# Aspects of Image Reshading 

by<br>CHRISTOPHER ANTHONY ROMANZIN<br>B.Sc. The University of Calgary, 1990

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of

Master of Science
in the Faculty of Graduate Studies
Department of Computer Science

We accept this thesis as conforming to the required standard

## THE University of British Columbia


#### Abstract

In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.


Department of COMPUTER SCIENCE
The University of British Columbia
Vancouver, Canada

$$
\text { Date MARCH } 71995
$$


#### Abstract

Old images are often used in the creation of new images, either to enhance the appearance of the result or to achieve a manual or computational savings. Without ample care this practice can lead to missing or conflicting visual cues in the result, since an old image may exhibit shading artifacts that are inconsistent with the scene it is incorporated into. Therefore there is a need to process a source image so that it is consistent with the way it is to be used. Current methods for altering the shading artifacts found in an image are largely ad hoc, pixel based and are somewhat unintuitive. This work explores methods for enabling a user to manipulate 3D shading artifacts in an image, that is, performing image editing operations that relate to physical processes such as moving and dimming a light source, or changing the reflectance properties of objects in an image-without having full knowledge of the scene properties. We call this goal one of image reshading, and it is closely tied between the disciplines of computer graphics and computational vision as it involves generating images and inferring properties of the scene that give rise to an image.

Image reshading is an enormous problem of its own, and this work explores only a few aspects of it. The first is the detection and removal of specular highlights from image data alone. Current techniques are explored and applied to textured images that are commonly used in computer graphics. The second image reshading task examined is to solve for the geometry of a light source illuminating a scene given an image of the scene and the geometry of the visible objects. A series of constraints formed by the shading of Lambertian and Phong reflectors is presented and a strategy for determining the position, orientation, and size of a rectangular source is demonstrated. Finally, given an image, a geometric model of the objects in the image, and the light source distribution, a method for solving for the relative emissive strengths and the reflectance parameters of surfaces in the image is given. This final reshading operation allows a large number of useful image editing operations to be performed.


## Table of Contents

Abstract ..... ii
Table of Contents ..... iii
List of Figures ..... v
Acknowledgements ..... vii
Chapter 1 Introduction ..... 1
1.1 Aims and Contributions of this Work ..... 6
1.2 Thesis at a Glance ..... 6
Chapter 2 A Little Reflection Background ..... 7
2.1 Definitions and Terms ..... 7
2.2 General Framework of Reflection ..... 10
2.3 Some Models of Light Reflection ..... 14
2.4 Sensor Model ..... 20
Chapter 3 Specular Highlight Analysis ..... 21
3.1 Previous Work in Adding and Removing Shading Cues from Images ..... 22
3.2 What is a Specular Highlight? ..... 24
3.3 Detection of Specular Highlights ..... 27
3.3.1 Previous Work: Specular Highlights as Signatures in Colour Space ..... 28
3.3.2 Previous Work: Specular Highlights as Violations ..... 32
3.4 Removal of Specular Highlights ..... 34
3.4.1 Get Real ..... 37
Chapter 4 Reshading Knowing Object Geometry ..... 45
4.1 Pixels Torn Asunder: Previous Work in Reshading ..... 48
4.2 Rectangular Source from Geometry ..... 50
4.2.1 How to Find That Light? Previous Work ..... 52
4.2.2 Problem Formulation ..... 53
4.2.3 Constraints from Diffuse Shading ..... 55
4.2.4 Constraints from Specular Highlights ..... 62
4.2.5 Experiment ..... 65
4.3 Solving for Radiometric Properties ..... 67
4.3.1 Reshading of Diffuse Surfaces ..... 67
4.3.2 Handling Specularities ..... 70
4.4 Summary ..... 71
Chapter 5 The End ..... 72
5.1 Conclusions and Contributions ..... 72
5.2 Future Work ..... 73
Bibliography ..... 75
Appendices ..... 79
A Notation ..... 80
B Fresnel Reflection ..... 81
C The Relationship Between Image Irradiance and Scene Radiance ..... 83

## List of Figures

1.1 Reasons for recycling images ..... 3
1.2 Problems in reusing images ..... 4
2.1 Model of light reflection at an inhomogeneous surface ..... 10
2.2 Local coordinate system for integrating over the surrounding hemisphere ..... 11
2.3 Notation for three point transport ..... 12
2.4 Local geometry for the Phong and Torrance-Sparrow shading models ..... 15
2.5 Distribution of pixel values in colour space following the NIR model ..... 16
2.6 Angular spectral reflectance curves for titanium and potassium bromide ..... 17
2.7 Illustration of theoretical partial reflectance for different angles of incidence ..... 19
3.1 An example of texture mapping ..... 23
3.2 Some examples of specular highlights ..... 26
3.3 The chromaticity convergence method ..... 29
3.4 The shape of the spectral cluster for a cylindrical object ..... 31
3.5 Detection à la Klinker et al. ..... 33
3.6 Effects of surface roughness on pixel distribution in colour space ..... 36
3.7 Skew geometry and reshading error ..... 37
3.8 A comparison of removal techniques on a BC Golden Delicious apple ..... 38
3.9 A comparison of removal techniques on a jalapeno pepper ..... 39
3.10 Effects of texture on colour space distribution ..... 40
3.11 Signature searching methods failing to detect specularities ..... 41
3.12 A comparison of removal techniques on a small-grain texture ..... 42
3.13 A comparison of removal techniques on a wood texture ..... 43
3.14 Highlight removal on a pop can ..... 44
4.1 An example of adding light and objects to a real scene ..... 49
4.2 Parameters for defining the geometry of a rectangular source ..... 54
4.3 Conditions where the reflected radiance at point $\mathbf{x}_{r}$ becomes zero ..... 56
4.4 Fixing the light plane from Lambertian boundaries ..... 57
4.5 Coordinate system for maximizing the form factor of a constrained rectangle ..... 58
4.6 Shading of a polygon, orthogonal source ..... 59
4.7 Shading of a polygon: oriented and perpendicular source ..... 60
4.8 Phong shading of a polygon ..... 63
4.9 Source-enclosing cone given by a highlight boundary ..... 65
4.10 Sharp versus diffused highlight boundary ..... 66
4.11 Solving for a rectangular source from the shading on a synthetic object ..... 67
C. 1 Image forming system ..... 84

## Acknowledgements

When's mental health week?
I am a sponge; lacking my own personality, I tend to take on the characteristics and mannerisms of those around me. As a result, everyone that I have met here during these too many years must take an equal share of the blame for this thesis being absurdly late. And there's a lot of blame to go around, as I've been blessed with a lot of good friends. Beeell, to the Yukon River we must go. Jeff, Raveen's got nothing on you, including fashion. Kevin. Need I say more? Words cannot describe the joy in my heart when xbiff goes off and I know it's you. Raza Minda Gone, you will lead thousands of people to their death one day, effortlessly and without remorse. Gwen, your furniture saved my summer but beware the batik, it spoke to me on occasion. Rob and Scott, thanks for many mellow evenings at your pad, and for introducing me to the glory of scotch. Pierre, the grad centre was never the same without you. I listened to your English for years, but ha ha now the shoe's on the other foot. Kori, a word of advice: get off the turpentine, Davis Inlet isn't far behind. Chris and Hiroko, my stomach sings when I drive near your place. Carl, you made coffee not just a beverage, but an experience. Elchi, never, never forget: sometimes you just gotta fuck shit up. Keep moving those pixels Yg , and maybe like all great cereals they'll spell something. I will sorrily miss ESB, who provided comfort on many a dry day.

Alain. Oh Alain. I have tested the bounds of your seemingly infinite patience. Yes, I do realize that it took me a year and a half to write my introduction, then I threw it away and rewrote it in a few days. But it's a pretty good intro. Thanks for your guidance, good-natured support, for going slowly when I wasn't catching on, and for not beating me with the stick of shame even though I really deserved it. You treated me like a peer at all times even though I rarely felt like one. I feel fortunate to have been your student.

Thanks to Jim Little, my second reader, for his constructive comments on this work under a tight deadline. Kori, Kevin C., Yunfei: thanks for letting me share your office these past few months, hog three bookshelves, and for generally putting up with me. Derek Upham, I don't know if you're lying dead in a ditch somewhere or what, but I used your desk for the last six months. You're kinda weird so maybe that bothers you, but that's just the way things go.

Friends come and go, but you always have family. Mom and Dad, Rob, Ger, Valerie, Dave and Glenda; we may not talk often, or very seriously, but you are all important to me. The older I get and the more people I meet, the more I realize what a great and special family we have. Hats off to Mom and Dad. Carol and Kelly, you're family too.

Well have a fun life everybody. Can I please go now? I really want to.

Financial support for this research and much of my CD collection was provided in part by the foolish but well-meaning taxpayers of Canada. Thanks welfare state.

## Chapter One

## Introduction



Standing on the shoulders of giants leaves me cold

REM

RENDERERS, magicians, stereo equipment salesmen and other shysters ply their trade with a careful balance of truth and trickery. The "truth" in image synthesis-from a puritanically photorealistic standpoint-is in rigorously modelling the phenomena or objects to be depicted, describing the parameters affecting their appearance and behaviour, and simulating the propagation of light through the environment and towards the viewer. Done well, the results obtained can be strikingly realistic ${ }^{\dagger}$.

[^0]Unfortunately, the effort required to produce this level of quality is great, and there are limits to what can be convincingly portrayed. Modelling and animating are tedious, exacting tasks, which quickly become overwhelming for complex scenes. Further, the descriptive power of synthetic models and rendering algorithms are not yet fully developed. Many important effects (such as the natural, worn appearance of objects or realistic looking and acting people) are effectively beyond the capability of current techniques. Practically, these shortcomings mean that there are significant limits to the kind of imagery that can be produced solely with computer graphics.

Hence the need for "trickery"-approaches that complement the conventional modelling and rendering process. One way to portray an effect while avoiding the pain of simulation it is to take it from another picture. Since images are costly to generate and usually contain similar elements, using image data in lieu of synthetic models is a natural strategy, one that often gives comparable results at considerable savings. Several methods for reusing real and synthetic imagery ${ }^{\dagger}$ have been developed, each varying in cost, versatility, and in their treatment of the source image. The fundamental yet conceptually simple operation of overlaying portions of two or more images to create a single montage-image compositing-forms the core of most postprocessing packages in use today, while a common render-time approach is to modulate the shading calculation at a surface with an image that is mapped onto it (texture mapping). Both indispensable for creating interesting and realistic pictures by computer, these two techniques use a source image as visual spice to enhance the flavour of a bland result. Alternatively, an image can be used as a canvas which is transformed through painting, colour correction, or one of numerous other pixel-based operations. A more recent role for existing images is as an example which does not directly dictate the appearance of a result but only suggests its look for goal driven renderers.

Recycling old pixels is lucrative as it yields visual detail without incurring the full expense of modelling and rendering that detail. It also expands the range of imagery that can be produced, since by combining real images and computer graphics one can portray objects and effects that would otherwise be difficult, if not impossible, to generate. It has its down side too, for the freedom to combine and manipulate imagery that gives rise to these benefits also endangers the integrity of the result. Whereas pictures taken with a camera (and, arguably, "photorealistic" computer generated images) conform to familiar notions of how the world works-light travels in straight lines and does not pass through opaque materials, energy is conserved,

[^1]

Figure 1.1: Reasons for recycling images: modelling and animating some objects is extremely difficult, while capturing them on film is easy.
etc.-an image that is combined with other elements or altered in an unconstrained way may lose physical plausibility, and with it, the illusion of realism. Derivative images may suffer from internal inconsistencies in the form of incomplete (noticeably absent) or incorrect (just dog wrong) visual cues ${ }^{\dagger}$. Incompleteness results from failing to account for some interactions between objects in a combined scene, and can take the form of missing shadows or lights that do not affect all objects. This is clearly a problem when images are brought together in a cut-and-paste fashion, since the elements of a montage only interact through occlusion. To heighten the appearance of congruity shading cues need to be added to the combined result. Incorrect visual cues are those that are either portrayed poorly or that should not appear, and range from the obvious (such as a shadow being cast in the wrong direction) to the subtle (a specular highlight being a bit too bright). In general, when an image is combined with another image or synthetic scene, any difference in surrounding object geometry, viewing or illumination conditions between the source scene and those present in the scene it is introduced into will lead to conflicting shading artifacts. For example, the self-shadowing found in the bark of a tree suggests an illumination direction to the viewer, which can cause problems if the image is later used as a bark texture. For similar reasons care must be taken when altering a realistic image, as any modification implies a physical process, which in the mind of a viewer should cause other effects in the image. For example, by adding a specular highlight onto the surface of an object a light source direction is suggested, which a viewer expects to influence the rest of the scene. Figure 1.2 serves to illustrate some of

[^2]the problems involved in maintaining consistency in a derived image.


Figure 1.2: Shading artifacts in images are a result of the viewing and lighting conditions present in the scene when the image was produced. If the image is subsequently used in another scene under different conditions, then visual inconsistencies may arise, requiring the (left) addition and (right) removal of shading cues from the source image. (left) The sardine can and bean can (barely visible under the plant) were composited into the countertop image. Shadows cast both onto and from these cans must be added manually. The artist must estimate how dark a shadow should be and how bright the cans should appear. Information about the strength and location of the real light sources would be helpful, and could lead to an automated techniques. (right) Noticeable shading artifacts on textures are bad, are sometimes preventable, and should be removed.

Modifying the shading effects in an image is largely an ad hoc process now, with cues improvised after the fact by a skilled graphic artist ${ }^{\dagger}$. To achieve consistency between many composited sources, colour correction is used to bring hues and intensities in line, and noise is added as needed. Missing cues are added through 2D painting, which offers the user no constraints. Render-time approaches such as texture mapping cannot remove shading effects from an image but can add them, although these are not based upon the content of the image. To minimize problems an image that is to be used as a texture is usually captured under neutral lighting conditions, but this requires a level of control that is not always possible to achieve. Also, except for the most simple of objects, it is not possible to image it entirely under normal illumination due to nonplanarity of the object and nondirectionality of the source, so shading artifacts are inevitably

[^3]present. Eliminating these problems requires gaining control over the appearance of the source image. This is turn requires an accurate understanding of the conditions in the source scene, because to fake or undo the physical processes of shading (either manually or algorithmically) one needs an understanding of those processes and the factors involved. A simple example is the case of adding a shadow: what direction should it be cast in, what should the spatial extent of the umbra and penumbra be, how dark must they appear, etc. The answers come from knowing the position and emissive strength of the light source and the geometry of the objects involved-which can either be measured beforehand or inferred by the artist. It would be nice if the computer could help determine these parameters.

Ideally, a user would have the means for manipulating 3D shading artifacts, or performing image editing operations that relate to physical processes such as moving or dimming a light source or changing the reflectance properties of a surface. Such higher-level effects are often precisely what one wants to achieve when working on an image and standard 2D image processing operations are merely a means to achieve these. We refer to the general problem of altering the appearance of an image realistically in terms of its effects as image reshading. Stated loosely, the challenge is given an image and some information about the scene conditions under which it was obtained, generate a new image of the scene under a different set of conditions. "Conditions" implies any rendering parameters other than camera pose: light source geometry, radiometric properties of lights and objects, object geometry, etc.

Although clearly related to rendering, reshading an image is nevertheless a distinct challenge. The fundamental difference is that rendering is a forward problem while reshading involves and inverse problem. When dealing with an image one has incomplete knowledge about the depicted scene, while global illumination algorithms require complete information about the environments they are applied to. To properly alter shading cues requires knowledge of the parameters that brought those cues about, which leads to the task of inferring scene properties algorithmically, or inverse shading. Since many factors are involved in image formation this is a highly underconstrained inverse problem, and is precisely the domain of computer vision. Another difference between reshading and rendering is instead of generating a new image from scratch one is calculating modifications to an image's appearance, which can be much more difficult (consider removing a reflection from a mirror or adding light to a dark region). Finally, there are a number of additional complications when working with real images, such as compensating for image noise and properties of the imaging device, dealing with visual effects that are not adequately described by computer graphics models, and the great complexity of detail in the real world.

### 1.1 Aims and Contributions of this Work

This work examines two closely related aspects of image reshading. The first is the detection and removal of specular highlights from image data alone. Previous work in computational vision is reviewed and applied to visually textured images typically used in computer graphics. Weaknesses in current approaches is discussed. The second area examined is to solve for some of the scene properties that will allow an image to be reproduced, apparently under different illumination conditions. In this task the frontal geometry of the source scene is assumed to be known, and the cases in which the light source geometry is both known and unknown are explored. In the unknown case, a strategy for determining the geometric and radiometric parameters of a single rectangular area source that illuminates the image is presented. It is based on minimizing an objective function relating the unknown light source parameters to the error in rerendering the source image. Many constraints on potential solutions are available from the shading of diffuse and specular surfaces in the input image; finding and exploiting these shading cues is discussed. Finally, assuming that the lighting geometry is known beforehand, a simple method for solving for the relative light source emissivities and object reflectance properties in terms of a simple reflection model is discussed. Determining these parameters allows many interesting reshading operations to be performed on the input image, such as moving or dimming a light source, and changing the reflectance properties of surfaces.

### 1.2 Thesis at a Glance

This chapter has briefly motivated the problem of image reshading. Chapter two contains some background information on the reflection models and terms used in this work. Previous work in computational vision regarding the problem of specular highlight detection and removal in covered in Chapter three, along with a brief evaluation of two of the more promising approaches applied to real images.

The fourth chapter describes the problem of reshading images when the facing geometry of the objects in the image is known. A minimization strategy utilizing constraints on the shading found in an input image is presented, with the goal of locating a single rectangular source that lights the scene. Finally the case of when the lighting geometry is known beforehand is covered. Chapter five summarizes conclusions and future work.

## Chapter Two

## A Little Reflection Background



This chapter covers some of the basics of reflection useful for an understanding of shading simulation and analysis. This material in no way remotely approaches a complete treatment of the subject but hopefully gives some minimum coverage so assumptions made by this and other works can be better criticized. Of the many reflection models developed, only those used by work discussed in this thesis are presented.

### 2.1 Definitions and Terms

Reflection is the process by which light flux (energy per unit time) is redirected from an object into the incident side without change in frequency [50,1]. The simplest form of reflection is specular (or regular) reflection. This is the "angle of reflection equals the angle of incidence" law that everyone learns in high school: in three dimensions, the angle of reflection is equal to the angle of incidence and lies in the plane
formed by the normal to the surface and the incident ray ${ }^{\dagger}$. In formula ${ }^{\S}$,

$$
L_{e}\left(\theta_{e}, \phi_{e}\right)=\left\{\begin{array}{cl}
k L_{i}\left(\theta_{i}, \phi_{i}\right) & \text { if } \theta_{e}=\theta_{i}, \phi_{e}=\phi_{i}+\pi \\
0 & \text { otherwise }
\end{array}\right.
$$

with $0 \leq k \leq 1$ accounting for the reflectance of the surface. $L$ is the radiance of the transported light, the subscript $i$ denotes an incident quantity and $e$ denotes an exitant quantity. Local reflection geometry is shown in Figure 2.2, and Appendix A summarizes the notation used. A good description of radiometric quantities and their importance in rendering is given by Hanrahan [22].

