target object were kept constant. Altogether six images were taken. The illumination level was measured with a luminance spot meter aimed at a white card. A graph showing the luminance measurement at each image in the sequence is presented in figure 7.5. The units are candles per square meter. The spot measurements were estimated to be repeatable to within 5%.

Again, the program analyzed the histograms to produce measurements of length, width and intersection for each image. The polynomial equation to calculate illumination intensity was then applied to these measurements. The results are shown in figure 7.6. The horizontal axis shows the intensity measured by the spot meter, while the vertical axis shows the intensity estimated by the histogram analysis. The gain of the camera has not been calibrated, so the program gives a relative estimate of intensity. The dotted line shows the best linear fit to the data. If the slope of that line is considered to be the gain of the camera, then the average error in estimating illumination intensity is 5.07%.

The results in figure 7.6 were computed using only the length, width, and intersection measurements of the highlight cluster. The graph in figure 7.7 shows the roughness estimate for each image in the sequence,

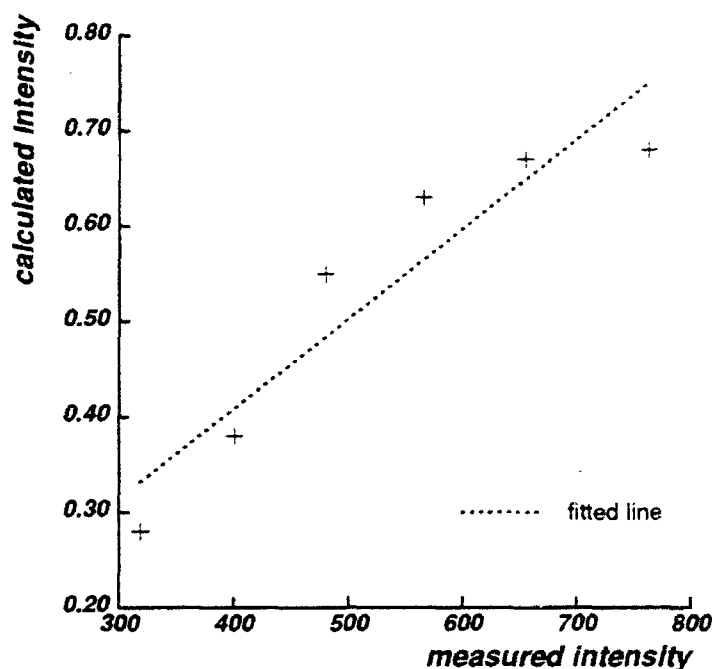


Figure 7.6: Results for calculating illumination intensity from real images

computed at the same time as the intensity estimates. Figure 7.8 shows the estimated phase angle for the same sequence. Both measurements are fairly stable across the sequence as would be expected since the imaging geometry and target object were the same in every case. The graph in figure 7.7 is particularly important, since both increasing smoothness and increasing illumination intensity will lengthen the highlight cluster. These graphs show that for the most part, the algorithm **does not** confuse changes in the histogram shape due to increasing intensity levels with the effects of roughness and phase angle.

The ground truth measurements for this experiment came from a *luminance* meter, which were compared with estimates of *intensity*. Unfortunately these measurements are somewhat different. Luminance is measured by integrating with the human luminous efficiency function $V(\lambda)$, shown in figure 7.9. A luminance meter shows how bright a spot will appear to humans, but not necessarily how bright it will appear to a color camera. A narrow band red illumination will have a much lower luminance than a narrow band green one, yet both may have the same intensity as measured by a color camera.

The difference between luminance and intensity is relevant to the results presented here, because as the voltage fed to the spotlight is decreased, the light becomes noticeably redder in hue along with becoming dimmer. This means that luminance may not correlate well with the intensity in this experiment. While the histogram analysis did correctly calculate increasing illumination intensities as the light level was increased,