Specular reflection occurs at a material boundary. The percentage of light incident from air at an angle $\theta_{i}$ that is reflected specularly by a smooth surface can be derived directly from Maxwell's equations relating electric and magnetic fields, and is given in Appendix B. The Fresnel equations (B. 1 and B.2) give the ratio of the total radiant flux reflected from a smooth surface to the radiant flux incident upon the surface from a given direction. The classic real-world example of a specular reflector is a mirror.

The wide angular redistribution of light into the incident hemisphere is referred to as diffuse reflection. Diffuse reflection arises from several processes, including specular reflection from a rough surface, subsurface scattering, and the absorption and subsequent reemission of light within the body of a material. An ideal or Lambertian diffuser uniformly scatters light according to Lambert's cosine law:

$$
L_{e}\left(\theta_{e}, \phi_{e}\right)=k L_{i}\left(\theta_{i}, \phi_{i}\right) \cos \theta_{i}, \quad 0 \leq k \leq 1, \quad 0 \leq \theta_{i} \leq \frac{\pi}{2}
$$

Since the apparent radiance reflected from an ideal diffuse reflector is constant throughout the surrounding hemisphere, the surface appears equally bright from all viewing directions.

Few materials are entirely diffuse or specular reflectors, and none are ideal. Diffuse reflectors often exhibit directionality, and most real surfaces contain imperfections which cause specular reflection to spread out over a number of viewing angles. There are also macroscopic effects that are neither diffuse nor specular

[^4]in appearance. In general, reflection from an object of a given temperature and pressure is a function of several variables: the wavelength and polarization of the light, surface profile, the direction of incidence, and the direction of exitance. Reflection from a material can be expressed in the form of a bidirectional reflection distribution function (BRDF), which relates irradiance falling upon a surface from one direction to the radiance reflected into another direction.
$$
f_{r}\left(\theta_{i}, \phi_{i}, \theta_{e}, \phi_{e}\right)=\frac{L_{e}}{E}=\frac{L_{e}\left(\theta_{e}, \phi_{e}\right)}{L_{i}\left(\theta_{i}, \phi_{i}\right) \cos \theta_{i} d \omega_{i}}
$$

In practice, since there are few analytic forms describing reflection from real surfaces and few reflectance datasets that are complete, BRDFs are commonly represented by reflection models which are designed to be flexible enough to accurately describe wide classes of materials without being overly expensive to evaluate. These can be theoretical (derived from physical laws), empirical (coming close to fitting measured data), or suppositional (neither derived from a physical basis not shown to fit real data). The vast majority of the numerous models in machine vision and computer graphics represent reflectance as a linear combination of variants to the two ideal reflectors. How to weight the two components to best approximate a given material and what form these components should take on is often unclear, since the parameters for these models may not correspond to any physically measurable quantity of the object. It is difficult to evaluate reflection models experimentally since reflectance data is scarce and difficult to collect. Simpler models exploit trends in the behaviour of many materials, while more complex models can portray more effects but are harder to apply to inverse shading problems due to their larger number of parameters.

Although many factors are involved in reflection, some characteristics are common across a large number of materials, and it is useful to make a distinction between these categories. Optically homogeneous materials have, for a given wavelength, a constant index of refraction throughout the material. These can be further divided into conducting and dielectric materials. Conductors contain many free electrons which scatter light equally at all wavelengths, and have electrons that exist only in specific energy zones (Brillouin zones). Incident light whose energy lies outside of the allowed energy zones of the material will be absorbed and mostly converted into heat, but above a certain critical wavelength the energy is not absorbed. The spectrum of the reflected light is influenced by this selective absorption process. Reflection from conductors is predominantly specular, and the more conductive the material, the more light it reflects. Dielectric or insulating materials do not have many free electrons and are poor conductors of electricity. Light is able to travel further into the body of the material, and is only slightly absorbed and scattered. As a result dielectrics are largely transparent, reflecting little, and what light is reflected is the same colour of the incident light. Metals and glass are examples of conductors and dielectrics, respectively [62].


Figure 2.1: Model of light reflection at an inhomogeneous surface.

Optically inhomogeneous materials are composites of homogeneous materials. They are commonly characterized as being composed of a substrate or medium comprising most of the mass of the object that is embedded with pigments of a different index of refraction (Figure 2.1). The specular reflectance properties are that of the dielectric substrate, which does not affect the spectral composition of the reflected light. Inhomogeneous materials exhibit significant amounts of diffuse reflection caused by the internal scattering of light as it strikes the pigments. Selective absorption of light passing through the pigments is responsible for the colour noticeable in the diffusely reflected light. Inhomogeneous materials make up the vast majority of the materials we see in everyday life. Paint, plastics, and ceramics are common examples [20, 28].

### 2.2 General Framework of Reflection

For objects that do not emit light the equation relating the light energy incident upon a surface area from the surrounding hemisphere to the exitant radiance from that surface is [26]

$$
L_{e}\left(\theta_{e}, \phi_{e}\right)=\int_{-\pi}^{\pi} \int_{0}^{\frac{\pi}{2}} L_{i}\left(\theta_{i}, \phi_{i}\right) f_{r}\left(\theta_{i}, \phi_{i}, \theta_{e}, \phi_{e}\right) \cos \theta_{i} \sin \theta_{i} d \theta_{i} d \phi_{i}
$$

where $f_{r}(\cdot)$ is the BRDF of the viewed surface, $\theta_{i}, \phi_{i}, \theta_{e}, \phi_{e}$ are the incident and exitance angles (hereafter represented as $\angle$ ) at the surface of area $d A, \lambda$ denotes wavelength and $L(\cdot)$ is the radiance passing through the solid angle about the specified direction.


Figure 2.2: Local coordinate system for integrating over the surrounding hemisphere.

In local illumination only the light energy from light sources is considered when computing the shading at a surface element, with the contribution of other reflecting objects accounted for in an ambient term. Direct illumination is accounted for by replacing the integral over the surrounding hemisphere with a summation over the light sources. To do this we must relate the incident radiance $L_{i}$ to that of the outgoing radiance from the source $L_{\ell}$ [22]

$$
L_{i}\left(\mathbf{x}_{r}, \vec{\omega}_{\ell}\right)=L_{\ell}\left(\mathbf{x}_{\ell}, \vec{\omega}_{r}\right) V\left(\mathbf{x}_{r}, \mathbf{x}_{\ell}\right)
$$

where
$\mathbf{x}_{\ell} \quad$ position of light source
$\mathbf{x}_{r} \quad$ position of receiver
$\vec{\omega}_{r} \quad$ vector from light emitter to the receiving point $\mathbf{x}_{r}$
$\vec{\omega}_{\ell}$ vector from receiving point to the light emitting point $\mathbf{x}_{\ell}$
$V$ unitless visibility term (either 0 or 1)
which follows from the conservation of energy within a thin pencil of light. This leads to

$$
\begin{equation*}
L\left(\mathbf{x}_{r}, \vec{\omega}_{t}\right)=\sum_{\ell=1}^{n}\left(L_{\ell}\left(\mathbf{x}_{\ell}, \vec{\omega}_{r}\right) \int_{A_{\ell}} f_{r}\left(\mathbf{x}_{r}, \vec{\omega}_{\ell}, \vec{\omega}_{t}\right) G\left(\mathbf{x}_{r}, \mathbf{x}_{\ell}\right) V\left(\mathbf{x}_{r}, \mathbf{x}_{\ell}\right) d A_{\ell}\right) \tag{2.1}
\end{equation*}
$$

with

$$
G\left(\mathbf{x}_{r}, \mathbf{x}_{\ell}\right)=G\left(\mathbf{x}_{\ell}, \mathbf{x}_{r}\right)=\frac{\cos \theta_{r} \cos \theta_{\ell}}{\left|\mathbf{x}_{\ell}-\mathbf{x}_{r}\right|^{2}}
$$

and the integration is now performed over the surface area $A_{\ell}$ of the $n$ light sources. Since all radiances are now outgoing, the subscripts $i$ and $e$ will be dropped. $L_{\ell}$ is outside the integral since it is assumed that all lights emit diffusely, which means radiance is constant over the emitting angles. $\theta_{\ell}$ and $\theta_{r}$ are the angles between the respective normals and the vector between $\mathrm{x}_{\ell}$ and $\mathrm{x}_{r}$.


Figure 2.3: Notation for three point transport.

Finally, most models in computer graphics represent reflection from an object as a sum of specular and diffuse components, which reduces the $\operatorname{BRDF} f_{r}(\cdot)$ to the form $k_{s} f_{s}(\cdot)+k_{d} f_{d}(\cdot)$ where $f_{s}(\cdot)$ and $f_{d}(\cdot)$ are the specular and diffuse BRDFs and $k_{s}, k_{d}$ are the specular and diffuse weighting factors multiplied by the albedo of the surface, $k_{s}+k_{d} \leq 1$. The amount $1-k_{s}-k_{d}$ is the amount of light absorbed by the surface. This results in

$$
L\left(\mathbf{x}_{r}, \vec{\omega}_{t}\right)=\sum_{\ell=1}^{n}\left(L_{\ell}\left(\mathbf{x}_{\ell}, \vec{\omega}_{r}\right) \int_{A_{\ell}}\left(k_{d} f_{d}\left(\mathbf{x}_{r}, \vec{\omega}_{\ell}, \vec{\omega}_{t}\right)+k_{s} f_{s}\left(\mathbf{x}_{r}, \vec{\omega}_{\ell}, \vec{\omega}_{t}\right)\right) G\left(\mathbf{x}_{r}, \mathbf{x}_{\ell}\right) V\left(\mathbf{x}_{r}, \mathbf{x}_{\ell}\right) d A_{\ell}\right)
$$

If we assume that the surfaces under examination exhibit Lambertian reflection, then $f_{d}=1 / \pi$ and $f_{s}=0$, which leaves us with

$$
\begin{equation*}
L\left(\mathbf{x}_{r}, \vec{\omega}_{t}\right)=\frac{k_{d}}{\pi} \sum_{\ell=1}^{n}\left(L_{\ell}\left(\mathbf{x}_{\ell}, \vec{\omega}_{r}\right) \int_{A_{\ell}} G\left(\mathbf{x}_{r}, \mathbf{x}_{\ell}\right) V\left(\mathbf{x}_{r}, \mathbf{x}_{\ell}\right) d A_{\ell}\right) \tag{2.2}
\end{equation*}
$$

Until now we have been treating the receiving area as a single point, but in a discrete environment or for pixel values we must deal in terms of the surface area surrounding this point:

$$
L\left(\mathbf{x}_{r}, \vec{\omega}_{t}\right)=\frac{k_{d}}{\pi} \sum_{\ell=1}^{n}\left(L_{\ell}\left(\mathbf{x}_{\ell}, \vec{\omega}_{r}\right) \frac{1}{A_{r}} \int_{A_{r}} \int_{A_{\ell}} G\left(\mathbf{x}_{r}, \mathbf{x}_{\ell}\right) V\left(\mathbf{x}_{r}, \mathbf{x}_{\ell}\right) d A_{\ell} d A_{r}\right)
$$

Finally, we can group the purely geometric terms and constants together to obtain

$$
\begin{equation*}
L\left(\mathbf{x}_{r}, \vec{\omega}_{t}\right)=k_{d} \sum_{\ell=1}^{n} L_{\ell}\left(\mathbf{x}_{\ell}, \vec{\omega}_{r}\right) F_{A_{r} \rightarrow A_{\ell}} \tag{2.3}
\end{equation*}
$$

with

$$
\begin{equation*}
F_{A_{r} \rightarrow A_{\ell}}=\frac{1}{A_{r}} \int_{A_{r}} \int_{A_{\ell}} \frac{\cos \theta_{r} \cos \theta_{\ell} V\left(\mathbf{x}_{r}, \mathbf{x}_{\ell}\right)}{\pi\left|\mathbf{x}_{\ell}-\mathbf{x}_{r}\right|^{2}} d A_{\ell} d A_{r} \tag{2.4}
\end{equation*}
$$

$F_{A_{r} \rightarrow A_{\ell}}$ is the form factor and represents the amount of energy from $\mathbf{x}_{r}$ that $\mathbf{x}_{\ell}$ receives, averaged over the surface of the source. A closed form exists for the inner integral of equation (2.4) for the case where $A_{\ell}$ is a polygon, which is arrived at through application of Green's theorem:

$$
\begin{equation*}
F_{d A_{r} \rightarrow A_{\ell}}=\frac{1}{2 \pi} \sum_{i} \vec{\Gamma}_{i} \cdot \overrightarrow{\mathbf{N}}_{d A_{r}} \tag{2.5}
\end{equation*}
$$


where $\overrightarrow{\mathrm{N}}_{d A_{r}}$ is the unit length vector normal to the infinitesimal surface area $d A_{r}, \gamma_{i}$ is the angle in radians between vertex $i, \mathbf{x}_{r}$, and vertex $i+1$ of the source, and $\vec{\Gamma}_{i}$ is the vector of magnitude $\gamma_{i}$ normal to the face* defined by vertex $i, \mathbf{x}_{r}$, and vertex $i+1$ of the source (see figure). Hottel gives a clear and complete derivation [27]. The outer integral can be evaluated numerically.

[^5]
### 2.3 Some Models of Light Reflection

One of the most widely used models is due to Phong [52]. It describes reflection at a surface as a sum of ambient, diffuse, and specular components:

$$
\begin{gathered}
L\left(\theta_{e}, \phi_{e}, \lambda\right)=k_{\mathrm{amb}} L_{\mathrm{amb}}\left(\theta_{i}, \phi_{i}, \lambda\right)+\sum_{\ell} L_{\ell}\left(\theta_{i}, \phi_{i}, \lambda\right) f_{r}(\lambda, L) \\
f_{r}(\lambda, \angle)=\left(\rho_{d} \cos \theta_{i}+\rho_{s} \cos ^{n} \varphi\right)
\end{gathered}
$$

where $\varphi$ is the angle between the reflection vector $\left(\theta_{e}, \phi_{e}\right)$ and the direction of ideal specular reflection $\left(\theta_{i}, \phi_{i}+\pi\right)$. The $\cos ^{n} \varphi$ term is meant to account for the spread of specular highlights due to surface roughness, and allows specularities (which are images of the light sources) formed by point and directional sources (all that was used in Phong's day) to appear as though they were formed by area sources, as realistic highlights are. Although the Phong model is very popular in the computer graphics community it is not physically plausible and has not been shown to fit reflectance data well, which makes it of questionable value for use in shading analysis [73]. Physically plausible reflectance functions must conserve energy (not reflect more energy than is received), be nonnegative over all directions, and must obey Helmholtz's principle of reciprocity (the reflectance properties for a given reflection geometry are independent of the direction that the light flows). Phong's model can be modified to be physically plausible [41].

Torrance and Sparrow proposed a model to explain reflection from rough surfaces [67] (later adopted by Blinn [7] and Cook and Torrance [12]). They model a rough surface as a collection of smooth microfacets of constant area and varying orientation. Neighbouring facets can block both incoming and outgoing light, and a surface's roughness dictates the degree to which microfacet orientations vary from the macroscopic surface normal. Reflectance is modelled as a sum of Lambertian and specular reflection, with the specular contribution given by
where

$$
f_{s}(\lambda, L)=\frac{A d \omega_{i}}{4} \frac{F\left(\theta_{i}^{\prime}, \bar{n}(\lambda)\right) G(\angle) P(\alpha)}{\cos \theta_{e}}
$$

$A$ is the area of each facet
$F(\cdot)$ is Fresnel equation (B.1) expressing percentage of light flux reflected specularly
$\theta_{i}^{\prime} \quad$ angle between light beam and facet normal
$G(\cdot)$ Geometric attenuation factor representing percentage of light flux not blocked
$P(\cdot)$ Probability function giving percentage of facets at given angle
$\alpha$ angle between macroscopic normal and facet normal


Figure 2.4: Local geometry for the (left) Original Phong shading model (right) Torrance-Sparrow shading model.

This model successfully explains the phenomena of off-specular peaks, or maxima in reflected radiance from rough surfaces at angles greater than the specular angle $\theta_{e}=\theta_{i}$, and predicted reflectance distributions using it have been shown to closely fit measured data. Although the model includes a Fresnel term, it too is based upon geometric optics. It is physically plausible [41], although its moderate complexity makes it unwieldy for use in computer vision, and is used under simplifying assumptions.

Shafer [60] formulated a theory of reflection for opaque inhomogeneous materials that he named the Dichromatic Reflection (DR) model:

$$
L(\lambda, \angle)=L_{\ell}(\lambda, \angle)\left(\rho_{d}(\lambda) g_{d}(\angle)+\rho_{s}(\lambda) g_{s}(\angle)\right) .
$$

where $\rho_{*}(\cdot)$ is a reflectivity term and $g_{*}(\cdot)$ is a geometric scale factor. The notable feature of Shafer's model is it represents reflection as being separable in geometry and wavelength. Shafer noted that the distribution of pixel values corresponding to an object of constant diffuse reflectance, plotted in RGB space, should form a parallelogram with the arms defined by the colours $\rho_{d}(\lambda)$ and $\rho_{s}(\lambda)$. This property holds in any colour space that encodes luminance as a linear dimension, which is true for many popular spaces (RGB, CIE, YIQ, CMY all hold this property). Lee narrowed the DR model with the additional assumption that
specular reflectance is equal for all wavelengths of light [37]:

$$
L(\lambda, \angle)=L_{\ell}(\lambda, \angle)\left(\rho_{d}(\lambda) g_{d}(\angle)+\rho_{s} g_{s}(\angle)\right)
$$

Due to its modelling of $\rho_{s}$ constant for all wavelengths it is referred to as the Neutral Interface Reflectance (NIR) model. It follows from this assumption that the light reflected specularly has the same spectral composition as the incident light. This property is key to several algorithms for determining the chromaticity of the illuminant from image shading.


Figure 2.5: Distribution of pixel values in colour space following the NIR model.

The important assumption that both these models make is that the spectral composition of the light reflected specularly and diffusely is independent of geometry. This assumption is in opposition to the Fresnel equations and therefore these models do not account for materials that exhibit significant colour shift over a range of viewing and illuminating angles. However, for dielectric materials the Fresnel equations only depend upon wavelength through the index of refraction, which Shafer argues is nearly constant for inhomogeneous materials (whose substrate is typically a dielectric) in the visible spectrum. Establishing the aptness of this assumption may be an open problem, for as Nicodemus et al. remarked several years earlier in discussing separability of the BRDF, "We are not aware of any data that will establish the extent to which there may be interaction between geometrical dependance and spectral dependance, except for the knowledge that it is a significant factor in some internal-scattering situations..." ([50], p. 124). Figure 2.6 (a) shows the hemispherical reflectance curve for titanium for several angles of incidence. The curves
are not scaled versions of each other and the largest differences appear in the visible spectrum (although titanium is not an inhomogeneous material and therefore the DR model does not claim to cover it). It is worth noting that the Phong model can be formulated in terms of the DR model, but the Torrance-Sparrow model cannot be.


Figure 2.6: Angular spectral reflectance at varying $\theta_{i}$ for (left) titanium and (right) potassium bromide. For titanium note the changes in the visible spectrum and that the curves are not scaler multiples of each other (data obtained from [68, 69]).

An analysis of the Dichromatic reflection model was carried out by Healey, who concluded that it was a reasonable model for many inhomogeneous dielectrics [24]. Healey noted, as did Shafer, that the index of refraction for dielectrics varies little over the visible spectrum, although no figures were given in either case to support this claim. Based on this observation he concluded that the separation of geometry and wavelength terms for specular reflection was sound, and concentrated on examining diffuse reflection. In this vein, Healey sought two functions $F_{1}(\lambda)$ and $F_{2}(\angle)$ that minimized the error measure

$$
\text { Error }=\sum_{\substack{\theta_{i}=0 \\ \theta_{i} \text { even }}}^{\pi / 2} \sum_{\substack{\lambda=400 \\ \lambda \bmod 10=0}}^{700}\left(f_{d}(\lambda, \angle)-F_{1}(\lambda) F_{2}(\angle)\right)^{2}
$$

where $f_{d}$ is a model for diffuse reflectance stemming from subsurface scattering for inhomogeneous materials as developed by Reichman ${ }^{\dagger}$ [58]. This model is not separable in wavelength and geometry and has been shown to give good agreement with more complicated models and experimental data [15]. The experiment

[^6]was performed on 25 simulated inhomogeneous materials, whose reflectance properties were calculated at increments of $2^{\circ}$ of polar angle and at 10 nm intervals in the visible spectrum. Each material was constructed by randomly selecting a spectral reflectance from the set of Munsell colour chip samples and taking that to be its spectral response at normal incidence, and assuming a constant index of refraction of 1.5. $F_{1}$ and $F_{2}$ were computed through principal component analysis of the matrix of $f_{d}$ values. For all 25 samples the variance in the simulated $f_{d}$ accounted for by the $F_{1} F_{2}$ model was in excess of $99 \%$. This high number suggests that the Dichromatic model is a good one if the Munsell samples are representative of real materials and if the diffuse reflectance model chosen describes real materials well.

The validity of the Lee's NIR model has been tested experimentally on a number of real objects [66, 39]. Tominaga and Wandell analyzed the spectral power distributions of four inhomogeneous objects (a red plastic cup, a green plastic ashtray, an apple and a lemon) to see if they were well described by a plane in colour space and if the intersection of two or more planes corresponding to objects illuminated by the same source gave an accurate estimate of the colour of the illuminant. Reflected light was measured every 10 nm at between eight and ten places on the objects under conditions of no interreflection. The principal components of the normalized data were determined through SV decomposition, and they found that the variance accounted for by first two vectors was over $99.9 \%$ for each object, and that the estimated illuminant colour was a good estimate. Their experiments did not give an exhaustive coverage of incident and viewing angles. Lee et al. ran similar experiments on 8 real objects and concluded that the NIR model is accurate for some materials (cloth, plastic, wood) but not for others (coloured paper).

It is important to remember that Lambertian reflection, used by virtually every model in computer graphics and vision, is only an approximation to how diffusely reflecting materials behave. Despite its popularity, there have been few studies on the aptness of the Lambertian model, which seems to be taken largely on faith in many cases. One of the causes of diffuse reflection is scattering within the body of the material-an encounter between a photon and another particle which can result in the redirection of the photon and a change in its energy [62]. Scattering is an important and complicated process that depends upon numerous factors, including size, shape, and cross-section of the particles (see Figure 3.2 of [28] for an good example of the dependance of scattering power on particle size). Given an arbitrary distribution of particles of unknown
tering particle is scattered equally in one of two directions, either coincident or directly opposed to the incident beam (two flux approximation). The instantiation of Reichman's theory for the case of an isotropic scattering material is

$$
f_{d}(\lambda, \angle)=\left(1-f_{s}(\angle)\right) \frac{A\left(\theta_{i}, \lambda\right) k_{1}\left(B(\lambda)-C\left(\theta_{i}\right)\right)}{\cos \theta_{i}\left(2-k_{2} B(\lambda)\right)}
$$

$f_{d}$ is defined as a ratio to a Lambertian surface, and in the Lambertion case reduces to $f_{d}=1$.
size and composition within a substrate, there seems at first glance to be little reason to believe that the profile of reflected flux, viewed macroscopically, will appear Lambertian. This skepticism is corroborated by theoretical results and measured data. A somewhat useful exercise is to examine the theoretical behaviour of an isotropically scattering material through the equations of transfer, for which quite accurate solutions have been obtained. Figure 2.7 shows the ratio of reflectance from an isotropically scattering semi-infinite slab to that of a Lambertian diffuser for several angles of incidence and exitance [27]. An ideal diffuser would be represented by a horizontal line; visually, it is a fair approximation for many reflectance geometries, although the deviation from the Lambertian case is apparent with increasing incidence angle. This has stirred work on alternate models for diffuse reflection. Wolff extends the ideas of a rough surface being composed of tiny microfacets by assuming the subsurface beneath each microfacet acts as a isotropically scattering slab [76].


Figure 2.7: Theoretical derivation of the ratio of reflectance from an isotropic scatterer to that from a lambertion diffuser for different angles of incidence and exitance, illumination from a directional source, and scattering accounting for $90 \%$ of the combined effects of scattering and absorption (adapted from [27], pp. 145).

The importance of all of this is that in detecting specular highlights, classifying anything that doesn't fit some notion of what diffuse reflection should be as specular will lead to errors if the diffuse model used is
inaccurate. Looking explicitly for specular characteristics will avoid this problem.

### 2.4 Sensor Model

Until now reflection has been dealt with in terms of radiance impinging upon and leaving a surface. To perform an analysis of shading, one must relate the light energy leaving a surface to the pixel values in an image formed by a camera. The radiance leaving a surface patch is related to its image irradiance by (see Appendix C):

$$
E(x, y)=\frac{\pi}{4}\left(\frac{d}{f}\right)^{2} \cos ^{4} \alpha L_{e}\left(\theta_{e}, \phi_{e}\right)
$$

where $E(x, y)$ is the image irradiance at pixel $P(x, y), d$ is the diameter of the camera lens, $f$ is the focal length of the lens, $\alpha$ is the angle between the optical axis and the vector from the center of the lens to the center of the area being viewed, and $L_{e}(\cdot)$ is the exitant radiance from the surface in the direction of the lens. We want to go the other direction, that is, we wish to relate the image irradiance at a pixel to the scene radiance. If the pixel is formed by the projection of a single object that is not 'too large', then the above relationship holds. The value at a pixel is somehow related to the irradiance at a position inside the camera's sensor; this relationship is assumed to be linear.

$$
P(x, y)=k_{1} E(x, y)+k_{2} .
$$

There are several reasons why this relationship may not be linear. Physical limitations of the camera (chromatic aberration, limited dynamic range resulting in colour clipping and blooming) cause nonlinearities in response, as does the electronics built into the camera (automatic intensity histogramming). These factors can be measured and accounted for through a proper camera calibration. Novak has addressed many of these issues [51].

## Chapter Three

## Specular Highlight Analysis



We have had "real" and "imaginary" axes explained in accepted terms. So it is but a short step to apply the same logic to our "material" and "spiritual" axes, the Argand Plane representing the membrane between these two realms.

Pat Delgado<br>Crop Circles: Conclusive Evidence?

Developing image processing tools that operate on shading cues rather than a collection of pixels necessitates ingraining these tools with an understanding of the physical processes behind these cues. This chapter addresses the detection and removal of one such cue-specular highlights-from image data alone. Specular highlights are worthy of interest for two reasons. First, highlights are very noticeable and can lead to glaringly (no pun intended) inconsistent results when the image is combined with other elements or used as a texture. An image processing or paint system that could remove specularities would be useful for touching up texture images. Second, since specular reflection is a relatively simple process, highlights can give clues about scene properties and thus are a valuable aid in inverse shading. Detecting where specularities lie in an image is the first step to making use of such cues.

### 3.1 Previous Work in Adding and Removing Shading Cues from Images

The idea of using portions of existing images to influence the shading of new images heralds back to the golden days of computer graphics. Texture mapping is a general technique for altering the surface characteristics of geometric objects according to some function assigned over the surface [9, 6]. When rendering, each point on an object's surface is influenced by the function mapped onto the surface evaluated at that point. This function can be a real image of a visually textured object, and by setting an object's diffuse reflectance properties in accordance with an image whose shading gradations are a result of changes in albedo alone (say a wood grain), the object can be imbued with a realistic looking visual texture. The texture only defines one parameter in the shading computation for the surface, so other cues can be added to the image, such as a shadow being cast on part of the object or having a specularity pop up. The parameters to produce these cues must be specified by the user, though, and therefore may be incorrect, and to be accurate the geometry of what the texture represents must be modelled. There is no provision for removing shading artifacts from a texture image, so any present in the image will be taken into the approximation of reflectance as well and can lead to conflicting visual cues. Figure 3.1 shows a simple example of altering the appearance of a synthetic object in this manner. The advantage of texture mapping is that it gives the appearance of surface detail without requiring that detail to be modelled. Of course, since a mapped texture has no substance, it cannot cast a shadow or reflect light as the object it portrays would.

A well established postprocessing technique for overlaying or blending many images into one is image compositing. Wallace [70] introduced a compositing system for merging multiple digital cartoon images in an order independent fashion, as long as a relative depth ordering is maintained (that is, the operations were associative but not commutative). Using this technique, a depth ordering must be established between the images to be combined and every pixel assigned an opacity value. This opacity value equals the percentage of colour that a pixel contributes to the result, with the remaining percentage coming from the corresponding pixels in images that are behind it. Porter and Duff [54] and later Duff [14] reformulated Wallace's work in terms of an alpha and depth channel, and introduced a set of compositing operators which come at the expense of associativity. Additional work in compositing has since followed [47]. Standard compositing cannot facilitate lighting or shading interactions between objects from different images, nor does it insure consistency of illumination between the separate components. This can lead to images in which objects appear to "float" on top of the background or have shading inconsistencies, causing the illusion of three dimensionality to be lost. Furthermore, the user has no means for changing the appearance of the source images to compensate for potential inconsistencies, and segmenting these images can be


Figure 3.1: An example of texture mapping.
laborious. Compositing's film-based cousin blue-screening works on similar principles, with the relative depth ordering determined by filming in front of a background of a known colour which indicates positions where the images behind it should be shown. One nice trick for capturing lighting interactions between components from different images is to take the luminance channel from the blue-screened portion and use it to alter the luminance channel of the occluding image. In this way shadows cast upon the occluded region can be transferred to the final image, and the effect can be quite convincing if the blue screen scene and the occluding viewed scene have similar geometry in the shadow region.

Painting systems are widely used to alter images. These offer the user the greatest amount of control but no constraints to ensure that a plausible result will be produced-it rests solely in the skill of the user. These systems usually offer primitive operations which can be used to produce simple effects like drop shadows or softening of shadows.

Shading discontinuities formed by shadows are prevalent, very noticeable, and give many important clues as to light source position and emissive strength. Like specular highlights, shadow detection and extraction has been studied in computer vision for years, albeit lightly. Most previous work deals with aerial photographs as shadows throw off missile guidance systems. Shadow analysis is closely related to specularity analysis as to remove a shadow, light from the blocked illuminant must be added to a region. To be done properly,
the chromaticity and strength of the illuminant must be determined, and corresponding shadow/non-shadow regions can give clues to these properties. Shadow detection and removal has been carried out through thresholding $[30,71]$ and by edge detection $[75,61]$.

Although this work deals solely with illumination, the issue of establishing common viewing parameters must be resolved when several image sources are combined. This problem has been traditionally addressed in photometry, computer vision, and more recently human-computer interaction where there is a herd of researchers exploring ways in which computers can usefully supplement a user's interactions with the real world. In the past their work has been classified as virtual reality, ubiquitous computing, and computer augmented reality. Typically work in these areas is concerned with augmenting a user's view of the world in real time, so performing a fast and accurate registration between real and virtual viewpoints is of critical importance, while issues of common illumination and portraying interactions between the real and synthetic components (which are usually kept quite simple) are not. A sampling of recent work is [4, 3].

### 3.2 What is a Specular Highlight?

What exactly are specular highlights, anyway? Hunter ${ }^{\dagger}$ dubs the "attribute of surfaces that causes them to have a shiny or lustrous appearance" as gloss, and goes on to distinguish between six types of gloss ([28], p. 75). These are:

- Specular gloss: characterized by shininess of reflection at the specular angle
- Absence-of-bloom gloss: characterized by a hazy appearance adjacent to specular highlights (bloom: the scattering of light by a deposit that can be wiped away)
- Distinctness-of-image gloss: measures the sharpness of mirror images
- Surface uniformity gloss: measures the nonuniformity of reflection from a surface which is not a function of the reflectance (such as texture)
- Sheen: characterized by shininess at grazing angles
- Luster: characterized by high contrast between the specularly reflected light and the surrounding, non-specular areas

[^7]These categories are intended for specimens observed under laboratory conditions. In terms of modelling or detecting specularities, some of the distinctions between these categories are unimportant. The difference between specular gloss and sheen is one of geometry, and surface uniformity gloss is one of the other five types at a different scale. Differences within categories may be more important than those between categories, as for example how a surface rates in the measures of absence-of-bloom and distinctness-of-image gloss reflects its nature of producing diffused, hazy highlights or sharp highlights.

In terms of highlight analysis on everyday objects, several characteristics suggest themselves as being useful:

- sharply delineated vs. spread out highlights
- mirror image vs. hazy highlights
- texture, or variance of $\rho_{d}$ on object and in highlight region
- variance of $\rho_{s}$ in highlight region
- highlights formed by multiple illuminants of varying colour and intensity
- reasons why a highlight ends: abrupt changes in surface orientation, light source blocked, highly reflective surface, etc.

Clearly there is not any one kind of specular highlight, and there cannot be one detection or removal strategy. In its most obstinate form a highlight is a mirror image of its surrounding environment, which clearly cannot be handled without some form of high-level knowledge not available in the image. The techniques presented in the next few sections deal with a small subset of imaginable highlights, some of which are shown in Figure 3.2.

Based on the physical underpinnings of reflection, a few observations and heuristics can made about specularities:

## 1. Brightness of specularities

- highlights tend to appear brighter than their surrounding non-specular regions. For a given point on a Lambertian surface the maximum reflected radiance is the incident radiance times $d \omega / \pi$, while for a specular surface of the same reflectance the same light energy is reflected in a more concentrated solid angle of directions


Figure 3.2: Some examples of specular highlights. (top left) Spread-out highlight on a textured surface, $\rho_{s}$ varies somewhat in highlight region (top right) Highlight whose boundary is formed on three sides by a change in surface orientation. $\rho_{s}$ is constant in highlight region. (bottom left) Highlight from a metallic surface. Neither of the two specularities exhibit colour shift (bottom right) mirror-like specular reflection, varying illuminant colour and intensity.
2. Chroma of specularities

- for composite materials, it is generally the chroma of the incident light, since the main material boundary with air is that of the dielectric substrate
- for homogeneous dielectrics, it is the chroma of the incident light
- for homogeneous conductors, it is the chroma of object, as governed by its absorptive band

Relatively few bare metallic surfaces appear in everyday life, though.
3. Geometry of specular reflection

- governed by the law of angle of reflection equals the angle of incidence in the plane formed by the surface normal and the incident beam, and is therefore a function of incident direction, surface normal, and viewing direction. At reasonable distances viewing direction varies slowly over an image, so a rapid falloff in specular reflection may indicate either a discontinuity in illumination direction or surface orientation or a mirror-like reflector.

Practical considerations, such as ambient light, camera blooming, and clipping of measured irradiance can render these observations false in any one case. Notwithstanding these rude intrusions of reality, the above-mentioned properties are exploited in many of the schemes presented in the following sections.

### 3.3 Detection of Specular Highlights

The goal of specularity detection is to identify those pixels whose value arises in some part from specular reflection. There are a few different approaches to this task. One class of approaches characterize specularities by distinctive "signatures" or geometric distributions in colour space. The alternate strategy is to define specularities not by what they are, but by what they are not: well-behaved diffuse reflectors. Under this approach constraints are formed based on a model of a well-behaved diffuse surface, and any image regions where these constraints are violated are potential specular highlight zones. These two strategies are presented below. Approaches that require additional information or constraints, such as operating on multiple images [40] or using polarized filters [49] are not covered.

### 3.3.1 Previous Work: Specular Highlights as Signatures in Colour Space

Lee used his NIR model to develop the chromaticity convergence method for calculating the (assumed unique) colour of the illuminant from the shading of inhomogeneous objects exhibiting specularities [37]. According to the NIR model, if there is one predominant illuminant colour in a scene (i.e. all lights are the same colour and there is no significant discolouration of light due to interreflection) then the chromaticity at any point on an object is a linear combination of the object colour and the illuminant colour. Since in the CIE chromaticity diagram any colour that can obtained by combining two other colours together lies on the line between those two colours, the line defined by the chromaticities of the light reflected from an object with a specularity should point towards the colour of the illuminant. Therefore the lines formed by plotting the colours of two or more objects of uniform and different diffuse reflectance in the CIE chromaticity diagram should intersect at the chromaticity coordinates of the illuminant, with a least-squares fit needed in the overconstrained case (see Figure 3.3). Unfortunately, given an arbitrary distribution of chromaticities from an image, it is very difficult to find these lines, especially without knowing how many lines should be fit. To alleviate this problem, Lee suggested a simple segmentation scheme of performing colour edge detection and comparing the ratio of the colour signals on both sides of edges to determine if the edge occurs due to a change in material. With the illuminant chromaticity known, specular highlights are located as those pixels lying on lines pointing towards the illuminant colour that contain some portion of the illuminant colour. Lee's only experiment was on a synthetic image containing spheres. Nevertheless, the flavour of Lee's approach lives on in much of the work that has followed in both illuminant colour estimation and specularity detection. This approach was later extended to the three dimensional CIE space by Gunawan after he applied it to real images and discovered that it did not produce useful results due to the irrelevant clustering of pixels in two dimensions [19].