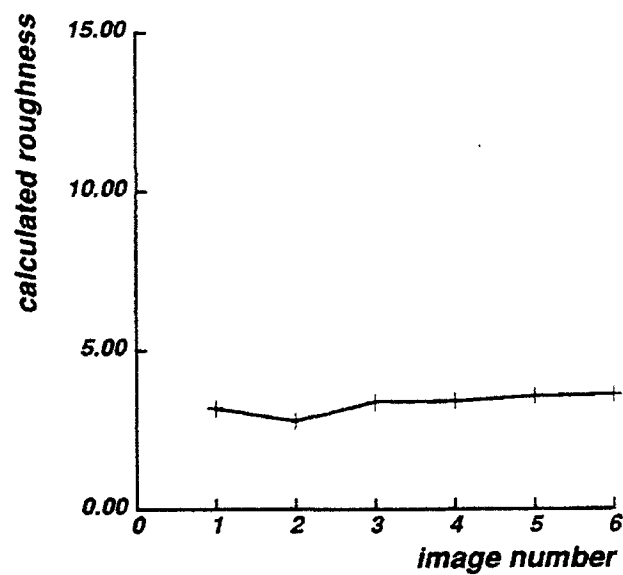


Figure 7.7: Calculating roughness when intensity changes

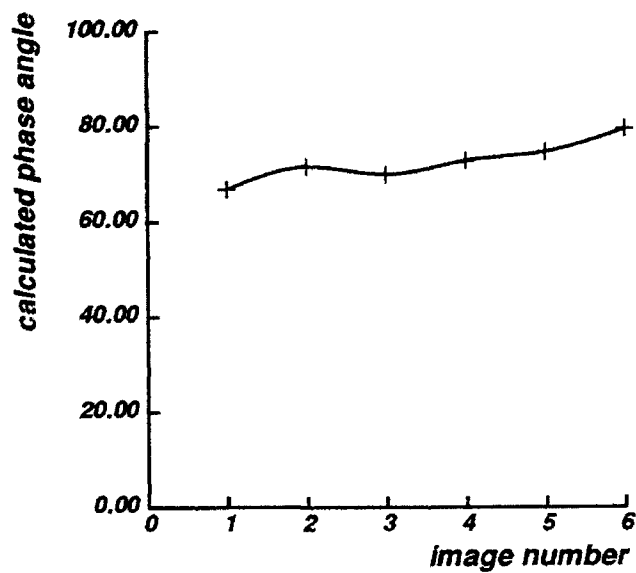


Figure 7.8: Calculating phase angle when intensity changes

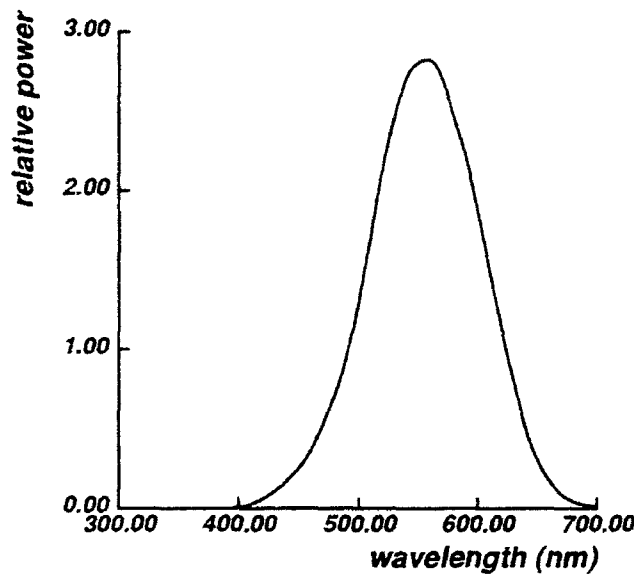


Figure 7.9: Curve used for calculating luminance

the calculated values do not appear as linear as one might hope. The use of a luminance meter might be causing some of the problems.

7.4 Estimating Roughness

A third experiment was performed to show how the system estimates surface roughness from color histograms. Color figure A.14 shows an image of a plastic inflated pool toy in the shape of an alligator. Color figure A.15 is an image of a plastic pumpkin for trick-or-treating. The pictures of the terra-cotta ball in figure A.11, the red plastic ball in figure A.8, and a red plastic pail in figure A.9 were also analyzed for this experiment.

Figure 7.10 shows the roughness calculated by the system for each of these objects. The objects are listed in order of decreasing roughness, as estimated by human observation. The calculated roughness number is the standard deviation of facet angles, in degrees.

There is no error measure for these results, since there is no ground truth data for the actual roughness values. Nevertheless, the roughness ranking from the program agrees with that produced by a human observer. The pumpkin presents a particularly interesting case, since it shows roughness at more than one scale. At the large scale where roughness is judged by touch, it is clearly the roughest object. This roughness

Object	Calculated Roughness
Alligator	10.07°
Pumpkin	8.93°
Terra-cotta ball	3.61°
Red ball	0.40°
Red pail	0.10°

Figure 7.10: Table of estimating roughness results

is large enough to be considered "texture". At the smaller scale of optical roughness, it is considered to be more shiny than the alligator. All the images are shown together in color figure A.16 for better comparison.

The result of this chapter is that the method developed in this thesis can be applied to real images. The algorithm produces reasonable estimate of phase angle, illumination intensity, and surface roughness by analyzing the dimensions of color histograms. The method is able to succeed even with such camera problems as noise, clipping, and chromatic aberration. Furthermore the method is independent of the shape of the object, and works on shapes ranging from a pumpkin to an alligator. The model used to develop the lookup tables is fairly general, and was not calibrated to match the actual imaging conditions such as light source extent, camera noise characteristics, actual distribution of facet angles, etc. It is likely that the method would perform even better if the lookup tables used measurements from real histograms rather than from simulated ones. Nevertheless, the method works reasonably well, even without prior reference to real data.

Chapter 8

Summary and Conclusions

This thesis describes the relationship between scene parameters and the color variation on an object. This variation is analyzed in a systematic way by examining the shape of the color histogram formed by the variation. Equations that describe the physical laws of reflection lead naturally to a mathematical relationship between scene properties and histogram shape. By characterizing the shape of histogram clusters in terms of simple measurements such as length, width, direction, and point of intersection, this thesis shows that it is possible to recover estimates of scene parameters from a single color histogram. The resulting algorithm works very quickly and requires only a single color image.

8.1 Contributions

This thesis make contributions to several aspects of computer vision, including highlight analysis, roughness estimation, and color constancy.

- **This thesis shows that color variation has a direct, quantitative relationship to scene properties.**

Traditional methods treat color variation as a statistical quantity, or as a source of noise to be overcome.

This thesis shows that the color variation may be modeled as a regular process that has structure which can be exploited to recover information about the scene.

- **The physical laws of reflection determine that relationship.**

Physical laws show that inhomogeneous dielectrics have two types of reflection: body reflection and surface reflection. In most cases, the two different processes of reflection have different spectral distributions, so that the color variation will have a planar distribution in color space. However the two different processes also have different *geometrical* distributions. These different distributions give rise to a particular shape of histogram within the plane.

- **The color variation is related to both color and non-color properties.**

It is obvious that the colors recorded by the camera are related to the color of the objects in the scene. The Dichromatic Reflection Model also shows how these colors are related to the color of the illumination. However, this thesis shows how the color variation is related to the non-color properties of surface roughness and the phase angle between the camera and light source.

- **This thesis presents a method for estimating roughness that is independent of the object shape.**

The physical laws of reflection show that colors seen on rough surfaces are not present on smooth surfaces and vice versa. This means that roughness is an important factor in explaining color appearance. It also means that color variation may be analyzed to recover estimates of surface roughness. The results in this thesis show that the method works for sphere-shaped objects, but also for surfaces with irregular shapes, such as a pumpkin or alligator.

- **This thesis shows how imaging geometry affects color appearance.**

Colors seen under one imaging geometry may not be present for different imaging geometries. As the camera and light source are separated by increasing phase angles, the amounts of body reflection and surface reflection at each point on the surface will shift, but in different ways. Some of the colors observed when the camera and light source are close together (small phase angle) will not be present when the camera and light source are far apart. This contributes to the difficulties caused by highlights in such tasks as stereo matching. When the viewing angle changes, not only does the *location* of the highlight on the object change, but also the actual colors seen in it.

- **There is a difference in color appearance between a rough object illuminated by a bright source, and a smooth object illuminated by a dim source.**

Previous methods to analyze color were only able to recover *relative* measurements of illumination color. However, by considering the dichromatic components of reflection, this thesis shows that the shape of the color histograms may be analyzed to recover estimates of both roughness and intensity.

Experiments demonstrate that the method is able to detect changes in illumination intensity, and is able to distinguish them from changes in surface roughness.

- Interreflection also has a regular, quantifiable effect upon the color histogram.