Gershon and Klinker et al. both jump to three dimensional colour space and search for L-shaped distributions. Both assume a single light source chromaticity and require highlights to exhibit colour change in order to detect them. Gershon searches in a colour constant space which requires an estimate of the light source colour to get into [17]. He expresses the colour in an image as the product of three spectral distributions: that of the light, the reflectance of the viewed surfaces, and the response of the sensor:

$$
E(\lambda)=\operatorname{Sensor}(\lambda) \operatorname{Light}(\lambda) \operatorname{Reflectance}(\lambda)
$$

With knowledge of the sensor's spectral response the problem becomes one of disambiguating the light and surface reflectance spectra. He adds sufficient constraint to his model to compute the illuminant spectra


Figure 3.3: The chromaticity convergence method.
by approximating the average surface reflectance in an image with a standard surface reflectance and by selecting an 'average' colour value in an image. With the illuminant spectra computed Gershon transforms pixel values into a colour constant space ( $C$-space) by discounting the contribution of the illuminant colour and sensor sensitivity from the value at each pixel. Specular highlights in this space arise from a neutral reflector, and these are located by searching for a "dog-leg" distribution of values which mark a transition from an object's reflectance spectrum to that of a neutral reflector.

Gershon defines the 'standard' surface reflectance as the mean of a collection of reflectance spectra of naturally occurring materials as approximated by a small number of basis functions ${ }^{\dagger}$. The average colour

[^8]value in an image is calculated by segmenting the image into regions representing different materials, selecting a single pixel value from each region, and computing the unweighted mean of the samples. The segmentation was performed through a split-and-merge process using chromatic similarity as the criterion for merging. In discussing his segmentation scheme, Gershon noted "it creates many small regions which cannot be merged into bigger ones, particularly around areas where there is a change from one color (sic) to another, such as material changes or changes caused by highlights." ([17], page 118). The latter failing is serious if the whole point of the exercise is to match highlights with the object it appears on. Once the standard surface reflectance and average colour value are determined, the illuminant colour is calculated and each pixel is transformed into C -space by removing the contribution of the illuminant colour. The distribution of C -space values for each neighbouring segmented region in the image is examined by fitting a least squares line to each region. Several tests are performed to classify a pair of distributions as a dog leg, such as determining if the two lines actually intersect and whether one of the regions forms a line pointing to the region in C -space corresponding to the perfect reflector.

Gershon's experiments on real images show reasonable results. Errors of omission and inclusion are evident in tests of the highlight detection algorithm, which he attributes to inaccuracies in segmentation. Obviously, the success of his technique depends upon the quality of the assumption that the standard surface reflectance as calculated from the spectral data set reflects the average reflectance in the imaged scene. It is interesting to note that Gershon estimates the illuminant colour first and then locates specularities later as a separate process, while in other approaches the methods for satisfying these two objectives are intertwined.

Klinker et al. search for T-shaped planar distributions in RGB space in their attempts to decompose a colour image into diffuse and specular components [34,35]. Along with everyone else, they reason that if the NIR model is accurate, then for an object of constant diffuse reflectance the vectors $\rho_{s}(\lambda)$ and $\rho_{d}(\lambda)$ will span a dichromatic plane in a spectral vector space in which every measured colour from this object will lie. They note that the plot of pixel values forms a dense cluster in this plane and that the geometric distribution of the points conveys useful information ${ }^{\dagger}$ Measured colour values from an object which exhibit diffuse reflection only will form a matte line along the $\rho_{d}(\lambda)$ axis when plotted in the RGB colour space. Likewise, colour

[^9]values arising from both specular and diffuse reflection will form a highlight line. The degree to which the highlight line is parallel to the vector $\rho_{s}(\lambda)$ depends upon how sharp the specularity is, since variation in the amount diffuse reflection underneath a highlight causes the highlight line to spread out in colour space and a fitted vector to stray from the true $\rho_{s}(\lambda)$ axis. Figure 3.4 shows a hypothetical example of reflection for a monochrome cylinder illuminated by a point source.


Figure 3.4: The shape of the spectral cluster for a cylindrical object assuming a point light source (adopted from [35]).

Given an image region corresponding to a single object, Klinker et al. fit lines to the distribution of pixel values in RGB space. All lines having an endpoint in close proximity to the black corner of the colour cube are classified as matte lines, and the first line sprouting from the upper portion of a matte line is classified as a highlight line. Pixel values that lie near the highlight line are then marked as specular pixels. In the spirit of Lee, multiple $\rho_{s}(\lambda)$ vectors can be intersected to give an estimate of the illuminant colour. Klinker et al.'s tests on pristine real objects show good results in detection but that the estimate of the illuminant colour is prone to error for reasons discussed below. They address the problems of dealing with real cameras and methods for correcting for their limitations. Their technique can be adapted to detect highlights of different colours on the same object, but for any reasonably complex image the distribution of pixel values in colour space will likely be sufficiently dense and varied to give automatic line or plane fitting schemes fits. Textures will cause the pixel values from an object to not lie in a plane in colour space, requiring some form of accurate segmentation. Their later work presents an automatic segmentation scheme based on grouping regions together whose distribution in RGB space have similar principal components. Their algorithm chops the image up into small non-overlapping chunks, determines the principal components of
these regions and then grows those regions which are well described by one vector. These linear regions are considered either the matte lines or highlight lines predicted by the DR model.

### 3.3.2 Previous Work: Specular Highlights as Violations

By manipulating the NIR equations, Lee expresses the difference of two correctly scaled irradiance signals solely in terms of diffuse or specular reflection [38]. The implication is that if the proper scaling factors can be found, the shapes of the two reflection components can be recovered and separated. Happily, it turns out that the scaling factors used to determine the body reflection shape is the CIE chromaticity coordinates of the illuminant. Lee shows that

$$
\begin{gathered}
E_{\lambda}(x, y)-\left(\frac{L_{\lambda}}{L}\right) E(x, y)=k L_{\lambda}\left(\rho_{d}(\lambda)-\rho\right) g_{d}(x, y), \quad \lambda \in\{r, g, b\} \\
E=E_{r}+E_{g}+E_{b} \\
L=L_{r}+L_{g}+L_{b}, \\
\rho=\left(L_{r} \rho_{d}(r)+L_{g} \rho_{d}(g)+L_{b} \rho_{d}(b)\right) / L
\end{gathered}
$$

where $L$ indicates light source radiance and $E$ indicates image irradiance (assumed to be equally linearly proportional to pixel values in all channels). Note that the scale factors $\left(\frac{L_{r}}{L}, \frac{L_{g}}{L}, \frac{L_{b}}{L}\right)$ are the chromaticity coordinates of the illuminant colour. The problem is thus one of how to find the right scaling factors $\left(s_{r}, s_{g}, s_{b}\right)$. Lee chose scaling factors that produce the smoothest function for $g_{d}(x, y)$, and justifies this criteria by noting that diffuse reflection is generally smoother than specular reflection, which usually occurs at a smaller scale in an image. Also, since peaks in diffuse and specular reflection do not in general occur at the same position on an object, any mixing of diffuse reflection with specular reflection will produce an irradiance curve that is less smooth than an irradiance curve arising from diffuse reflection alone. Therefore, if an improper $s_{\lambda}$ is selected, $E_{\lambda}-s_{\lambda} E$ will not be as smooth as in the case when $s_{\lambda}$ is selected properly.

Lee chose the Laplacian operator to define smoothness. The proper scaling factors were subsequently chosen such that

$$
\begin{aligned}
& \sum_{x} \sum_{y}\left[\nabla^{2}\left(E_{r}(x, y)-s_{r} E(x, y)\right)\right]^{2} \\
& \sum_{x} \sum_{y}\left[\nabla^{2}\left(E_{g}(x, y)-s_{g} E(x, y)\right)\right]^{2}
\end{aligned}
$$



Figure 3.5: Specularity detection ala Klinker et al. (top left) Original image (top right) Core detected area before pixel growing (bottom) RGB plot of pixels outlined region and its first two principal axes, viewing direction orthogonal to these. Note that these axes must be orthogonal and do not define $\rho_{s}$ and $\rho_{d}$. Also note clipped pixel values at $R=255$ causing a colour shift in the image that is not present when viewed with the eye.
were minimized. One disadvantage of this technique is that the Laplacian operator is likely to heighten shading and colour discontinuities in the input, making an accurate segmentation beforehand important (although this is a problem for signature searching approaches as well since line fitting algorithms are sensitive to outliers). Also, since Lee defines smoothness over the image plane, his technique is not viewpoint-invariant (even though diffuse reflection is) and must assume smooth geometry on the surface. Detection of specularities and separation of the diffuse and specular components follows the chromaticity convergence approach: the endpoint of the line in CIE space that is farthest away from the estimated illuminant colour is taken as the diffuse colour, with all others on a line starting from this point belonging to a specularity. Since this approach is based on finding the smoothest possible representation of diffuse reflection it is sensitive to even the slightest visual texture, and hence requires a very good segmentation of image into regions of similar $\rho_{d}$.

Brelstaff formulates three constraints (one global, two local) on the shading from Lambertian surfaces and explores the effectiveness of detecting specularities in intensity images by searching for image regions which violate these constraints [8]. In the retinex test a bound on the dynamic range of the measured irradiance from a Lambertian surface is established. This bound is calculated by assuming bounds on the dynamic ranges of albedo (assumed to be piecewise constant), reflectance due to directional variations of illumination, and normal orientation. To obtain the bound in illumination a retinex process is applied on the image in order to discount gradual spatial variations. Neither this process nor the bound on illumination variation handles shadows. The cylinder test checks whether a peak in shading is too sharp to come from a diffuse surface by assuming that the local surface geometry is cylindrical. The third test is based on local contrast, and compares intensity values across edges in search of regions where the local shading contrast is too great to be caused by a Lambertian surface. Brelstaff concludes that the retinex and cylinder tests give good evidence for specularities while the local contrast test is of limited use. His experiments on real images were somewhat successful but in every case failed to find some genuine specularities.

### 3.4 Removal of Specular Highlights

To remove a specularity is to separate the colour signal in the affected region into a specular reflection component and a non-specular component. Some detection techniques naturally suggest a removal strategy while others do not. Schemes that compute the colour of the illuminant try to subtract the contribution of the illuminant from the colour signal, but have problems in determining how much to take away and where to take it from.

Klinker's detection scheme was to locate a matte and a highlight line in a distribution of pixels values in RGB space. These lines are estimates of the $\rho_{d}(\lambda)$ and $\rho_{s}(\lambda)$ vectors, so the specular and diffuse components are separated by projecting points lying on the highlight line onto the matte line. For very sharp highlights that have a thin, prominent highlight line this projection produces good results. However, specularities of this kind are the exception and not the rule, and the distribution in colour space of specular pixels can be abstruse for several reasons. First, the dynamic range of the camera causes pixel values to be clipped at a maximum value, which is a very common malady in specular regions, as they tend to be very bright compared to diffuse regions. This can be seen in the plot in Figure 3.5(c). Including these pixels when performing line-fitting will cause the line to deviate from its true orientation, although it is simple to detect and exclude these clipped points. More difficult is to remove the specular component from these pixels; Klinker et al. suggest linearly extrapolating the clipped colour channels based upon the observed variation in the non-clipped channels. A second and more serious cause of ambiguous distributions in colour space is "diffused" or spread-out specular highlights caused by rough surfaces, which is also shown in Figure 3.5. The variation in diffuse reflection in a spread-out highlight region can also cause a line fit to these points to be skewed away from the direction of the illuminant colour (see Figure 3.6). Figure 3.7 shows how the amount of error in reconstructed colour values depends upon the angle between the highlight line and the matte line, the angle of skewing, and the distance along the fit highlight line, all of which can be measured from the colour histogram except for the skewing angle. Effects of skew can be quite noticeable: for an object whose diffuse colour lies along the red axis and a true highlight line that leaves the matte line at an angle of $45^{\circ}$, a skew of as little as $2^{\circ}$ will result in an error of 9 units when the potentially brightest pixel in the highlight (without being clipped, $\operatorname{RGB}=(255,128,0)$ ) is projected onto the matte line. A $5^{\circ}$ skew will give an error of 24 units. Figure 3.8(d) shows errors in the reconstructed diffuse colour signal in the highlight region, due to both skew and not correcting for clipped colour values (left in for illustration).

Novak corrects for skewing by establishing a mapping between measurable histogram properties (length, width, and direction of highlight cluster, and the point of intersection of the matte and highlight lines) to variables in the Torrance-Sparrow reflection model (optical roughness, phase angle, intensity and chromaticity of the light source) [51]. She generated 480 simulated images while varying the reflection model parameters and created a lookup table to relate these to the measurable histogram properties. To analyze a given image the task is then to measure its histogram and interpolate the lookup table entries for these points. In generating these images the Fresnel term was assumed to be 1.5 for all wavelengths, and aside from phase angle (the angle between the viewing and illumination vectors with respect to the surface area) there is no accounting for object shape. The synthetic images used to create the lookup table exhibited a wide variety


Figure 3.6: Effects of surface roughness on pixel distribution in colour space (adopted from [51]).
of surface orientations, and for the histogram measurements to be reliable for a new image it must as well. Novak only tested her approach on 98 synthetic images generated with the same Torrance-Sparrow model and found the average error in the skew computation to be $1.73^{\circ}$. Her only experiments on real images were to compute single parameters.

Figures 3.8 and 3.9 show the result of removing the specular component through one signature searching approach and one specularity-as-violation approach (Klinker [34] and Lee [38]). In the apple example the highlight is spread out which would be expected to cause a skewing problem. Effects of this skewing are seen in Figure 3.8(d) in the reconstructed highlight region. Lee performs well since the diffuse component of the apple is smooth in the image plane. In the pepper example there are no clipped colour values and the highlight region is not spread out in RGB space, with most of the variance accounted for by one dimension. The diffuse colour signal also lies on a vector in RGB space, although the colour signal is not smooth in the image plane due to variations in reflectance and surface orientation. Since Lee searches for the "smoothest" possible description for the diffuse component, his approach incorrectly estimates the illuminant chromaticity. Figure 3.9(c) shows the result of this algorithm. The reconstructed diffuse region is smoother than it should be, and this leads to the discolouration of the region. Klinker's approach (Figure $3.9(\mathrm{~d})$ ) works well since the highlight region is well delineated.


Figure 3.7: Skew geometry and reshading error.

### 3.4.1 Get Real

These techniques sound all very nice in theory, but they have only been demonstrated on either synthetic or squeaky clean untextured real objects, which are precisely those objects that computer graphics can render very well and there is little need for algorithms to alter images of them. Reshading algorithms need to be able to deal with the visual complexity that computer graphics has trouble producing.

Arguably the most important quality for a useful highlight analysis technique is that it be impervious to visual texture (varying $\rho_{d}$ on a surface). None of the approaches discussed in the proceeding sections were designed specifically with texture in mind. A varying diffuse signal on a surface causes problems for the highlights-as-violations approaches since they make assumptions as to how a well-behaved diffuse surface acts. Texture violates the assumptions of piecewise constant reflectance made by Brelstaff in his retinex and contrast tests. Lee's method is very sensitive to changes in reflectance as it is based upon finding the smoothest possible description for the diffuse component. Its success hinges critically on an accurate segmentation of the image into regions constant reflectance, which at best may be possible for coarse, large, and clearly delineated texture regions (such as are found on a chessboard). For most textures one or more of these properties is not the case (wood grain, marble).


Figure 3.8: A comparison of highlight removal strategies on a BC Golden Delicious apple, 89 cents/lb. This example shows the effect of a spread-out highlight and clipped colour values on an object of smooth diffuse reflectance. (top left) Original image (top right) Plot in RGB space of highlight region and its first two principal components, excluding clipped colour values. Plot of entire outlined area is shown in Figure 3.5. (bottom) Reconstructed diffuse colour signal as computed through the algorithm of (left) Lee [38] (right) Klinker et al. .


Figure 3.9: A comparison of highlight removal strategies on a jalapeno pepper. In this example the variance in distribution of highlight pixels in RGB space is mostly one dimensional and the diffuse component exhibits small variations. (top left) Original image (top right) Plot of highlight region in RGB space and its maximally variance-accounting vector (bottom) Reconstructed diffuse colour signal as computed through the algorithms of (left) Lee [38] (right) Klinker et al. .

Texture causes problems for signature searching methods as well. Highlights are detected and removed as L-shaped distributions in colour space: linear clusters emanating from other linear clusters. With varying reflectance on a surface the distribution of diffuse pixels in colour space will not form a nice line, but may be a point, planar, or even a volumetric cluster. This causes two problems. First, it makes it difficult to detect a specularity, since the distribution of specular pixels will likely not form a nice line either, but will spread out from various points in colour space. Second, having detected a highlight, removing it is a challenge since there is no longer an obvious direction to project specular pixel values along, and the diffuse component they should be projected onto is likely not a line. Diffuse linear clusters form when significant shading gradation is present on a region of constant reflectance; on a textured surface, these gradations may only occur for a few of the reflectances in the image. Hypothetical colour distributions are shown in Figure 3.10, two "hidden" diffuse distributions due to a lack of shading gradations are shown in Figure 3.11(b), and a highly spread out distribution is shown in Figure 3.11(c).


Figure 3.10: Effects of texture on colour space distribution. (left) Idealized distribution from a single coloured object. (right) Hypothetical distribution from a textured object (a red, green, and blue texture pattern with a white highlight). The diffuse regions may not form lines to project the specular pixels onto since they may lack shading gradations. Clearly, the closer in chroma the texture patterns are, the more confused the distribution in colour space becomes, and knowing where to project onto becomes more difficult.

Clearly, the closer in chroma a texture pattern is in a given colour space, the more ambiguous the distribution of pixel values from that texture will be. Skewing of the highlight line is a problem on textured surfaces as well. A strategy like Novak's to correct for skew will run into difficulties since histogram measurements are less accurate when a texture is present (for example, the measured "length" of a highlight line may change


Figure 3.11: Reasons why signature searching methods fail in detection and removal of visibly noticeable specularities. (left) [plot of tripod stem from Figure 3.2] The highlight may exhibit little chroma change typical of reflection from metallic objects. (middle) [plot of entire wood texture from Figure 3.2] The variance in colour in the diffuse regions is small compared to that in the highlight region, causing the distribution to be misleading. More shading gradations would stretch the matte line out of the cluster, showing two lines leading towards the black corner of the colour cube. (right) [plot of highlight region on pop can from Figure 3.14] A highly textured object can give a distribution in colour space that is too busy to be of use.
when the highlight occurs on a texture, as the line is bent or muted by varying reflectance). Signature searching approaches also rely upon a segmentation of the input image into regions of constant $\rho_{d}$. Some, like Klinker et al., perform their own segmentation, but the danger in using an underlying theory for segmentation that happens to be the same theory used in removal is that the segmentations tend to be self-fulfilling. Klinker et al.'s segmentation strategy has difficulties in the presence of texture as diffuse regions may fail to form linear clusters (Figure 3.11).

Figures 3.12-3.14 show the effectiveness of Lee and Klinker et al.'s approaches on three different textured surfaces, from fine grained to coarse.

Signature searching approaches have two nice properties that are useful for handling textures. First, the cases in which an approach can be expected to work well are quantifiable through the principal components of the highlight and matte lines. It is therefore possible to determine when the algorithm will fail and switch strategies. Second, since they put no constraints on how far they project a specular pixel along an axis when removing a highlight, they are somewhat insensitive to varying $\rho_{s}$ in the highlight region.