Interreflection occurs even in simple scenes with a small number of objects. It can have a pronounced effect upon the image appearance, and also upon the structure of the histogram. This thesis shows how the Dichromatic Reflection Model can be extended to cover interreflection. Considering each component of reflection and interreflection separately makes it possible to quantify the effects in a straightforward manner.

8.2 Directions for Future Research

The results presented in this thesis demonstrate the effectiveness of analyzing color histograms to estimate scene properties. There are several ways in which the method could be extended.

The model used in this thesis makes several limiting assumptions, and although they do not prevent the method from producing reasonable results, it is likely that the results could be even better if more realistic assumptions are incorporated into the model.

For example, the illumination is modeled as a single point light source. There are very few real images for which this model captures the nature of the illumination. Some ambient illumination is present in most real images, since there is usually interreflection from the floor, ceiling, and walls. Also, typical light sources have some extent, since most point light sources do not emit enough light to take good pictures. The light used to take the pictures in chapter 7 was not a point light source, and it is possible that the estimates of scene parameters could be improved if this were taken into account in the model.

Ambient light in the scene will have the effect of shifting the whole histogram away from the black point of the color cube. The body reflection cluster would begin, not at the black point, but at some brighter point that depends upon the ambient illumination. This might require very little adjustment to the algorithm: the length of the body reflection cluster could be measured from the maximum down to the ambient point rather than down to the black point of the color cube.

An extended light source will have the effect of making the highlight cluster wider than it would have

been for a point source. This is because the number of reflectance angles on the object where the camera and light source are lined up to produce a highlight will be greater. The extent of a light source can be incorporated into the simulation that creates the lookup tables. However, this assumes that the same size light source will be used every time, so this approach is not completely general. A more general method might be to estimate the extent of the light source by measuring the width of the highlight cluster at the upper end, where the highlight is at a maximum, and incorporate this estimate into the algorithm.

A further extension to the model would be to handle multiple extended sources of the same color. This type of lighting is common in many situations, such as office and home interiors. If the light sources are at different phase angles with respect to the camera, the resulting highlights will occur at different reflectance angles on the object. Just as the interreflection can produce a second highlight cluster, each of the extended light sources would produce a separate highlight cluster whose intersection with the body reflectance cluster depends upon the phase angle for that light.

A final extension to the illumination model is the analysis of histograms where interreflection is present. Interreflection presents the most general model for illumination, since it can act like ambient illumination (interreflection from walls); or multiple extended light sources (surface reflection from a nearby object); but can also affect the appearance in a completely different manner (e.g. body-to-body reflection). Chapter 4 showed how interreflection from a nearby object affects the appearance of the histogram, but did not present a method for automatically analyzing images with interreflection. Thus a direction of future work is to turn the predictive model into an algorithm, as has already been done for the simple histogram without interreflection.

The structure of the histogram that is described in this thesis assumes that objects are small in comparison to the distance to the camera and light source, which is the orthographic model. Under this assumption, the phase angle between the camera and light source is the same for all points on the object, so that highlights always occur at the same reflectance angle, even for objects with multiple highlights. While the orthographic assumption is a good approximation for some images, it is a poor model for images where objects are large compared with their distance from the camera. In such situations a perspective model should be used. Therefore another direction for future research is to examine how the shape of the histogram is affected by perspective projection.

Another simplifying assumption made in this work is about the nature of the surface roughness. The Torrance-Sparrow model assumes isotropic, Gaussian distribution of facet slopes. This type of scattering

is likely to characterize the differences between gloss, semi-gloss and eggshell paint finishes, where the roughness is caused by a deliberate process. However it may not describe very well roughness caused by wear or age. It is likely that scattering caused by these latter processes will have a preferred direction. A direction for future research is the extension of the current model to handle non-isotropic scattering and non-Gaussian distributions of facet angles.

The method presented here produces reasonable results on real images, even though the measurements for the lookup table came from simulated data. However, it is possible that better results could be achieved by using histogram measurements from real images to fill the lookup tables. This type of data would require calibrated samples of roughness which are difficult to obtain. However, it would provide the benefit of ground-truth data on the effect of surface roughness on histogram appearance. Measurements from real, calibrated images would also mean that the lookup tables will accurately reflect the changes in histogram shape due to changes in illumination intensity and imaging geometry. Finally, histograms from real images do not have the weakness of relying on idealized models of reflection, such as Lambertian shading. Therefore another direction of future research is comparison of the models used here with actual images of known parameters.

8.3 Applications

One possible application for the method described in this thesis is automated inspection. The roughness of surfaces may be an important measurement for determining quality in manufacturing new objects, where flaws in the finish may cause scattering of surface reflection [Cielo, 1988]. Roughness estimation might also be useful for detecting excessive wear in frequently used parts, where abrasions cause a formerly smooth object to become rough in places. The fact that the method presented here works quickly, and from a single color image, indicates that it may prove useful for moving assembly lines or other setups where the inspection must be done in a short period of time.

Another application of the type of analysis performed by this thesis is in computing color constancy. The colors in a scene measured by a camera change whenever the illumination color changes, so it is useful to be able to compute the illumination color and the colors of objects in the scene as they would appear under "white" light. Previous methods for computing illumination color from highlights rely on the intersection of more than one dichromatic reflection plane to determine the illumination color [Klinker *et al.*, 1988].

[Lee, 1986], or cannot determine illumination intensity [Lee, 1988]. By modeling the effects of the body reflection and surface reflection distributions, the method presented in this thesis is able to estimate both illumination chromaticity and intensity, even when only a single surface with highlights is present. Color constancy is useful for tasks ranging from target recognition for moving robots, to automated film processing and printing.

Estimates of the phase angle between the camera and the light source seem less intrinsically useful at first. However there is an application to automated film processing: automatically determining whether or not a flash was used in taking the picture. A built-in flash means that the phase angle between the camera direction and the illumination direction will be small. A large phase angle means that it is very unlikely that a flash was used. (An exception to this rule is provided by professional photographers who use remote flashes, but they are unlikely to trust their film to automated processing in any case.)

Estimates of the phase angle may also be useful for estimating the amount that highlights are likely to shift between images in a stereo pair. The shift in the highlight's position on the object is likely to be small if the phase angles for the cameras are similar. Also, the phase angle determines how much the highlight colors will change between stereo pairs. Changes in the intersection of the highlight cluster are equivalent to changes in the colors observed in the highlight. The intersection of the highlight cluster changes slowly for small phase angles, and more quickly for large phase angles. Estimates of phase angle could be used to give stereo matching algorithms error tolerances in matching positions and colors of highlights.