Figure 3.12: A comparison of highlight removal on a small-grain texture. (top left) Original image (top right) Plot in RGB space of entire texture (bottom left) Reconstructed diffuse colour signal as computed by Lee [38]. Since the texture is too fine to be accurately segmented, the smoothness operator removes most of the texture to obtain a smooth diffuse component. (bottom right) Reconstructed diffuse colour signal as computed by Klinker et al.


Figure 3.13: A comparison of highlight removal on a wood texture. (top left) Original image. (top right) Plot in RGB space of highlight region. (bottom) Reconstructed diffuse colour signal as computed through the algorithm of (left) Lee [38] (right) Klinker et al.


Figure 3.14: Highlight removal on a pop can. (left) Original image (right-top) Reconstructed diffuse component by Klinker et al. (right-bottom) Reconstructed diffuse component by Lee.

It is very common for a bright highlight to exceed the dynamic range of a camera, leading to sharp, clipped specular pixel values. In this case there is no signal underneath the highlight to reconstruct. Since the spatial information inherent in an image is lost in colour space, signature searching methods have trouble. To reconstruct the colour signal in these cases some form of discontinuity-preserving surface reconstruction across a region of no information is needed. Interpolation from the boundaries of the highlight region is the only clue as to what should be reconstructed. Figure 3.14 shows highlight removal with Klinker et al. and Lee's methods with the user manually outlining the diffuse regions to reconstruct under the highlight.

## Chapter Four

# Reshading Knowing Object Geometry 



The classical computer graphic approach for creating an image is to use some modelling system to construct a three dimensional geometric model of the objects to be rendered, assign surface properties to each object, place a number of lights in the scene and a viewpoint. A global illumination algorithm is then applied to the environment which simulates the propagation of light through the scene and onto the image plane using one of any number of local models of light reflection. This process of rendering an image can be formalized as finding an approximate solution to the integral equation of light transport (the rendering equation) as seen from the position of the viewer [31]:

$$
\begin{aligned}
P\left(\mathbf{x}_{t}\right)= & \text { Sensor }\left(\text { Geometry } ( \mathbf { x } _ { r } , \mathbf { x } _ { t } ) \left[\operatorname{Emission}\left(\mathbf{x}_{r} \rightarrow \mathbf{x}_{t}\right)+\right.\right. \\
& \left.\left.\int_{V} \underset{\text { Properties }}{\text { Reflectance }}\left(\mathbf{x}_{t}, \mathbf{x}_{r}, \mathbf{x}_{\ell}\right) \quad \underset{\text { Light }}{\text { Incoming }}\left(\mathbf{x}_{\ell} \rightarrow \mathbf{x}_{r}\right) d \mathbf{x}_{\ell}\right]\right)
\end{aligned}
$$

[^10]The goal of reshading is to allow a user to alter an image in terms of scene properties, such as moving or dimming a light, changing the geometry of an object or the reflectance properties of a material. These alterations will have an effect on some or all of the original image, and a new image whose pixel values $P^{\prime}$ are altered versions of the old values $P$ must be generated:

$$
P^{\prime}\left(\mathbf{x}_{t}\right)=P\left(\mathbf{x}_{t}\right)+\Delta P\left(\mathbf{x}_{t}\right)
$$

The challenge is to compute the $\Delta P$ (or $P^{\prime}$ ) for a wide variety of scene changes. What is involved in computing $P^{\prime}$ depends in part upon the operation being performed; some possible reshading operations are shown in table 4.1. These assume the standard computer graphics model of local illumination:

$$
P\left(\mathrm{x}_{t}\right)=\text { Sensor }\left(\text { Ambient }+\sum_{\ell=1}^{n} f_{r} L_{\ell} F_{r, \ell}\right)
$$

Not all operations are equal in computational cost and effect. Changing the illumination conditions potentially affects all regions of a scene, and necessitates (re)rendering the scene, requiring knowledge of how bright the light sources are, where they are, and the geometry and reflectance properties of the visible objects. On the other hand, changing an object's reflectance affects primarily only that object alone (although to be complete, the ambient term should be adjusted), and only requires knowledge of what an "object" is and where in the image it lies.

Another factor determining the difficulty of computing $P^{\prime}$ is the amount of knowledge available about the source scene. In reshading we are confronted with the inverse problem of rendering, in that one starts with the pixel values and must solve for some unknown parameters on the right hand side of the rendering equation, which in turn are changed to produce the new image. Since the value formed at a pixel depends upon many factors, doing this solely from image data is a heavily underconstrained problem. Additional information or constraints are required but what is reasonable to ask the user to provide, and what should be solved for? Assuming too much leaves the problem impractical, while assuming too little leaves it intractable.

This chapter examines the task of reshading objects when the facing geometry of the objects in the source image is known, as well as the imaging and lighting geometry. In terms of the reshading operators in table 4.1 this means the $F_{i, j}$ are known. It may seem unrealistic to expect this much to be known about a real scene,

| Alter... | Variables Affected | Compute |
| :---: | :---: | :---: |
| Emissivity | $L_{\ell} \rightarrow k L_{\ell}$ | $\Delta P=f_{r}(k-1) L_{\ell} F_{r, \ell}$ |
| for each pixel |  |  |

Table 4.1: Some reshading operations.
but there are many instances when these conditions can be met. The growing accessibility of rangefinders and the increasing robustness of computer vision algorithms promises that geometric information about objects will become easier to obtain in the future. For many real objects used in computer graphics the first level approximation of their geometry is simple (a tabletop wood grain is planar, a coke can is a cylinder) and techniques now exist for registering synthetic models with images [18]. Determining the lighting geometry is more complicated business, although some progress in this area has been made and is discussed in section 4.2. In this work it is assumed that there are no shadows present on the objects being operated on. Solving for the relative light source emissivities of direct and indirect light of a scene of known light and object geometry is discussed in section 4.3.

Computing the frontal geometry of viewed objects and the parameters of the camera which took the image are well-studied topics. A recent survey of pose estimation from image features is given by Lowe [43]. The shape from * literature is vast; probably a good place to start is [77].

### 4.1 Pixels Torn Asunder: Previous Work in Reshading

The first attempt in computer graphics at simulating lighting interactions between elements from different images was made by Nakamae et al. [46]. They use digital photographs of outdoor landscapes as a setting into which computer generated objects are inserted. Thirion later addressed the similar problem of generating arbitrary images of a city block given a number of photographs taken from known locations and a complete 3D model of the scene [65]. Both Nakamae et al. and Thirion exploit the simplicity inherent to outdoor scenes viewed at a distance: they approximate the illumination conditions with a directional source and a spatially invariant ambient light, assume all objects reflect diffusely, and require the coarse 3D object geometry and the sun's position to be provided. Nakamae et al. calculate the relative strength of the light sources as the average ratio of pixel values in several manually chosen sunlit and shaded regions. These estimates are used for rendering synthetic objects that are composited into the original photograph, and shadows cast by these objects are presumably simulated by adjusting the affected regions to reflect only the ambient source. The limitations of their approach are the reliance on manual intervention, the implicit assumption that all objects have the same albedo, and the restriction of a single directional source. Thirion estimates the relative albedo at each pixel by linearly interpolating image values, and uses these estimates to simulate the scene under varying lighting conditions with a simple Lambertian shading model. Thirion's approach implicitly assumes that all objects have the same geometry with respect to the light source, and relies upon the single directional source assumption.

Schoeneman et al. present an interactive system that allows a user to paint the desired appearance of a synthetic scene which then solves for the light source emissivities that gives a radiosity solution that best fits it in a least-squares sense [59]. Their challenge is to constrain the user to remain in the realm of consistent or perhaps even physically realizable solutions, and to produce the solution at interactive speeds. For a given scene the unknowns are patch radiosities and light source emissivities; the exact scene geometry, light source geometry, and patch reflectances are all known. The current solution and the user's constrained painting assigns radiosity values to patches, which in turn are used to solve for the emissivities and a new, consistent result for the entire scene. They minimize the weighted sum of squared error of patch radiosities, with the weights proportional to the area of the patch and the user's interest in that patch. The resulting linear system is solved iteratively with negative radiosity values clipped to zero to constrain the solution to a physically meaningful result. The sum of squared error measure is an absolute difference measure, while the human eye is sensitive to relative differences, so visible changes in dark regions may not be weighted heavily enough. They weigh each of the red, green, and blue channels equally.


Figure 4.1: An example of adding light and objects to a real scene (taken from [16]).

Fournier et al. deal with issues involved in establishing common illumination between real indoor scenes and an inserted synthetic component [16]. Given an image of a real scene, a coarse 3D model of the real objects, the positions of the light sources and an estimate of the area-weighted albedo of the surfaces, they solve for light source emissivities and patch albedos and use these values to drive a radiosity solution for the global redistribution of light in the combined scene. The radiosities of visible patches are computed as the average of the pixel values corresponding to the patch; hidden patches are assigned an ambient radiosity computed as the mean radiosity value in the image. Patches' albedos are set to the estimated average provided, with the reflectances of visible patches scaled based on a local examination of radiosity values. Light source emissivities are determined as those that give the best least-squares fit between the estimated radiosity values and those given with a radiosity computation. They then insert a number of additional lights and objects into the scene, compute the change in radiosity for each patch and produce a new image from these differences. They make no attempt to account for non-Lambertian shading artifacts, and their provided model was already subdivided into faces which they treated as regions of constant reflectance. They offer no quantitative evidence for the accuracy of their approach but rely upon empirical evaluation.

Kawai et al. present a lighting design system for a diffuse scene of known geometry [33]. They allow a user to specify the lighting conditions of a synthetic scene by both placing constraints on rendering parameters and by specifying subjective goals such as the "privacy" or "pleasantness" of the solution. Free variables in the system are light source emissivities, element reflectances and the direction and falloff term for spotlights. They pose the problem as a constrained nonlinear optimization problem, convert the solution constraints into
penalty functions, and minimize the resulting objective function with Broyden-Fletcher-Goldfarb-Shanno (BFGS) multidimensional quasi-Newton minimization that uses gradient information. They offer the user a library of objective functions, several of which can be linearly combined into the general objective function. Their initial solution is provided with a baseline radiosity rendering of the scene with no free variables. The user then selects the free variables and places constraints on them at each iteration of the minimization. Kawai et al. note that before they introduced the psychophysical constraints their system "required quite a bit of unintuitive tweaking of the objective function weights in order to achieve lighting that had the right subjective appearance". Examples provided were produced in a matter of minutes.

Poulin presents an interactive system for defining surface properties by painting desired colours or ranges of colours onto the object [56]. Housed in a modelling system, the knowns are the geometry of the object being painted, geometry and emission of the light sources, viewing geometry, and specular roughness. The unknowns depend upon which shading model is used, which also dictates whether the constraints formed by the user's painting are linear or nonlinear. The local shading models considered by Poulin account for diffuse and specular reflection from direct illumination and represent indirect illumination with an ambient term. In the case where the system of equations is overconstrained, nonlinear weighted least squares fitting based on the Levenberg-Marquardt algorithm is used, with the constraints transformed into penalty functions. Each painted colour is assigned an arbitrary weight depending upon if it lies in shadow, exhibits diffuse reflection only, or corresponds to a highlight region. In the underconstrained case a unique solution is found by minimizing an objective function which the user may personalize through a library of objective functions. Poulin recommends a two stage process of first minimizing the ambient term and then maximizing the diffuse component as giving intuitive behaviour and interesting results. A feasible initial solution is determined by analyzing the boundary conditions on the shading parameters.

### 4.2 Rectangular Source from Geometry

This section discusses how to solve for the position, orientation, size, and emissive strength of a rectangular light area from the shading of an object of known geometry, assuming it is illuminated from the source alone.

| Approach | Scenario | Assumptions / Knowns |  |  | Solved For |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | Object/Viewing <br> Geometry | Illumination | BRDF |  |
| Nakamae et al. | Common illumination between synthetic objects and an outdoor scene; direct illumination | Viewpoint computed by minimizing image space displacement error of a number of known landmarks <br> Geometry of real scene manually approximated with horizontal and vertical planes | Real illuminants approximated with directional source of known direction and a spatially invariant ambient light | Lambertian + ambient, reflectance assumed constant | Relative intensities of illuminant analytically computed by manually examining sunlight and shady regions in photograph |
| Thirion | Real aerial images reproduced under varying lighting conditions direct illumination | Viewpoint known <br> Full model known, not necessarily finely detailed | Real illuminants approximated with directional source of known direction and a spatially invariant ambient light | Lambertian + ambient, reflectance assumed constant for all surfaces | Surface albedo at each pixel computed through interpolation of pixel values |
| Schoeneman et al. | User paints radiosity onto patches | Viewpoint known, under user control <br> Full 3D geometry known | Light source geometry known | Lambertian, known | Patch radiosities and light source emissivities; overconstrained; solved via linear least squares |
| Fournier et al. | Synthetic objects and lights inserted into a real scene, common illumination through radiosity | Camera parameters computed by manually tracking known landmarks <br> Geometry of real scene provided, coarse approximation tolerable, synthetic objects registered to this | Light source positions known, strength of ambient light computed as energy required to produce image given estimated reflectance | Lambertian, model of real scene arbitrarily subdivided into patches of constant reflectance, average reflectance estimated by user | Relative emissive strengths; overconstrained; solved through linear best-fit radiosity solution |
| Kawai et al. | Interactive lighting design for radiosity environments | Viewing known <br> Full 3D geometry known | Position of emitters known | Lambertian, model subdivided into patches of constant reflectance | Light source emissivity, patch reflectance, spotlight directionality and falloff; underconstrained; BFGS nonlinear optimization, objective functions include radiometric and psychophysical constraints made into penalty functions |
| Poulin | Painting surface characteristics onto 3D objects; local illumination | Viewing parameters known <br> Object geometry known | Light positions, emissivities known | Includes specularities, specular roughness parameter known | Surface characteristics which depend upon shading model employed; both over and underconstrained systems; nonlinear least squares and FSQP optimization with potentially nonlinear constraints |

### 4.2.1 How to Find That Light? Previous Work

Most everybody deals with a single light source. The polyhedra boys [63] ignore graduated shading effects and use only greater than/less than comparisons between diffusely reflecting planar facets to constrain the position of a single directional light source. Hanrahan and Haeberli note that the point of maximum intensity of a highlight defines a light direction [21]. Poulin and Fournier use the point of maximum reflection and a point on a highlight's boundary to define the light direction and the solve for the surface's roughness coefficient using the Phong model [55]. They also present interactive techniques for defining point, directional, and extended sources by manipulating the shadows cast by object primitives. Poulin later shows how the point of maximum brightness from a diffuse surface can be used to define a directional source, or a point source constrained to lie on a plane [56].

Ikeuchi and Sato estimate the diffuse and specular reflectance weights, specular roughness parameter and the orientation of a directional source from a range and intensity image [29]. They base their shading analysis upon a simplified Torrence-Sparrow model. They ignore regions in which the incident and reflected angles exceed $60^{\circ}$, assume the Fresnel term is constant for the imaged surface, and that microfacets do not mask or shadow each other. They also assume that the object in the image has constant reflectance (implying a segmentation into regions of common reffectance) and that the images are captured through orthographic projection. Their analysis proceeds in two stages. In the first stage the illuminant direction and diffuse reflectance are estimated through an iterative fitting/segmentation process. Each pixel's intensity is assumed to arise from diffuse reflection; a linear least squares fit for the unknowns is performed and outliers are discarded until convergence. In the second stage those pixels which pass a brightness and geometric threshold are tagged as potentially exhibiting specular reflection. This leaves the specular roughness and reflectance parameters unknown. Rather than perform a multidimensional nonlinear fit to determine these Ikeuchi and Sato alternatively solve for one unknown while holding the other constant. The initial solution for the specular reflectance is estimated from the pixel whose normal is closest to the perfect specular angle. Their two-stage approach requires that the image arises predominantly from diffuse reflection and makes use of several thresholds. It assumes that the imaged surface has constant reflectance, and cannot be easily extended to handling multiple light sources.

Solving for a light source direction from just an image has been addressed by Liu and others [42]. Solving for multiple sources has only recently shown some activity [78].

### 4.2.2 Problem Formulation

A gap exists between what the approaches in the previous section attempt to accomplish and what exists in reality. First, there often is more than one light source. Second, real illuminants are not directional or point sources, but have definite area. This subsection addresses this possibility.

Assume we have an object that is illuminated by only a single rectangular source. Can we determine the properties of this light source from the shading of the object? The geometry of a rectangular source in 3D can be described by 8 parameters: for example, its center $\vec{C}$, orientation $\vec{O}$, and scale $\vec{S}$ relative to a canonical square of length 1 lying in the XZ plane, as diagrammed in Figure 4.2.

$$
\begin{aligned}
\vec{C} & =(x, y, z) \\
\vec{O} & =(\phi, \Theta, \vartheta) \\
\vec{S} & =(l, w)
\end{aligned}
$$

In addition the rectangular source has radiometric properties. If we assume that the source is diffuse (constant emission on all points of the source and in all directions) then these properties can be represented by $\vec{E}_{\ell}$, the emissive strength for all wavelengths. We will work in terms of a single wavelength $\lambda$ and emissivity $E_{\ell, \lambda}$.

A rectangular source can then be represented as a point $\vec{u}=\left(\vec{C}, \vec{O}, \vec{S}, E_{\ell, \lambda}\right)=\left(u_{1}, \cdots, u_{i}, \cdots, u_{9}\right)$ in the space spanned by the unknown parameters.

We have 9 unknowns to determine from the shading in the image. No closed form solution exists, and if there were one it would not give an exact solution due to differences between our idealized model and that which exists in the real scene, indirect illumination, noise in the imaging process, etc. Therefore we formulate the problem in terms of a search to minimize an objective function. To evaluate the quality of our estimated source we use the sum of the squared difference between the original image and a new image generated with the estimated source:

$$
\begin{equation*}
\text { Residual }=\sum_{\forall \text { pixels } i} \sum_{\lambda \in\{r, g, b\}}\left(P_{i \lambda}^{\prime}(\vec{u})-P_{i \lambda}\right)^{2} \tag{4.1}
\end{equation*}
$$

where $P_{i \lambda}$ is the value of pixel $i$ in colour channel $\lambda$ and $P_{i \lambda}^{\prime}$ is the computed value for the pixel using the


Figure 4.2: Parameters for defining the geometry of a rectangular source.
estimated light source $\vec{u}$ (from equation (2.1)):

$$
\begin{gathered}
P_{i \lambda}^{\prime}(\vec{u})=\operatorname{Sensor}\left(L_{\lambda}\left(\mathbf{x}_{r}, \vec{\omega}_{t}\right)\right) \\
L_{\lambda}\left(\mathbf{x}_{r}, \vec{\omega}_{t}\right)=L_{\ell, \lambda} \int_{A_{\ell}} f_{r, \lambda}\left(\mathbf{x}_{r}, \vec{\omega}_{\ell}, \vec{\omega}_{t}\right) G\left(\mathbf{x}_{r}, \mathbf{x}_{\ell}\right) V\left(\mathbf{x}_{r}, \mathbf{x}_{\ell}\right) d A_{\ell}
\end{gathered}
$$

The sought after solution is the parameter vector $\vec{u}$ that minimizes equation (4.1). However, not every solution is acceptable, as negative values for emissive strength are not physically possible. Constraints on the emissive strength $L_{\ell, \lambda}$ are converted to penalty functions of the form:

$$
\text { Penalty }=e^{-L_{\ell, \lambda}} \quad \text { if } L_{\ell, \lambda}<0
$$

which is combined with the residual measure to give the objective function $f(\vec{u})$ :

$$
\begin{equation*}
f(\vec{u})=\sum_{\forall \text { pixels } i} \sum_{\lambda \in\{r, g, b\}}\left(\left(P_{i \lambda}^{\prime}(\vec{u})-P_{i \lambda}\right)^{2}\right)+\text { Penalty } \tag{4.2}
\end{equation*}
$$

Finding a point $\vec{u}$ which best satisfies our objectives $f(\vec{u})$ is then an unconstrained nonlinear minimization
problem. However, equation (4.2) has a lot of local minima, and to have any hope of converging to a global minimum or finding a satisfying local minimum we need a good initial solution. What clues can the image give us?

### 4.2.3 Constraints from Diffuse Shading

Working on an image region which arises solely through diffuse reflection reduces $f_{r, \lambda}$ to $k_{d, \lambda} / \pi$ (from equation (2.3) and (2.5)):

$$
\begin{gather*}
L_{\lambda}\left(\mathbf{x}_{r}, \vec{\omega}_{t}\right)=L_{\ell, \lambda} k_{d, \lambda} F_{r, \ell} \\
F_{r, \ell}=\int_{A_{\ell}} \frac{\cos \theta_{r} \cos \theta_{\ell}}{\pi\left|\mathbf{x}_{\ell}-\mathbf{x}_{r}\right|^{2}} d A_{\ell}=\frac{1}{2 \pi} \sum_{i} \vec{\Gamma}_{i} \cdot \overrightarrow{\mathrm{~N}}_{d A_{r}} \tag{4.3}
\end{gather*}
$$

The visibility term $V$ has been dropped from equation (4.3) as we only consider the case where the source is unoccluded. In what follows we will treat the surface area that projects to the pixel we are studying as an infinitesimal $d A_{r}$ that is centered about the point $\mathbf{x}_{r}$. All terms are the same as defined in Chapter 2.

## BOUNDARIES

Every point in an image contributes some constraint upon the light source geometry, although it may not be very useful. For a point $\mathbf{x}_{r}$ to exhibit some Lambertian reflection means that $\cos \theta_{r}, \cos \theta_{\ell}$, and $1 /\left|\mathbf{x}_{\ell}-\mathbf{x}_{r}\right|^{2}$ must all be greater than zero for at least one point on the light source (and significant enough so their product is greater than zero). This means that at least one vertex of the source must lie above the tangent plane at $\mathbf{x}_{r}$ and that $\mathbf{x}_{r}$ must lie in the hemisphere above the light plane (the plane that the rectangular source lies in).

Conversely, for a point to not be shaded by a light source, one of the above terms must be zero for all points on the source. So each point in the image imposes some constraint which discounts a portion of the 8 -dimensional space spanning the geometry of all rectangles from the realm of potential solutions. Across smooth surfaces, any one point does not add much new information since the small changes in position and orientation from its neighbours will bring only small changes in constraints.

Some points give a lot of information, though. Suppose we have found a region in the image where reflection from the sought-after source falls to zero smoothly. If the point $\mathbf{x}_{r}$ has no reflection from the source and a nearby point $\mathbf{x}_{r}^{\prime}$ has some, then this implies that either $1 /\left|\mathbf{x}_{\ell}-\mathbf{x}_{r}\right|, \cos \theta_{\ell}$, or $\cos \theta_{r}$ goes from being
nonzero at $\mathbf{x}_{r}^{\prime}$ to zero at $\mathbf{x}_{r}$. We can discount the first possibility if the points in question lie on a connected surface. The second case occurs when the source falls below the tangent plane at $\mathbf{x}_{r}$. Since the point $\mathbf{x}_{r}^{\prime}$ reflects some light, at least one vertex of the source must lie in the space bounded by the tangent planes to $\mathbf{x}_{r}$ and $\mathbf{x}_{r}^{\prime}$. The third case occurs when $\mathbf{x}_{r}$ falls below the light plane. This case offers the constraints that $\mathbf{x}_{r}$ lies in the light plane and the normal to the source is such that $\mathbf{x}_{r}^{\prime}$ receives light. An infinite number of planes satisfy the constraints given by one $\mathbf{x}_{r}$ and $\mathbf{x}_{r}^{\prime}$, but if a reflection/no-reflection contour on the surface can be localized, then the family of planes can be further reduced. All points on such a boundary should lie in the light plane, so the plane will only have one rotational degree of freedom (about this line). The 2D versions of these cases are shown in Figure 4.3. Similar constraints are formed by discontinuities in surface orientation and by the silhouette of an object. These constrain the position of the light source less but are much easier to detect.


Figure 4.3: Conditions where the reflected radiance at point $\mathbf{x}_{r}$ becomes zero. (left) $\vec{N}_{\ell} \cdot \vec{L}=0$. (right) $\vec{N} \cdot \vec{L}=0$.

As each region where $\cos \theta_{\ell}$ goes to zero defines a family of plausible light planes, then by locating two or more such regions the plane that contains all of the boundary lines will define the light plane (Figure 4.4 shows the 2D case). This plane will likely not exist due to errors in localizing the boundaries, so some best-fitting strategy will need to be employed. This strategy is not reliable in itself since localizing a contour that goes to zero smoothly is difficult, especially with indirect illumination falling on the surface. We must
be able to determine when the signal goes to zero because of $\cos \theta_{\ell}$ and not $\cos \theta_{r}$. Note that a $\vec{N} \cdot \overrightarrow{\mathrm{~L}}$ boundary cannot occur on a planar surface.


Figure 4.4: Fixing the light plane from Lambertian boundaries.

The shading in many regions of an image does not change smoothly. Discontinuities in image irradiance arise from abrupt changes in surface orientation, distance from the source, visibility of the light source, and reflectance. Discontinuities in surface orientation and position can be determined since we know the frontal geometry of the scene. Discontinuities in visibility arise from self-shadowing and blocking and can cause discontinuities in the first and second derivative of image shading [25]. Determining lighting information from shadows has been addressed by Poulin and Fournier [55].

## Maxima

Since Lambertian reflection varies with the cosine of the angle between a surface normal and the direction of incoming light, the surface normal at the point of maximum diffuse reflection (the point on a diffusely reflecting surface that corresponds to the pixel with the largest intensity) has in the past been used to define a light direction in lighting design systems. For a synthetic scene this direction defines a parallel light source and constrains the position of a point source. To use the point of maximum diffuse reflection to estimate the geometric properties of a rectangular source, note that equation (4.3) attains a maximum for a material of constant $k_{d, \lambda}$ when $F_{r, \ell}$ is at its maximum. This occurs when the source is an infinitely large rectangle which touches the shaded surface at $\mathrm{x}_{r}$. This is not too useful a solution, but if we hold the size of the source constant as well as its distance to the surface then we can then make a semi-useful conclusion about the
remaining parameters. Assume that the source has length $2 l$ and width $2 w$ and its center is at a distance $D$ from $\mathbf{x}_{r}$. To find where $F_{r, \ell}$ takes on its maximum value in terms of the remaining free parameters, we first need to locate its critical points $\left(\vec{\nabla} F_{d A_{r} \rightarrow A_{\ell}}=\overrightarrow{0}\right)$. The source has six degrees of freedom (three positional $\vec{C}$ and three rotational $\vec{O}$ ) and one constraint: $\left\|\vec{C}-\overrightarrow{\mathbf{x}_{r}}\right\|=D$. Rather than solve a constrained extrema problem we reparameterize the source in terms of spherical coordinates: the center of the source lies on the hemisphere of radius $D$ above $d A_{r}$ and is given by $\vec{C}=\left(D, \theta_{c}, \phi_{c}\right)$. Without loss of generality assume that $d A_{r}$ sits at the origin of our coordinate system and its normal is aligned with the $Y$ axis (see Figure 4.5).


Figure 4.5: Coordinate system for maximizing the form factor of a constrained rectangle.

This reduces the expression for the form factor to

$$
\begin{equation*}
F_{r, \ell}=\frac{1}{2 \pi}\left(\Gamma_{1, y}+\Gamma_{2, y}+\Gamma_{3, y}+\Gamma_{4, y}\right) \tag{4.4}
\end{equation*}
$$

where $\Gamma_{i, y}$ is the y component of the vector $\vec{\Gamma}_{i}^{\dagger}$.

[^11]

Figure 4.6: Diffuse integral evaluated for the XZ plane illuminated by a rectangular source aligned with the XZ axes centered at $(0,10,0)$, surface and source parallel and facing each other $\left(\theta_{c}=0, \phi=0, \Theta=0, \vartheta=0, l=10, w=2\right.$, $D=10$ ). (left) view of scene setup in XY plane, $Z$ axis comes out of the page (right) contours of constant reflected radiance. The maximum radiance occurs at the center and decreases radially. (contours in the central region have been removed, horizontal axis is X , vertical axis is -Z ).

Equation (4.4) has a critical point at $\left(\theta_{c}, \phi, \Theta, \vartheta\right)=(0,0,0,0)$. This corresponds to the case when the rectangular source faces, is parallel to, and is centered above $d A_{r}$, which is the intuitive solution. Clearly from the geometric interpretation of the form factor this is a maximum, and indeed the global maximum. Figure 4.6 shows the relative reflected radiance of a planar surface illuminated in this configuration. Note that the light source still has one rotational degree of freedom since rotating it about its normal does not affect the form factor.

If we can find the "ideal" point of maximum diffuse reflection in the image we can constrain the orientation of the rectangular source and fix a direction which pierces its center. This is a strong constraint and would allow us to estimate the 3D position of the source within a narrow range if we obtain an additional, independent constraint upon its position. Unfortunately, finding this ideal point is in no way an easy task. There can only be one such point in any scene for each light source, and it may not exist in the scene, it may exist but not be visible in the image, or can just be hard to recognize due to texture or indirect illumination. Such a point need not be the global maximum in the image as it may lie on an object of low reflectance,


Figure 4.7: Diffuse reflection from the XZ plane illuminated by a square source centered at $(0,50,0)$. (left) source oriented at $45^{\circ}$ to the XZ plane (right) source perpendicular to XZ plane. (top) scene setup (middle) graph of reflection for line $\mathrm{z}=0$ (bottom) iso-reflection contours.
although it is a local maximum if reflectance is constant in its surrounding area. One is not safe in using local maxima to determine a direction towards the light source, however, since variations in distance and relative orientation to the source play a role in shading, as illustrated in Figure 4.7. The first column shows a Lambertian planar surface lying in the XZ plane illuminated by a square source oriented at $45^{\circ}$ to it (about the Z axis). Reflected radiance from this plane is shown in the middle and bottom plots. The point $(10,0,0)$ exhibits maximum reflection, and the normal at this point does not intersect the light source. The point $(0,0,0)$, whose normal pierces the center of the source, reflects $92.2 \%$ as much as the maximum. The right column shows the case when the source is perpendicular to the surface being shaded. At no point on this surface does the normal intersect the source: the point $(-28,0,0)$ exhibits maximum reflection.

The ideal point of maximum reflectance can be detected by its characteristic gradient surrounding the maximum. The magnitude of the gradient (with respect to world coordinates) is the same for all points on an isoreflection contour. Such a measure is very sensitive to noise and texture, however. Barring this one could treat each point in the set of local maxima as the ideal maximum to set constraints in several versions of the minimization, and choose the best result. Finding the point in an image which exhibits the greatest reflectance value is often not just a case of choosing the largest pixel value, since diffuse reflection changes slowly and large image regions may have similarly high pixel values.

For the shading of a single object where the distance from the source to points on the object are roughly equal, the normal at the point of maximum reflection may be a useful estimate of the direction towards the center of the source. In general, though, the surface normal at a local maximum is not a particularly trustworthy indicator of a direction to the source as it may not even intersect the light plane. What can we say about the maxima from diffuse surfaces? The reflection at each point on a shaded surface can be represented as the sum of the product of three smooth functions over the area of the source (one for $\cos \theta_{\ell}$, one for $\cos \theta_{r}$, and one for $r^{2}$ ). The point of maximum reflection will in general not coincide with a point on the surface where the integral of any of these functions attains its maximum (although for Figure 4.6 all of these points are the same). As we move away from a local maximum some term in the integral is decreasing, but it may be one, several, or all terms. The directional derivative of reflection can aid in adding constraint. For instance if we know the position and size of the light but not its orientation, then examining the change in reflection in the direction towards the source gives clues about the orientation. The gradient of irradiance in synthetic scenes has been examined in terms of known surface, illuminant, and blocker geometry, primarily in the context of adaptive meshing [74, 2].

### 4.2.4 Constraints from Specular Highlights

Suppose we have detected a specular highlight. The contribution of the specular component to the colour signal in this region is

$$
L_{s, \lambda}\left(\mathbf{x}_{r}, \vec{\omega}_{t}\right)=L_{\ell, \lambda} \int_{A_{\ell}} k_{s, \lambda} f_{s}\left(\mathbf{x}_{r}, \vec{\omega}_{\ell}, \vec{\omega}_{t}\right) G\left(\mathbf{x}_{r}, \mathbf{x}_{\ell}\right) d A_{\ell}
$$

In the world of Phong this becomes

$$
\begin{equation*}
L_{s, \lambda}\left(\mathbf{x}_{r}, \vec{\omega}_{t}\right)=L_{\ell, \lambda} k_{s, \lambda} \frac{2 \pi}{n+1} \int_{A_{\ell}} \cos ^{n} \varphi \frac{\cos \theta_{\ell}}{\left|\mathbf{x}_{\ell}-\mathbf{x}_{r}\right|^{2}} d A_{\ell} \tag{4.5}
\end{equation*}
$$

Specularities are nice in that they are very nearly mirror images of the light source, and therefore give many clues about it. Unfortunately they are viewpoint dependent, and give us and additional unknown $(n)$ to deal with. Also it is not correct to treat specular reflection from a surface in isolation, since there is a diffuse reflection component in a highlight region as well, although in many cases the diffuse component may be nearly constant in this region.

## MAXIMUM

Having a specularity at a point is pretty strong evidence that the direction of perfect specular reflection for that point passes near to a light source. Even so, one is not justified in assuming that the direction of perfect specular reflection at the point of maximum specular reflection intersects the center of the source. Figure 4.8 shows reflection from a planar surface using the Phong specular reflection model with two different values for the roughness coefficient $n$. In both these cases the center of the source is pierced by the vector of perfect specular reflection at the point $(0,0,0)$. In neither case does the direction of perfect specular reflection at the point of maximum reflection intersect the light source. We see that as the surface acts more like a mirror reflector (i.e. as $n \rightarrow \infty, \cos ^{n} \phi \rightarrow \delta(\vec{R}, \vec{L})$, where $\delta$ is the Dirac delta function), the $\cos ^{n} \varphi$ term dominates the integral and the point of maximum reflection migrates closer to the origin. Even so, this direction is a good constraint for finding an initial solution. Finding a "maximum" point in a highlight is often not easy, since irradiance values quite often exceed the dynamic range of the camera, and extrapolation of values in the highlight region must be performed.


Figure 4.8: Specular (Phong) reflection from the XZ plane illuminated by a square source centered at ( $50,50,0$ ), eye at $(-50,50,0)$, both normals intersect the XZ plane at the origin. (left) roughness coefficient $n=5$, maximum occurs at $(9,0,0)$ (right) roughness coefficient $n=10$, maximum occurs at ( $5,0,0$ ) (top) scene setup (middle) graph of reflection for line $\mathrm{z}=0$ (bottom) iso-reflection contours.

## Boundary

At the boundary of a highlight, equation (4.5) goes to zero (or more properly, below some small threshold $\epsilon$ ). This occurs for one of two reasons:

1. $\cos \theta_{\ell} d A_{\ell} /\left|\mathbf{x}_{\ell}-\mathbf{x}_{r}\right|^{2}$ goes to zero, which (discounting occlusion) means the light source falls below the horizon of the tangent plane at $\mathbf{x}_{r}$. This is a specularity at a grazing angle and this can be detected since diffuse reflection at grazing angles also goes to zero.
2. $\cos ^{n} \varphi$ goes to zero.

Since we know the surface geometry, we know if case (2) comes about by an abrupt change in surface orientation. For a smooth surface there is an ambiguity as to why $\cos ^{n} \varphi$ goes to zero: $\varphi$ could be large while $n$ is small, or $\varphi$ could be small but $n$ is large. If at some point we determine the true size of the light source, then we can use the boundary of the highlight to determine the roughness coefficient $n$. Without prior knowledge of $n$, though, all we can deduce is that at the boundary the vector $R$ does not intersect the light source (since in this case the value of $n$ is immaterial). Applying this constraint to all points on the highlight boundary gives a cone that bounds a volume of space that contains the source. A 2D example is shown in Figure 4.9.

How can we disambiguate big ( $\vec{R} \cdot \vec{L}$ ) and big $n$ ? The bigger $n$ is, the more mirror-like the highlight becomes. For some surfaces highlights have two distinct regions: a sharp, mirror-like region in the interior of the specularity, and a diffused region surrounding it. Usually the boundary between the mirror-like "inner" highlight and the diffuse "outer" highlight is somewhat distinct, but the boundary between the outer highlight and the non-specular region is less well-defined. For these surfaces the diffused highlight region comes from those $\vec{R}$ that do not intersect the source. Therefore if the boundary of the inner highlight can be accurately determined, we have a good estimate of the cone that the light source subtends (see Figure 4.10). Knowledge of $n$ or a distance to the source would then give a precise fix of the source position.

One should probably not put too much stock in gaining information from the outer boundary of diffused highlights by applying the Phong model, since it has never been shown to fit measured data well.


Figure 4.9: Source-enclosing cone given by a highlight boundary.

### 4.2.5 Experiment

As a simple proof-of-concept, we ran a test of the minimization strategy on a synthetic example of a teapot. Figure 4.11 shows the original image, an image rendered using the solved-for source, and the average error per pixel between the rendered image and the original. The teapot was originally rendered with a single area source and low levels of ambient illumination. To simulate loss of data common in using depth-finding mechanisms, we determined the depth to points on the model, stored them at a low resolution, and then reconstructed the surface using orthogonal multinomials [5]. This resulted in some errors in the surface geometry used in the minimization.