8.4 Concluding Remarks

The insights presented by this thesis have application to the general field of image understanding. This work has shown that not only do the colors of object and illumination affect color appearance of shiny objects, but also that the non-color parameters of surface roughness and imaging geometry play an important role in color appearance. The relationship between scene parameters and color variation is made explicit by examining physical models of reflection for each of the dichromatic components. A surprising amount of information can be recovered from examining a single color histogram. This indicates how complex and yet how rich the process of vision is, and why color is particularly valuable for image understanding.

Appendix A

Color Figures



Figure A.1: Histogram of shiny billiard ball



Figure A.2: Histogram of matte billiard ball

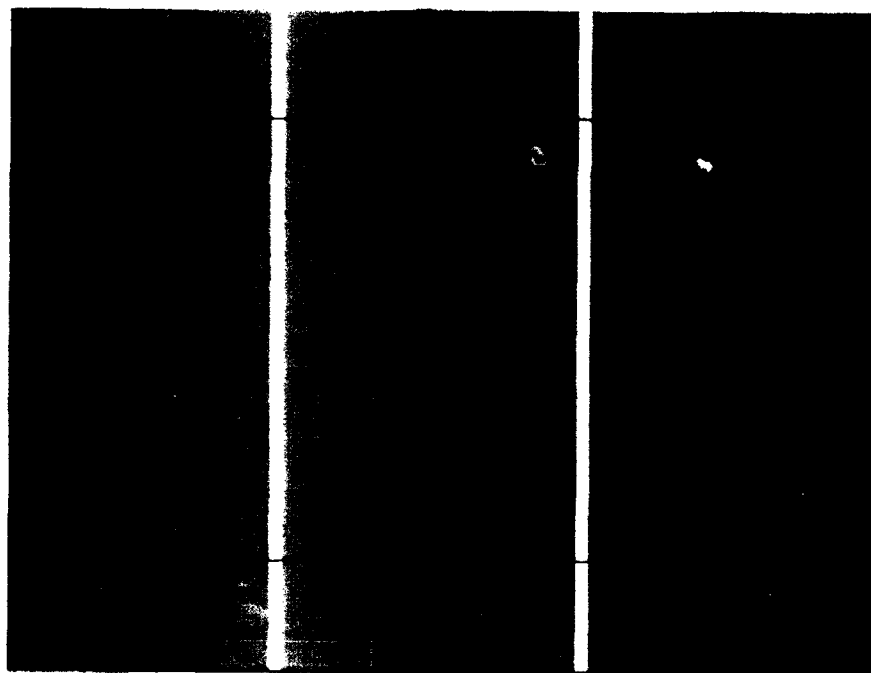


Figure A.3: Simulation results

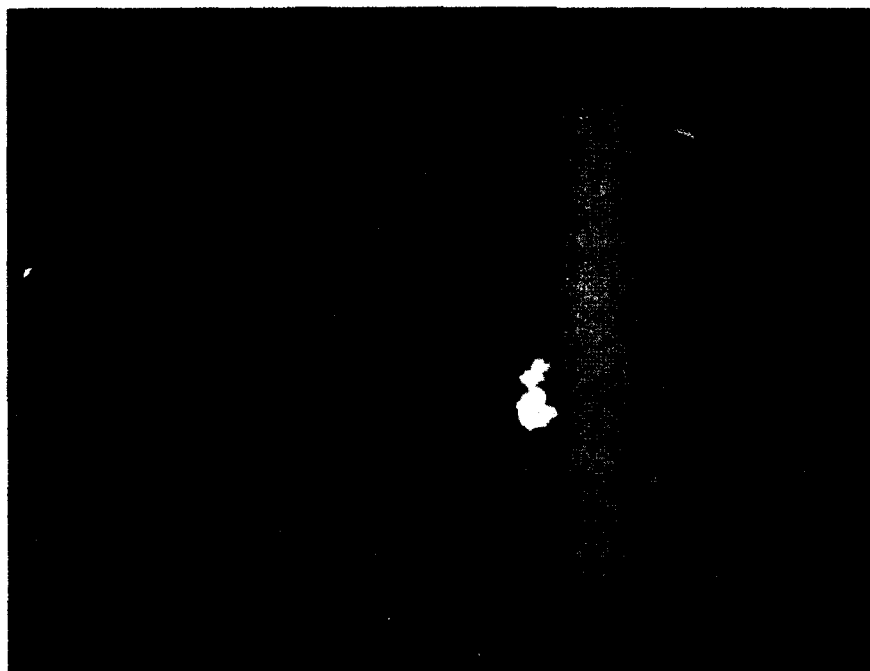


Figure A.4: Simulated test image

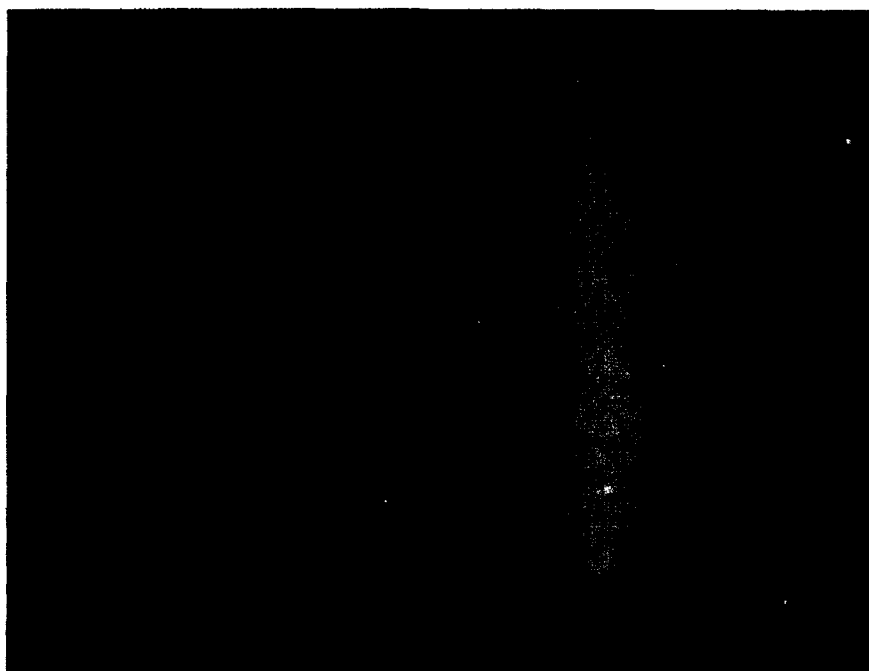


Figure A.5: Image simulated from recovered parameters



Figure A.6: Intrinsic images calculated with a skewed highlight color

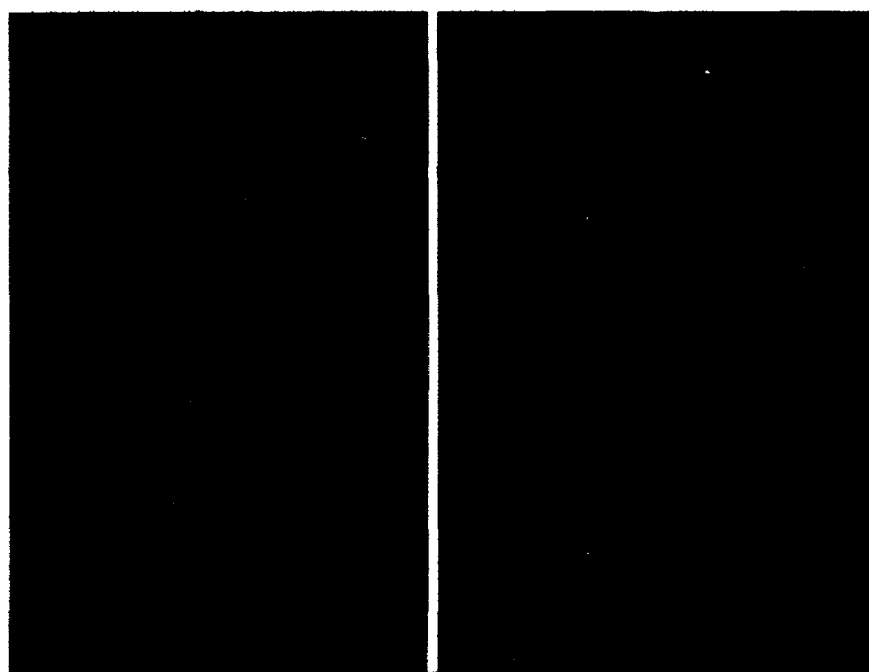


Figure A.7: Intrinsic images calculated with the corrected highlight color

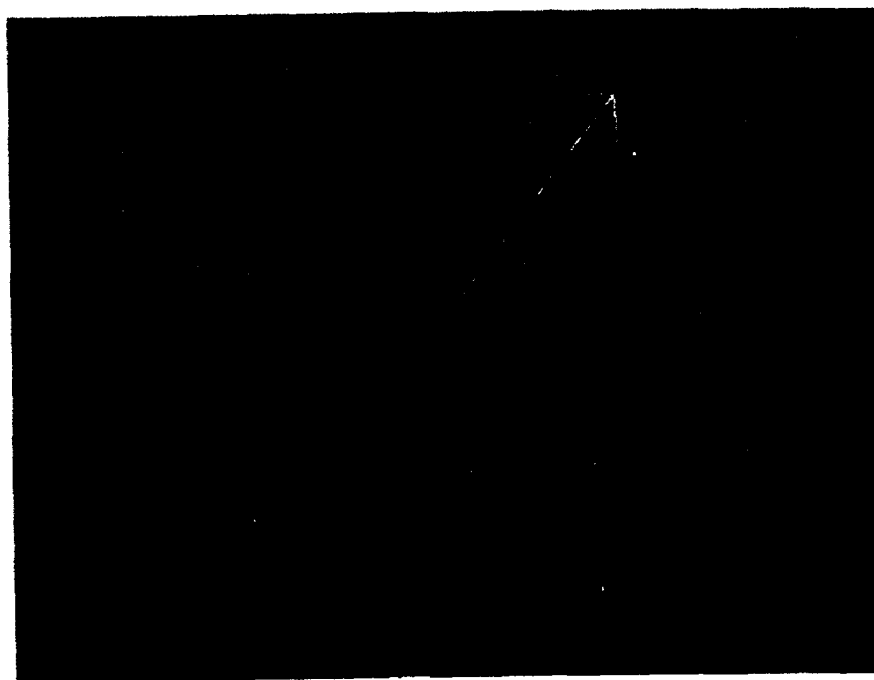


Figure A.8: Histogram for shiny ball

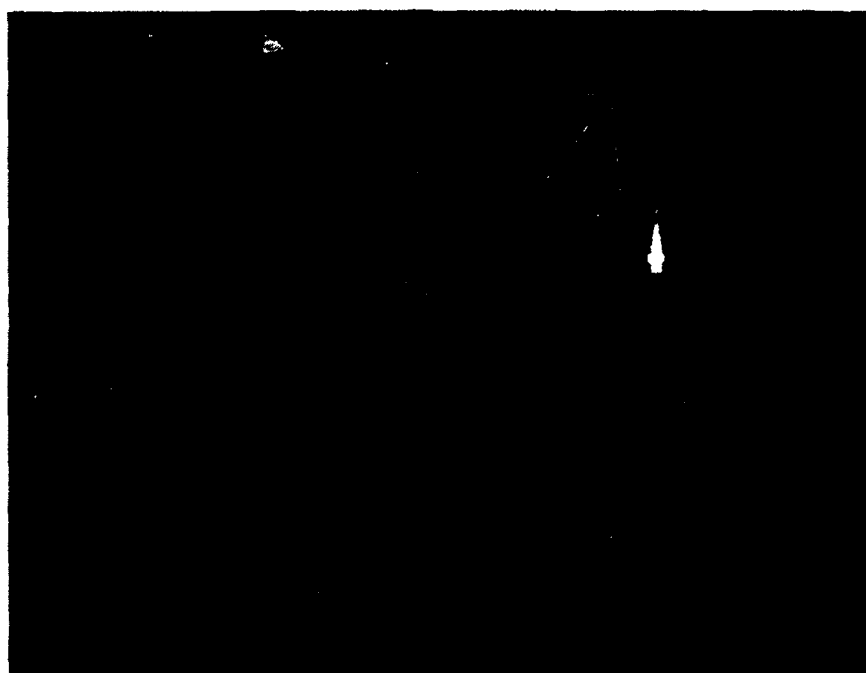


Figure A.9: Histogram for very shiny pail



Figure A.10: Closeup of bloomed area

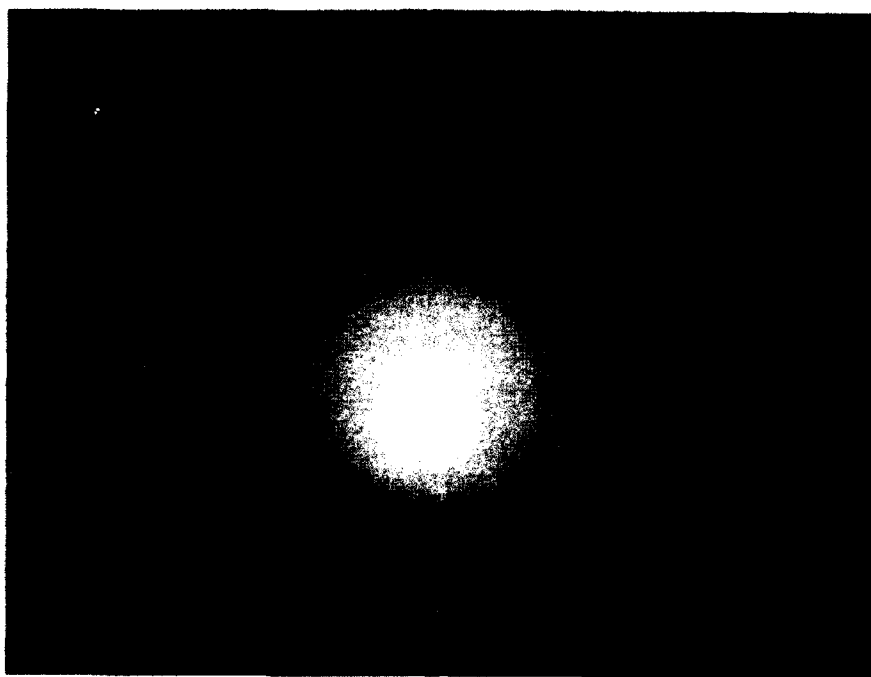


Figure A.11: Image with phase angle of 10 degrees

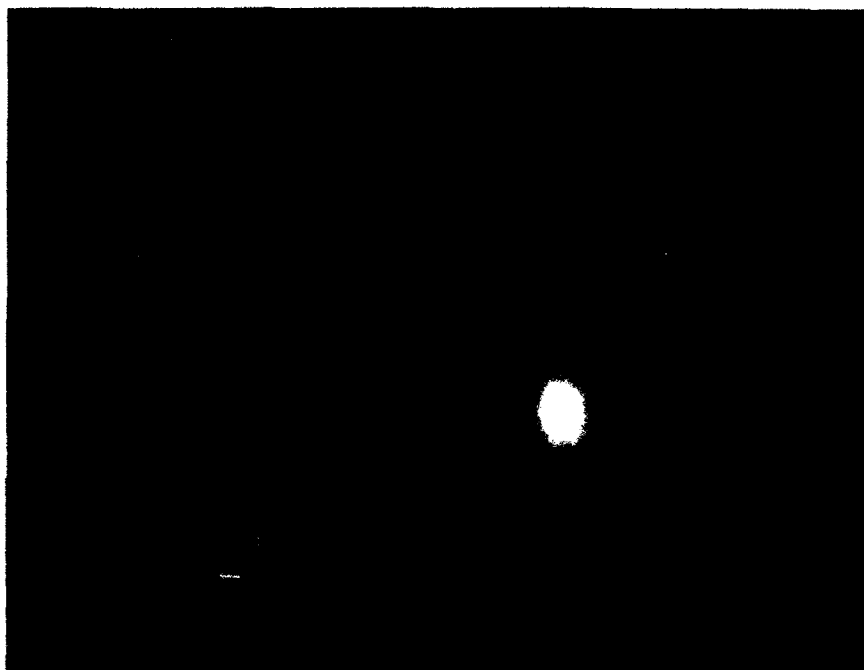


Figure A.12: Image with phase angle of 90 degrees

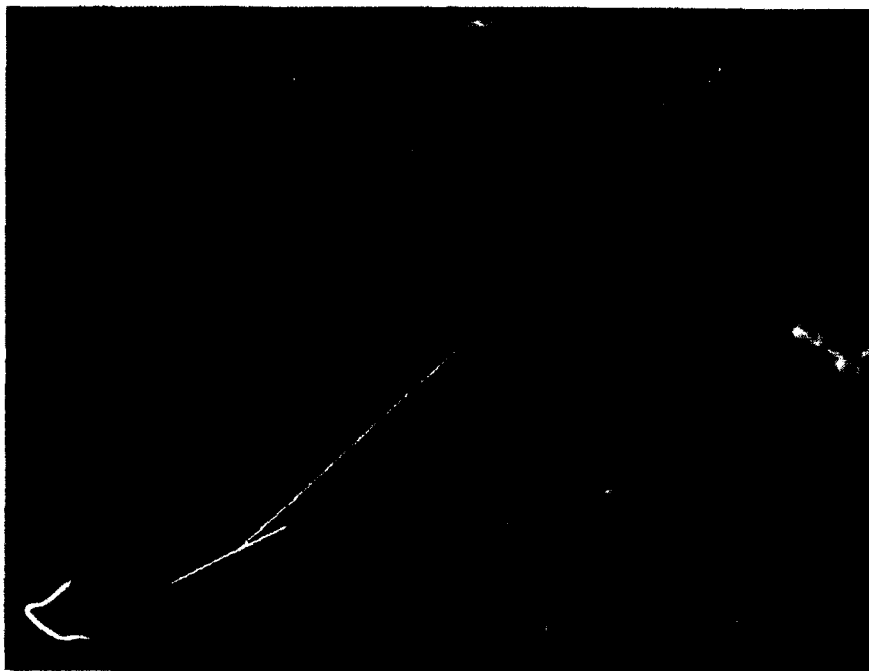