We first used Powell's method to minimize $f(\vec{u})$ and found that it did not converge to a reasonable minimum for synthetic tests unless it was given an initial solution that was unreasonably close to the true solution. To improve upon this we then used Broyden-Fletcher-Goldfarb-Shanno (BFGS) minimization which operates by performing a sequence of 1 D minimizations and uses gradient information [57]:

$$
\frac{\partial f(\vec{u})}{\partial u_{i}}=\sum_{\forall \text { pixels } i} \sum_{\lambda \in\{r, g, b\}} 2\left(P_{i \lambda}^{\prime}(\vec{u})-P_{i \lambda}\right) \frac{\partial P_{i \lambda}^{\prime}}{\partial u_{i}}
$$

We compute the partial derivatives analytically.


Figure 4.10: Sharp versus diffused highlight boundary.

The average error per pixel in each colour channel between the rerendered image and the original is shown in the table in Figure 4.11. To determine how close these values were to an ideal solution, we also rendered the reconstructed teapot model with the same area source used to generate the original image. This resulted in the error reported in column 2 of the table. From these results we see that the area source given by the minimization comes very close to the suspected global minimum.

The only constraints used to obtain an initial solution were the point of maximum reflection and the shadow region on the edge of the teapot. A distance from the center of the source to the teapot was estimated by the user and was in error on the order of $20 \%$. The initial configuration of the area source was significantly larger than that used to render the original image and was placed in a different orientation. The initial solution for $\vec{E}_{\ell} k_{d}$ was obtained by performing a simple segmentation, assuming that the light source colour is white, and found the values for $k_{d}$ that gave the best fit for the image values in the segmented region. This segmentation need not be very precise, as excluding values that should properly belong is not harmful, as long as their is enough variety in the pixel values chosen to give a significant estimate.

If specularities are present, they should be used above diffuse regions to obtain an initial solution as they are less affected by texture. Also in the case of multiple sources, specular highlights are usually formed by one
source while the diffuse signal underneath it is formed by several. In the minimization we only consider diffuse regions since we have no idea if the Phong model can really fit the highlight region.


| $\lambda$ | Ave. error / pixel |  |
| :--- | :---: | :---: |
|  | Ideal Solution | Computed Minimum |
| Red | 8.42 | 8.96 |
| Green | 1.54 | 1.57 |
| Blue | 1.54 | 1.57 |

Figure 4.11: Solving for a rectangular source from the shading on a synthetic object. (left) Original image (right) Rendered image using area source given by minimization. Black squares in the center of each image denotes areas excluded from the computation due to cracks in the patch model.

### 4.3 Solving for Radiometric Properties

This section explores methods for computing the unknown radiometric parameters of light sources and materials when the facing geometry of objects and the geometry of the light sources is known beforehand. In terms of the reshading operations in table 4.1, we know the $F_{i, j}$ but not the $f_{r}$ or $L_{\ell}$.

### 4.3.1 Reshading of Diffuse Surfaces

The value of pixel $P\left(\mathrm{x}_{t}\right)$ formed from reflection from a diffuse surface is given by (from equation ??):

$$
P\left(\mathbf{x}_{t}\right)=\text { Sensor }\left(k_{d} \sum_{\ell=1}^{n} L_{\ell} F_{r, \ell}\right)
$$

With known light source and object geometry, the only unknowns are the emissive strengths of the light sources ( $L_{\ell}$ ) and the reflectance at each point on the surface $\left(k_{d}\right)$. Each pixel in the image gives us one equation with two unknowns. Here we see the inherent ambiguity of emissivity and reflectance: if we shine twice as much light onto a given object and halve its reflectance, it will appear the same in the image. Fortunately, for most reshading operations it is not necessary to determine the absolute values for emission and reflectance, but merely the relative values. In table 4.1, wherever the new image relies upon $L_{\ell} k_{d}$ appearing together we can treat them as a single variable. The relative values of these parameters can be determined if we can locate a suitably large enough image region of uniform reflectance. If we have an image region of $p$ pixels formed by a surface of uniform $k_{d}$, determining the relative $L_{\ell}$ for each light source is a simple task of solving a linear system of equations:

$$
\underbrace{\left[\begin{array}{ccc}
F_{1,1} & \cdots & F_{1, n} \\
F_{2,1} & \cdots & F_{2, n} \\
\vdots & & \vdots \\
F_{p, 1} & \cdots & F_{p, n}
\end{array}\right]}_{\mathbf{A}} \underbrace{\left[\begin{array}{c}
L_{1} k_{d} \\
L_{2} k_{d} \\
\vdots \\
L_{n} k_{d}
\end{array}\right]}_{\mathbf{x}}=\underbrace{\left[\begin{array}{l}
\operatorname{Sensor}^{-1}\left(P_{1}\right) \\
\operatorname{Sensor}^{-1}\left(P_{2}\right) \\
\cdots \\
\operatorname{Sensor}^{-1}\left(P_{p}\right)
\end{array}\right]}_{\mathbf{b}}
$$

Taking the ratio of the various $L_{\ell} k_{d}$ gives us the relative strengths of the light sources. Arbitrarily setting $L_{1}$ to 1 , we can then solve for the $k_{d}$ at each pixel in the image to get an exact fit. This can cause unexpected behaviour from the user's point of view as an "object" will not have a constant value of $k_{d}$ over its surface but will have slight variations. To obtain reflectances in terms of "objects" requires a segmentation.

How can we find a region of uniform $k_{d}$ ? This is the turf of colour image segmentation, for which there are numerous techniques but no solutions. There are two main approaches and many hybrids. One main class are histogram techniques, which group image regions together based on similarity in some measurement space (greyscale, colour, frequency). These techniques are global. The second class of techniques are primarily neighbourhood based which operate solely in the local neighbourhood.

For reshading, the segmentation task is slightly easier in that we do not require the entire image to be partitioned into homogeneous regions, but only need one or a small number of regions to solve for the relative emissivities. If errors are to be performed in segmentation, it is better to be too conservative than overly greedy. Including image regions with varying reflectance will lead to incorrect estimates for $L_{\ell}$, while not getting enough pixels one risks not getting linearly independent system of equations.

There is no one best segmentation strategy, so several have been employed in this work. A good segmentation should handle changes in illumination levels, specular highlights, textures, shadows, and non-connected regions. We have found Klinker et al.'s approach [35] with similarity thresholds set quite low to avoid greediness to be satisfactory. It is good as it includes shadow regions in the segmentation, which is useful for solving for the contribution of indirect light falling on a surface. For textured surfaces a multiresolution approach might be better, as texture is a matter of scale [42]. Methods which give confidence measures with their solutions would be valuable, as then only a region that we are most confident in would be used in the emissivity computation [32]. The more regions used, the more statistically significant the results will be.

The linear system above only deals with direct illumination. The amount of light falling from other surfaces can be quite significant and should be accounted for. At one extreme we could say that the ambient light falling on each pixel is different, but this would lead to an underconstrained system as each equation would add a new unknown. Three ways of handling indirect illumination are

- constant
- piecewise constant
- piecewise smooth

The amount of indirect light falling on a surface depends upon its surrounding environment: the proximity of nearby surfaces, their reflectance, and the amount of light falling on them. If full geometric information about all objects is known, then an estimate of the indirection illumination at some surface area can be estimated using the size and proximity of nearby surfaces, perhaps assuming that the reflectivity of all nonvisible surfaces are the same. With only knowledge of the geometry of the surfaces that can be seen in the image (such as produced by a rangefinder), a more coarse estimate must be made. Indirect illumination depends upon the surrounding environment, which changes when there are abrupt changes in object geometry. So wherever there is a discontinuity in position or orientation in the scene, we treat it as a new region with different indirect illumination properties.

For a constant ambient term, we merely add another light source to the computation above. For a piecewise constant we add one to every region. When the geometric information provided is in the form of a list of faces or polygons finding regions where the indirect illumination value should change is possible since the places where these discontinuities occur is clear. When geometry is given as depth from the eye, this is more difficult as we need to determine a threshold (what a discontinuity means) and need to find bounded regions.

A piecewise smooth ambient term can be achieved by constraining the derivatives of the ambient illumination so that it varies smoothly over smooth surfaces. This problem bears some similarity to the surface reconstruction problem, for which regularization is a common solution strategy [53]. Trying to find the continuous and smooth function $f(x, y)$ which minimizes the fit to the discrete data $d(x, y)$ and satisfies some constraints upon its smoothness:

$$
\int_{x} \int_{y}(f(x, y)-d(x, y))^{2}+\lambda\left(f_{x x}^{2}+2 f_{x y}^{2}+f_{y y}^{2}\right) d x d y
$$

Fitting a smooth ambient term over a surface is akin to fitting a smooth surface to a sampled set of points on the surface. Data points for the ambient component come from shadow regions and from the discrepancy between the intensity signal computed from the emissivity estimates and what appears in the image, although these should be computed simultaneously. Terzopoulos, amongst others, has investigated using regularization in surface reconstruction handling discontinuities [64].

### 4.3.2 Handling Specularities

Handling specular highlights when the object, viewing, and lighting geometry is a tractable problem. With this information the positions on the viewed surfaces where specular highlights may occur can be determined, assuming some suitably large upper bound on how rough a surface can be. With this aid detection strategies should perform better, falling for fewer false matches. Once a highlight is located, the natural operation is to solve for the unknown specular parameters which describes it best. This allows for the removal and subsequent placement of a highlight on some other portion of the surface, and would be extremely useful for processing textures, since the texture image would then be a source of both diffuse and specular reflectance properties. What unknown parameters need to be solved for depends upon the reflection model used. For Phong, $k_{s}$ and $n$ are the only unknowns.

Regardless of the reflection model used, an exact fit between the model and the observed highlight will not occur. If one simply subtracts the colour signal predicted by the model from the image, then the resulting image will not look proper, as either too much or too little will be removed. If a technique is available for removing the highlight more accurately, then one should remove the highlight that way and define a mapping from the changes predicted by the model and the changes made by the technique used. That is, when performing some reshading operation, we do not replace the values in the affected image region, but instead compute a hypothetical image with our model, compute the changes to this image after reshading, and then apply those changes to the original image.

### 4.4 Summary

This chapter has addressed some of the issues involved in realistically altering the appearance of an image when the geometry of the scene is known. Section 4.1 reviewed previous work in this area. Current techniques for determining the geometric and radiometric properties of a single light source from the shading of objects of known structure was covered in Section 4.2. These approaches deal with illuminant models that do not capture the properties of real light sources, so extending these techniques to handle the case of a rectangular source was explored. A solution strategy formulated as a multidimensional minimization incorporating a number of constraints was presented, tested, and found to give encouraging results for the single synthetic case shown. With knowledge of both the illuminant and scene geometry a wide array of reshading operations is possible. Obstacles to performing these and solution strategies was discussed in Section 4.3.

## Chapter Five

## The End



He swung
his face went purple a roar came from the crowd.

Shane MacGowen, Garnet

### 5.1 Conclusions and Contributions

This work is a collection of ideas and is a proof-of-concept rather than a presentation of a working solution. Through these ideas hopefully some contribution has been made. This work has shown that current techniques for the detection and removal of specular highlights do not perform well on textured images, which are precisely those that are used most often in computer graphics. Current algorithms fail on textured images for several reasons. Several approaches classify specularities as violations to a smoothly varying diffuse signal; a texture (as well as abrupt changes in surface orientation) also satisfies this description, and
is erroneously detected and/or removed. Approaches that deal with the colour signal in an image but ignore the spatial distribution of colour changes are also confused by texture, since they cannot distinguish between the commonly well-structured and contiguous changes brought about by a highlight and the potentially haphazard and widely-distributed reflectance variations brought about by a texture. For these reasons a purely local approach will have difficulties with texture. Current approaches rely upon an accurate segmentation of the input image into regions of constant reflectance, which is generally not achievable for even moderately complex textures.

This work attempts to extend previous methods for solving for light source properties from image shading by considering the case of a single rectangular source illuminating the source scene. A series of constraints based on shading cues have been proposed to aid in this task. Solving for the geometric and radiometric properties of the rectangular source was formulated as a multidimensional minimization incorporating these constraints, and was tested on a single synthetic scene. The results of this experiment, preliminary as they are, are very accurate, promising, and give reason to explore this solution strategy further. Methods for solving for the light source emissivities and surface reflectance properties from an image of a scene of known illuminant and object geometry has been presented. Although these ideas are largely untested, the problems faced in this task are not ones of insurmountable theoretical difficulties, but merely ones of time.

### 5.2 Future Work

In terms of specular highlight detection and removal, many areas of exploration suggest themselves. Techniques that work solely in colour space lose spatial information given by the image, and no current technique can handle textures of moderate complexity. Shading changes occur at many scales. A heuristic that may be worth exploring is that changes in a colour signal occur at a number of frequencies, in order from highest to lowest:

- reflectance changes
- shading due to specular reflection
- shading due to diffuse reflection

A frequency based approach to highlight detection and removal seems worth a try. For some textures, standard Fourier analysis may be quite acceptable, while for others some representation of the colour signal
that encodes both frequency and scale information may be better.

For solving for a rectangular source from image shading, there are a number of future extensions. Most importantly, further testing is needed, especially on real images. To get a hold of the problem of locating multiple area sources, one can take advantage of the special constraints present in indoor environments. In such environments many objects are planar and perpendicular. Constraining an area source to be either perpendicular or parallel to a face imposes a valuable constraint on where that source can lie. In addition, asking the user to estimate the distance from a floor to a ceiling is not an overwhelming burden, and this distance can make life easier still.

In solving for scene parameters with a known light source distribution, removal and subsequent reapplication of shading cues such as specular highlights becomes very possible. For some tasks in computer graphics, knowing roughly the surface geometry and the light source direction is possible, so exploring methods for using this information would be useful. Further testing is needed.