Figure A.13: Histogram for phase angle of 90 degrees

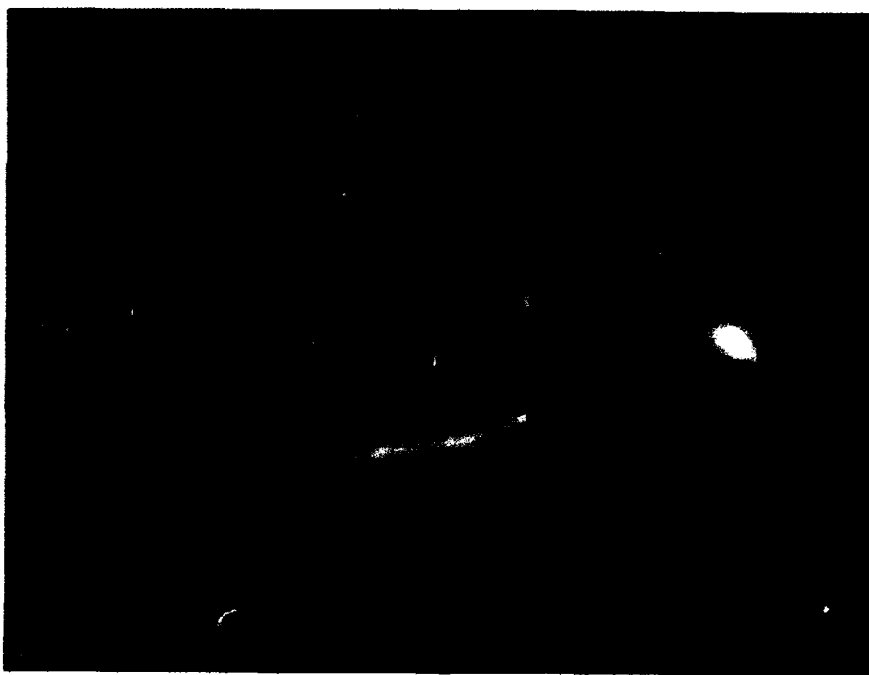


Figure A.14: Image of plastic alligator

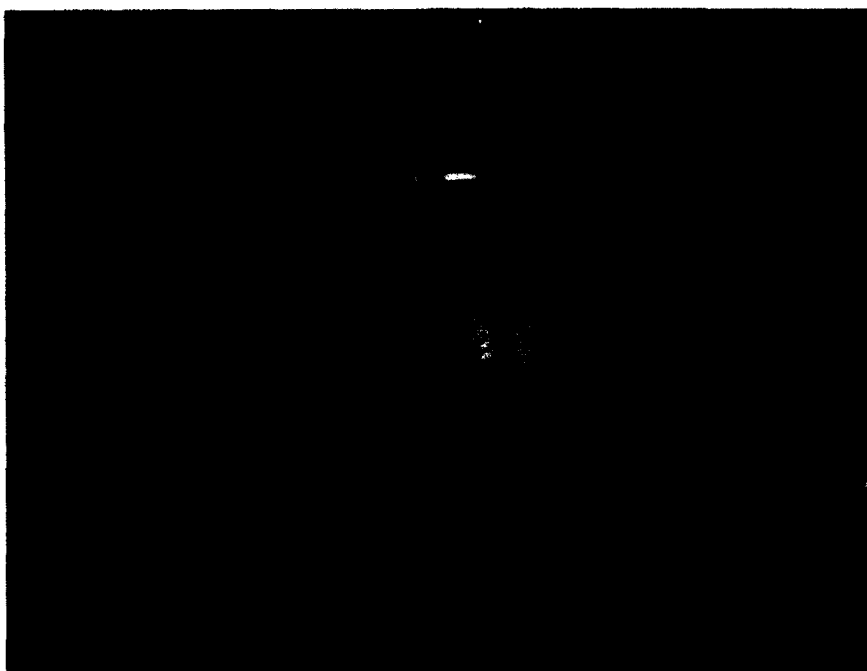


Figure A.15: Image of plastic pumpkin



Figure A.16: Five objects with different roughness values

Bibliography

- [Bajcsy *et al.*, 1989] R. Bajcsy, S. W. Lee, and A. Leonardis. Image segmentation with detection of highlights and inter-reflections using color. MS-CIS 89-39, University of Pennsylvania, Philadelphia, June 1989.
- [Bajcsy *et al.*, 1990] R. Bajcsy, S. W. Lee, and A. Leonardis. Color image segmentation with detection of highlights and local illumination induced by inter-reflections. In *Proc. 10th International Conference on Pattern Recognition*. IEEE, 1990.
- [Beckmann and Spizzichino, 1963] P. Beckmann and A. Spizzichino. *The Scattering of Electromagnetic Waves from Rough Surfaces*. The Macmillan Company, 1963.
- [Brill, 1989] M. H. Brill. Object-based segmentation and color recognition in multispectral images. *SPIE/SPSE*, January 1989.
- [Cielo, 1988] P. Cielo. *Optical Techniques for Industrial Inspection*. Academic Press, Inc., 1988.
- [Crisman and Thorpe, 1991] J. D. Crisman and C. E. Thorpe. Unscarf, a color vision system for the detection of unstructured roads. In *Proc. of International Conference on Robotics and Automation*, pages 2496–2499. IEEE, 1991.
- [Drew and Funt, 1992] M. S. Drew and B. V. Funt. Variational approach to interreflection in color images. *J. Opt. Soc. Am. A*, 9(8):1255–1265, August 1992.
- [D’Zmura and Lennie, 1986] M. D’Zmura and P. Lennie. Mechanisms of color constancy. *J. Opt. Soc. Am.*, 3(10):1622–1672, October 1986.
- [Forsyth and Zisserman, 1989] David Forsyth and Andrew Zisserman. Mutual illumination. In *CVPR*. IEEE, 1989.

- [Funt and Drew, 1991] B. V. Funt and M. S. Drew. Color space analysis of mutual illumination. CSS/LCCR TR 91-03, Simon Fraser University, Burnaby, B.C., Canada, 1991.
- [Funt *et al.*, 1991] B. V. Funt, M. S. Drew, and J. Ho. Color constancy from mutual reflection. *International Journal of Computer Vision*, 6(1):5–24, 1991.
- [Gershon, 1987] R. Gershon. *The Use of Color in Computational Vision*. PhD thesis, University of Toronto, 1987.
- [Goral *et al.*, 1984] Cindy M. Goral, Kenneth E. Torrance, Donald P. Greenberg, and Bennett Battaile. Modeling the interaction of light between diffuse surfaces. *ACM Computer Graphics*, 18(3), July 1984.
- [Haralick and Kelly, 1969] R. M. Haralick and G. L. Kelly. Pattern recognition with measurement space and spatial clustering for multiple images. *Proc. IEEE*, 57:654–665, April 1969.
- [He *et al.*, 1991] X. D. He, K. E. Torrance, F. X. Sillion, and D. P. Greenberg. A comprehensive physical model for light reflection. *Computer Graphics (SIGGRAPH '91 Proceedings)*, 25(4):175–186, July 1991.
- [Healey and Binford, 1987a] G. Healey and T. O. Binford. The role and use of color in a general vision system. In *Proceedings of the ARPA Image Understanding Workshop*, pages 599–613. DARPA, 1987.
- [Healey and Binford, 1987b] G. Healey and T.O. Binford. Local shape from specularity. *Proc. Image Understanding Workshop*, 2:874–887, February 1987.
- [Healey and Kondepudy, 1991] G. Healey and R. Kondepudy. Modeling and calibrating ccd cameras for illumination insensitive machine vision. In *SPIE No. 1614 Conference on Optics, Illumination, and Image Sensing for Machine Vision VI*. SPIE, 1991.
- [Healey, 1988] G. Healey. A color reflectance model and its use for segmentation. In *International Conference on Computer Vision (ICCV)*, pages 460–466. IEEE, 1988.
- [Horn, 1977] B. K. P. Horn. Understanding image intensities. *Artificial Intelligence*, 8:201–231, 1977.
- [Ikeuchi and Sato, 1990] Katsushi Ikeuchi and Kosuke Sato. Determining reflectance parameters using range and brightness images. *ICCV*, 1990.
- [Jang, 1991] Younghee Jang. Identification of interreflection in color histograms using a physics-based reflection model. In *Computer Vision and Pattern Recognition*, pages 632–637. IEEE, 1991.

- [Klinker *et al.*, 1987] G. J. Klinker, S. A. Shafer, and T. Kanade. Using a color reflection model to separate highlights from object color. In *International Conference on Computer Vision (ICCV)*, pages 145–150. London, June 1987. IEEE.
- [Klinker *et al.*, 1988] G. J. Klinker, S. A. Shafer, and T. Kanade. The measurement of highlights in color images. *International Journal on Computer Vision*, 2(1):7–32, 1988.
- [Klinker, 1988] Gudrun J. Klinker. *A Physical Approach to Color Image Understanding*. PhD thesis, Carnegie Mellon University, May 1988.
- [Lee, 1986] H. C. Lee. Method for computing the scene-illuminant chromaticity from specular highlights. *J. Opt. Soc. Am. A*, 3(10):1694–1699, October 1986.
- [Lee, 1988] H. C. Lee. Estimating the illuminant color from the shading of a smooth surface. A.I. Memo 1068, Massachusetts Institute of Technology Artificial Intelligence Laboratory, August 1988.
- [Lee, 1991] S. W. Lee. *Understanding of Surface Reflections in Computer Vision by Color and Multiple Views*. PhD thesis, University of Pennsylvania, 1991.
- [Mundy and Porter, 1980] J.L. Mundy and G.B. Porter. Visual inspection of metal surfaces. In *Proc. 5th ICPR*. IEEE, 1980.
- [Nayar *et al.*, 1990] S. K. Nayar, K. Ikeuchi, and T. Kanade. Shape from interreflections. In *International Conference on Computer Vision*, pages 2–11. IEEE, 1990.
- [Nayar *et al.*, 1991] S.K. Nayar, K. Ikeuchi, and T. Kanade. Surface reflection: Physical and geometrical perspectives. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 13(7):611–633, July 1991.
- [Novak *et al.*, 1990] C. L. Novak, S. A. Shafer, and R. G. Willson. Obtaining accurate color images for machine vision research. *Proc. SPIE*, 1990.
- [Photometrics, 1989] Ltd. Photometrics. Charge-coupled devices for quantitative electronic imaging. Tucson, AZ, 1989.
- [Porter and Mundy, 1981] G.B. Porter and J.L. Mundy. Automatic visual inspection of metal surfaces. *Proc. SPIE 281: Techniques and Applications of Image Understanding*, pages 176–181, 1981.
- [Sato and Ikeuchi, 1992] Y. Sato and K. Ikeuchi. The temporal-color space analysis of reflection. Technical Report CMU-CS-92-207, Carnegie Mellon University, 1992.

- [Shafer *et al.*, 1990] S. A. Shafer, T. Kanade, G. J. Klinker, and C. L. Novak. Physics-based models for early vision by machine. *Proc. SPIE*, 1990.
- [Shafer, 1984] S. A. Shafer. Using color to separate reflection components. Computer science department, University of Rochester, 1984.
- [Sloan, 1977] K. Sloan. *World Model Driven Recognition of Natural Scenes*. PhD thesis, University of Pennsylvania Moore School of Electrical Engineering, June 1977.
- [Solomon and Ikeuchi, 1992] F. Solomon and K. Ikeuchi. Extracting the shape and roughness of specular lobe objects using four light photometric stereo. In *Computer Vision and Pattern Recognition*, pages 466–471. IEEE, June 1992.
- [Swain, 1990] M. J. Swain. *Color Indexing*. PhD thesis, University of Rochester, November 1990.
- [Tominaga and Wandell, 1989] S. Tominaga and B. A. Wandell. Standard surface-reflectance model and illuminant estimation. *J. Opt. Soc. Am. A*, 6(4):576–584, April 1989.
- [Torrance and Sparrow, 1967] K. Torrance and E. Sparrow. Theory for off-specular reflection from roughened surfaces. *Journal of the Optical Society of America*, 57(9):1105–1114, September 1967.
- [Willson and Shafer, 1991] Reg G. Willson and Steven A. Shafer. Active lens control for high precision computer imaging. In *International Conference on Robotics and Automation*. IEEE, 1991.
- [Willson and Shafer, 1992] Reg G. Willson and Steven A. Shafer. Precision imaging and control for machine vision research at carnegie mellon university. In *SPIE Conference of High Resolution Sensors and Hybrid Systems*. SPIE, 1992.
- [Wolff, 1992] L. B. Wolff. Diffuse reflection. In *Computer Vision and Pattern Recognition*, pages 472–478. IEEE, June 1992.