## Bibliography

[1] ANSI. Nomenclature and Definitions for Illuminating Engineering. Technical Report ANSI/IES RP-16-1986, American National Standards Institute, 1986.
[2] Arvo, J. "The Irradiance Jacobian for Partially Occluded Polyhedral Sources". In Proc. SIGGRAPH 1994, pages 343-350. ACM, July 1994.
[3] Azuma, R. and Bishop, G. "Improving Static and Dynamic Registration in an Optical See-through HMD". In Proc. SIGGRAPH 1994, pages 197-204. ACM, July 1994.
[4] Bajura, M., Fuchs, H., and Ohbuchi, R. "Merging Virtual Objects with the Real World: Seeing Ultrasound Imagery within the Patient". Computer Graphics (Proc. SIGGRAPH), 26(2):203-210, July 1992.
[5] Bartels, R. H. and Jezioranski, J. J. "Least-Squares Fitting Using Orthogonal Multinomials". ACM Transactions on Mathematical Software, 11(3):201-217, September 1985.
[6] Blinn, J. F. and Newell, M. E. "Texture and Reflection in Computer Generated Images". Communications of the ACM, 19(10):542-547, October 1976.
[7] Blinn, J. F. "Models of Light Reflection for Computer Synthesized Pictures". Computer Graphics (Proc. SIGGRAPH), 11(2):192-198, July 1977.
[8] Brelstaff, G. J. Inferring Surface Shape from Specular Reflections. PhD Thesis, Dept. of Computer Science, University of Edinburgh, August 1989. Available as CST-60-89.
[9] Catmull, E. A Subdivision Algorithm for Computer Display of Curved Surfaces. PhD Thesis, Dept. of Computer Science, University of Utah, December 1974. Available as Report UTEC-CSc-74-133.
[10] Cavanagh, P. and Leclerc, Y. G. "Shape From Shadows". J. of Experimental Psychology: Human Perception and Performance, 15(1):3-27, February 1989.
[11] Cohen, J. "Dependency of the Spectral Reflectance Curves of the Munsell Color Chips". Psychonomic Science, 1:369-370, 1964.
[12] Cook, R. L. and Torrance, K. E. "A Reflectance Model for Computer Graphics". ACM Transactions on Graphics, 1(1):7-24, January 1982.
[13] Devore, J. L. Probability and Statistics for Engineering and the Sciences. Brooks/Cole Publishing Co., Monterey, Cal., 1982.
[14] Duff, T. "Compositing 3-D Rendered Images". Computer Graphics (Proc. SIGGRAPH), 19(3):41-44, July 1985.
[15] Egan, W. G., Hilgeman, T., and Reichman, J. "Determination of Absorption and Scattering Coefficients for Nonhomogeneous Media. 2: Experiment". Applied Optics, 12(8):1816-1823, August 1973.
[16] Fournier, A., Gunawan, A., and Romanzin, C. "Common Illumination Between Real and Computer Generated Scenes". In Proc. Graphics Interface '93, pages 254-262, May 1993.
[17] Gershon, R. The Use of Color in Computational Vision. PhD Thesis, Dept. of Computer Science, University of Toronto, 1987. Available as RBCV-TR-87-15.
[18] Gleicher, M. and Witkin, A. "Through-the-Lens Camera Control". Computer Graphics (Proc. SIG$G R A P H), 26(2): 331-340$, July 1992.
[19] Gunawan, A. S. "Estimating the Illuminant Color of a Scene from its Image Shading". In Proc. 1991 Western Computer Graphics Symposium, pages 29-30, Vernon, B.C., April 8-10, 1991.
[20] Hall, R. Illumination and Color in Computer Generated Imagery. Springer-Verlag, New York, 1989.
[21] Hanrahan, P. and Haeberli, P. "Direct WYSIWYG Painting and Texturing on 3D Shapes". Computer Graphics (Proc. SIGGRAPH), 24(4):215-223, August 1990.
[22] Hanrahan, P. Rendering Concepts. In Cohen, M. F. and Wallace, J. R., editors, Radiosity and Realistic Image Synthesis, pages 13-40. Academic Press, 1993.
[23] Harman, H. H. Modern Factor Analysis. University of Chicago Press, Chicago, 2nd edition, 1967.
[24] Healey, G. "Using Color for Geometry-Insensitive Segmentation". J. Optical Society of America A, 6(6):920-937, 1989.
[25] Heckbert, P. S. Simulating Global Illumination Using Adaptive Meshing. PhD Thesis, Computer Science Division, University of California, Berkeley, June 1991.
[26] Horn, B. K. P. and Sjoberg, R. W. "Calculating the Reflectance Map". Applied Optics, 18:1770-1779, 1979.
[27] Hottel, H. C. and Sarofim, A. F. Radiative Transfer. McGraw-Hill, Toronto, Ontario, 1967.
[28] Hunter, R. S. and Harold, R. W. The Measurement of Appearance. John-Wiley \& Sons, New York, Second edition, 1987.
[29] Ikeuchi, K. and Sato, K. "Determining Reflectance Properties of an Object Using Range and Brightness Images". IEEE Trans. Pattern Analysis and Machine Intelligence, 13(11):1139-1153, November 1991.
[30] Irvin, R. B. and McKeown Jr, D. M. "Methods for Exploiting the Relationships Between Buildings and Their Shadows in Aerial Imagery". IEEE Systems, Man, and Cybernetics, 19(6):1564-1575, November 1989.
[31] Kajiya, J. T. "The Rendering Equation". Computer Graphics (Proc. SIGGRAPH), 20(4):143-150, August 1986.
[32] Kaufman, L. and Rousseeuw, P. J. Finding Groups in Data: An Introduction to Cluster Analysis. John Wiley and Sons, 1990.
[33] Kawai, J. K., Painter, J. S., and Cohen, M. F. "Radioptimization - Goal Based Rendering". In Proc. SIGGRAPH 1993, pages 147-154. ACM, August 1993.
[34] Klinker, G. J., Shafer, S. A., and Kanade, T. "The Measurement of Highlights in Color Images". Int. J. of Computer Vision, 2(1):7-42, January 1988.
[35] Klinker, G. J., Shafer, S. A., and Kanade, T. "A Physical Approach to Color Image Understanding". Int. J. of Computer Vision, 4(1):7-38, January 1990.
[36] Krinov, E. L. Spectral Reflectance Properties of Natural Formations. Technical Report TT-439, National Research Council of Canada, 1947.
[37] Lee, H.-C. "Method for Computing the Scene-Illuminant Chromaticity from Specular Highlights". J. Optical Society of America A, 3(10):1694-1699, October 1986.
[38] Lee, H.-C. Estimating the Illuminant Color from the Shading of a Smooth Surface. Technical Report AI Memo 1068, MIT, August 1988.
[39] Lee, H.-C. "Modeling Light Reflection for Computer Color Vision". IEEE Trans. Pattern Analysis and Machine Intelligence, 12(4):402-409, April 1990.
[40] Lee, S. W. Understanding Of Surface Reflections In Computer Vision by Color and Multiple Views. PhD Thesis, Dept. of Computer and Information Science, University of Pennsylvania, February 1992. Available as MS-CIS-92-13.
[41] Lewis, R. R. "Making Shaders More Physically Plausible". Computer Graphics Forum, 13(2):109120, 1994.
[42] Liu, L. Shade from Shading. Masters Thesis, Dept. of Computer Science, University of British Columbia, August 1994.
[43] Lowe, D. Solving for 3-D Model Parameters from the Locations of Image Features. In Khatib, Craig, and Lozano-Perez, editors, Robotics Review 2, pages 137-143. MIT Press, 1992.
[44] Maloney, L. T. "Evaluation of Linear Models of Surface Spectral Reflectance with Small Numbers of Parameters". J. Optical Society of America A, 3(10):1673-1683, October 1986.
[45] Mitchell, W. J. "When is Seeing Believing?". Scientific American, 270(2):68-73, February 1994.
[46] Nakamae, E., Harada, K., and Ishizaki, T. "A Montage Method: The Overlaying of the Computer Generated Images onto a Background Photograph". Computer Graphics (Proc. SIGGRAPH), 20(4):207-214, August 1986.
[47] Nakamae, E., Ishizaki, T., Nishita, T., and Takita, S. "Compositing 3D Images with Antialiasing and Various Shading Effects". IEEE Computer Graphics and Applications, 9(2):21-29, March 1989.
[48] Nayar, S. K., Ikeuchi, K., and Kanade, T. "Surface Reflection: Physical and Geometrical Perspectives". IEEE Trans. Pattern Analysis and Machine Intelligence, 13(7):611-634, July 1991.
[49] Nayar, S. K., Fang, X.-S., and Boult, T. "Removal of Specularities Using Color and Polarization". In Proc. Computer Vision and Pattern Recognition, pages 583-590. IEEE, 1993.
[50] Nicodemus, F., Richmond, J., Hsia, J., and Ginsberg, I. Geometrical Considerations and Nomenclature for Reflectance. National Bureau of Standards Monograph 160, Washington, D.C. : U.S. Department of Commerce, October 1977.
[51] Novak, C. L. Estimating Scene Properties by Analyzing Color Histograms with Physics-Based Models. PhD Thesis, Dept. of Computer Science, Carnegie Mellon University, December 1992. Available as CMU-CS-92-222.
[52] Phong, B. T. "Illumination for Computer Generated Pictures". Communications of the ACM, 18(6):311-317, June 1975.
[53] Poggio, T., Torre, V., and Koch, C. "Computational Vision and Regularization Theory". Nature, 317:314-319, 1985.
[54] Porter, T. and Duff, T. "Compositing Digital Images". Computer Graphics (Proc. SIGGRAPH), 18(3):253-259, July 1984.
[55] Poulin, P. and Fournier, A. "Lights from Highlights and Shadows". In 1992 Symposium on Interactive 3D Graphics, Computer Graphics, volume 26, pages 31-38. ACM, March 1992.
[56] Poulin, P. Shading and Inverse Shading from Direct Illumination. PhD Thesis, Dept. of Computer Science, University of British Columbia, December 1993.
[57] Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P. Numerical Recipes in C: The Art of Scientific Computing. Cambridge University Press, Second edition, 1992.
[58] Reichman, J. "Determination of Absorption and Scattering Coefficients for Nonhomogeneous Media. 1: Theory". Applied Optics, 12(8):1811-1815, August 1973.
[59] Schoeneman, C., Dorsey, J., Smits, B., Arvo, J., and Greenberg, D. "Painting With Light". In Proc. SIGGRAPH 1993, pages 143-146. ACM, August 1993.
[60] Shafer, S. A. "Using Color to Separate Reflection Components". Color: Research and Application, 10(4):210-218, 1985. Also available as Tech. Report TR 136 (University of Rochester, April 1984).
[61] Shu, J. S.-P. "Cloud Shadow Removal from Aerial Photographs". Pattern Recognition, 23(6):647-656, 1990.
[62] Siegel, R. and Howell, J. R. Thermal Radiation Heat Transfer. Hemisphere Publishing Co., Washington, 2nd edition, 1981.
[63] Sinha, P. and Adelson, E. "Recovering Reflectance and Illumination in a World of Painted Polyhedra". In Proc. Computer Vision and Pattern Recognition, pages 156-163. IEEE, 1993.
[64] Terzopoulos, D. "Regularization of Inverse Visual Problems Involving Discontinuities". IEEE Trans. Pattern Analysis and Machine Intelligence, 8(4):413-424, July 1986.
[65] Thirion, J.-P. "Realistic Three Dimensional Simulation of Shapes and Shadows for Image Processing". CVGIP Graphical Models and Image Processing, 54(1):82-90, January 1992.
[66] Tominaga, S. and Wandell, B. A. "Standard Surface-Reflectance Model and Illuminant Estimation". J. Optical Society of America A, 6(4):576-584, April 1989.
[67] Torrance, K. E. and Sparrow, E. M. "Theory for Off-Specular Reflection From Roughened Surfaces". J. Optical Society of America, 57(9):1105-1114, September 1967.
[68] Touloukian, T. and DeWitt, D. Volume 7, Thermal Radiative Properties: Metallic Elements and Alloys. In Touloukian, Y. and Ho, C., editors, Thermophysical Properties of Matter. IFI/Plenum, New York, 1970.
[69] Touloukian, T. and DeWitt, D. Volume 8, Thermal Radiative Properties: Non-Metallic Solids. In Touloukian, Y. and Ho, C., editors, Thermophysical Properties of Matter. IFI/Plenum, New York, 1970.
[70] Wallace, B. A. "Merging and Transformation of Raster Images for Cartoon Animation". Computer Graphics (Proc. SIGGRAPH), 15(3):253-262, August 1981.
[71] Wang, C., Huang, L., and Rosenfeld, A. "Detecting Clouds and Cloud Shadows on Aerial Photographs". Pattern Recognition Letters, 12:55-64, 1991.
[72] Wanger, L. R., Ferwerda, J. A., and Greenberg, D. P. "Perceiving Spatial Relationships in ComputerGenerated Images". IEEE Computer Graphics and Applications, 12(3):44-58, May 1992.
[73] Ward, G. J. "Measuring and Modeling Anisotropic Reflection". Computer Graphics (Proc. SIGGRAPH), 26(2):265-272, July 1992.
[74] Ward, G. J. and Heckbert, P. S. "Irradiance Gradients". In Third Eurographics Workshop on Rendering, pages 85-98. European Association for Computer Graphics, 1992.
[75] Witkin, A. Recovering Intrinsic Scene Characteristics from Images. Technical Report TR SRI Project 1019, SRI International, September 1981.
[76] Wolff, L. B. "Diffuse Reflection". In Proc. Computer Vision and Pattern Recognition, pages 472-478. IEEE, 1992.
[77] Wolff, L. B., Shafer, S. A., and Healey, G. E. Shape Recovery. Physics-Based Vision: Principles and Practice. Jones and Bartlett Publishers, Boston, 1992.
[78] Yang, Y. and Yuille, A. "Sources From Shading". In Proc. Computer Vision and Pattern Recognition, pages 534-539. IEEE, 1991.

## Appendix A

## Notation

## Radiometric Symbols (Units)

$E \quad$ Irradiance ( $\mathrm{W} / \mathrm{m}^{2}$ )
$B \quad$ Radiosity ( $\mathrm{W} / \mathrm{m}^{2}$ )
$L \quad$ Radiance ( $\mathrm{W} / \mathrm{m}^{2} \mathrm{sr}_{f}$ )
$f_{r} \quad \mathrm{BRDF}\left(\mathrm{sr}^{-1}\right)$
$\Phi \quad$ Radiant flux (W)
$P$ pixel value
$\lambda$ wavelength
$\rho \quad$ reflectance

## Subscripts

## Angles

$\omega \quad$ Solid angle (sr)
$\theta$ polar angle from normal
$\phi \quad$ azimuthal angle
$\varphi \quad$ angle between specular reflection and viewing vectors
$\angle \quad$ local reflectance geometry $\left(\theta_{i}, \phi_{i}, \theta_{e}, \phi_{e}\right)$
$2 \pi$ hemispherical quantity

## Appendix B

## Fresnel Reflection

The percentage of light incident from air at angle $\theta_{i}$ that is reflected specularly by a smooth surface can be derived directly from Maxwell's equations, and is given by [62]

$$
\begin{align*}
\rho_{s \perp}\left(\theta_{i}, \phi_{i} ; \theta_{i}, \phi_{i}+\pi\right) & =\frac{a^{2}+b^{2}-2 a \cos \theta_{i}+\cos ^{2} \theta_{i}}{a^{2}+b^{2}+2 a \cos \theta_{i}+\cos ^{2} \theta_{i}}  \tag{B.1a}\\
\rho_{s| |}\left(\theta_{i}, \phi_{i} ; \theta_{i}, \phi_{i}+\pi\right) & =\frac{a^{2}+b^{2}-2 a \sin \theta_{i} \tan \theta_{i}+\sin ^{2} \theta_{i} \tan ^{2} \theta_{i}}{a^{2}+b^{2}+2 a \sin \theta_{i} \tan \theta_{i}+\sin ^{2} \theta_{i} \tan ^{2} \theta_{i}} \rho_{s \perp}\left(\theta_{i}, \phi_{i} ; \theta_{i}, \phi_{i}+\pi\right) \tag{B.1b}
\end{align*}
$$

with

$$
\begin{aligned}
& 2 a^{2}=\left[\left(n(\lambda)^{2}-\kappa(\lambda)^{2}-\sin ^{2} \theta_{i}\right)^{2}+4 n(\lambda)^{2} \kappa(\lambda)^{2}\right]^{1 / 2}+\left(n(\lambda)^{2}-\kappa(\lambda)^{2}-\sin ^{2} \theta_{i}\right) \\
& 2 b^{2}=\left[\left(n(\lambda)^{2}-\kappa(\lambda)^{2}-\sin ^{2} \theta_{i}\right)^{2}+4 n(\lambda)^{2} \kappa(\lambda)^{2}\right]^{1 / 2}-\left(n(\lambda)^{2}-\kappa(\lambda)^{2}-\sin ^{2} \theta_{i}\right)
\end{aligned}
$$

$n(\lambda)$ is the simple index of refraction for the material. $\rho_{s \perp}(\cdot)$ and $\rho_{s \|}(\cdot)$ are the specular reflectivities for the material irradiated by a polarized wave whose amplitude is perpendicular and parallel to the plane of incidence, respectively. $\kappa(\lambda)$ is the extinction coefficient of the material which describes the amount of attenuation radiation experiences while passing through a unit length through the material. It is composed
of an absorption and a scattering component, depends upon the pressure and temperature of the material, and is inversely proportional to the electrical resistivity of the material ${ }^{\dagger}$. Electrically insulating or dielectric materials have a high electrical resistivity ( $r_{e} \gg 0$ ) and therefore for these materials $\kappa(\lambda) \rightarrow 0$. Taking $\kappa(\lambda)$ to its limit simplifies the reflectivity formulas to

$$
\begin{align*}
\rho_{s \perp}(\cdot) & =\frac{\sin ^{2}\left(\theta_{i}-\theta_{t}\right)}{\sin ^{2}\left(\theta_{i}+\theta_{t}\right)}  \tag{B.2a}\\
\rho_{s \|}(\cdot) & =\frac{\tan ^{2}\left(\theta_{i}-\theta_{t}\right)}{\tan ^{2}\left(\theta_{i}+\theta_{t}\right)} \tag{B.2b}
\end{align*}
$$

where $\theta_{t}$ is the angle of transmission and is related to the angle of incidence through Snell's law:

$$
n(\lambda)=\frac{\sin \theta_{i}}{\sin \theta_{t}}
$$

Equations (B.1) and (B.2) are known as the Fresnel equations. They give the ratio of the total flux reflected from a smooth surface to the flux incident upon the surface from a given direction. For light of no particular polarization, the specular reflectivity is given by an equal weighting of the parallel and perpendicular components:

$$
\rho=\rho_{s}\left(\theta_{i}, \phi_{i} ; \theta_{i}, \phi_{i}+\pi\right)=\frac{\rho_{s\| \|}(\cdot)+\rho_{s \perp}(\cdot)}{2} .
$$

[^12]$$
\kappa(\lambda)^{2}=\frac{\mu \gamma c^{2}}{2}\left(\sqrt{\left(1+\frac{\lambda}{2 \pi c r_{e} \gamma}\right)^{2}}-1\right)
$$

## Appendix C

## The Relationship Between Image Irradiance and Scene Radiance

The irradiance $E_{p_{i}}$ at a sensor element $p_{i}$ is the radiant flux entering that element per unit area:

$$
E_{p_{i}}=\frac{d \Phi_{i}}{d A_{p_{i}}}
$$

where $d A_{p_{i}}$ is the area of the sensor element. Assuming a perfect lens (no loss or distortion), then the radiant flux incident upon $p_{i}$ is the total flux passing through a camera lens destined for $p_{i}$ :

$$
d \Phi_{i}=\int_{S} \int_{\mathrm{lens}} L_{d A_{s} \rightarrow \mathrm{lens}} \cos \beta d \omega_{\mathrm{lens}} d A_{s}
$$

The integrals are over all surfaces $S$ projecting to the sensor element and over the surface of the lens. $L_{d A_{s} \rightarrow \text { lens }}$ is the radiance emitted or reflected from surface element $d A_{s}$ to a portion of the lens subtended by the solid angle $d \omega_{\text {lens }}$. Assuming that the diameter of the camera lens is small compared to the distance from the lens to the objects being imaged allows us to drop the integral over the lens:

$$
\begin{aligned}
d \Phi_{i} & =\int_{S} L_{d A_{s} \rightarrow \mathrm{lens}} \cos \beta d \omega_{\mathrm{lens}} d A_{s} \\
& =\frac{d^{2} \pi}{4} \int_{S} L_{d A_{s} \rightarrow \mathrm{lens}} \cos \beta \frac{\cos ^{3} \alpha}{z^{2}} d A_{s}
\end{aligned}
$$

where $z$ is the distance along the optical axis from the lens to $d A_{s}$, and $\alpha$ is the angle between the optical axis and $d A_{s}$.


Figure C.1: Image forming system.

The solid angle subtended by a surface patch as seen from the lens equals the solid angle of the surface's projection onto the image plane as seen from the lens:

$$
\begin{equation*}
\frac{d A_{s} \cos \beta}{(z / \cos \alpha)^{2}}=\frac{d A_{p_{i}} \cos \alpha}{(f / \cos \alpha)^{2}} \tag{C.1}
\end{equation*}
$$

If only a single infinitesimal surface patch $d A_{s}$ projects to the sensor element, then the integral over $S$ drops, and substituting equation (C.1) produces

$$
\begin{aligned}
& d \Phi_{i}=\frac{\pi}{4}\left(\frac{d}{f}\right)^{2} \cos ^{4} \alpha L_{d A_{s} \rightarrow \mathrm{lens}} d A_{p_{i}} \\
& E_{p_{i}}=\frac{\pi}{4}\left(\frac{d}{f}\right)^{2} \cos ^{4} \alpha L_{d A_{s} \rightarrow \mathrm{lens}}
\end{aligned}
$$

which is the standard result [26]. For those pixels that have large surface areas or multiple objects projecting to them, the integral over $S$ cannot be discounted. $\alpha$ varies across all surfaces, but is constrained to be within the solid angle subtended by $p_{i}$ viewed from the centre of the lens, which may be very small; therefore, $\alpha$ may be usefully approximated with the angle $\bar{\alpha}$ through the centre of the solid angle:

$$
E_{p_{i}}=\frac{d^{2} \pi}{4 d A_{p_{i}}} \cos ^{3} \bar{\alpha} \int_{S} L_{d A_{s} \rightarrow \mathrm{lens}} \cos \beta \frac{1}{z^{2}} d A_{s}
$$

Once again, however, reality laughs at our foolish and naive assumptions. Perfect lens do not exist, and imperfect ones can exhibit a whole slew of abnormalities that may not be radially symmetric.


[^0]:    ${ }^{\dagger}$ e.g. Chair Affair

[^1]:    ${ }^{\dagger}$ Terminology time: hereafter, source image will denote an image or some portion of an image that is being transformed, and source scene will indicate the 3D environment from which it arose. A transformed image will be called a derivative. The desired result (what the source image is being used to create) will be referred to as the target image that arises from some hypothetical target scene. Since this work (futilely) deals with the realistic appearance of objects as brought about by the process of light transport, the domain is restricted to photorealistic images (or as good as it gets).

[^2]:    ${ }^{\dagger}$ Interestingly, one does not have to produce an entirely complete or consistent image in order to fool the viewer, but it is not known just how good a job has to be done. Two studies relevant to determining the relative importance of these criteria are that of Wanger et al., who explored the effect the presence of certain cues have in the understanding of spatial relationships in images [72], and the work of Cavanugh and Leclerc in evaluating the effects of implausible shadows in the perception of depth and shape [10].

[^3]:    ${ }^{\dagger}$ Some nice examples of the effects that can be produced solely with these methods and the effort involved in producing them convincingly is given in Mitchell's recent Scientific American article [45].

[^4]:    ${ }^{\dagger}$ All of the models in this thesis are based on geometric optics as opposed to physical optics. Geometric optics treats light as a particle and is only valid for explaining the gross behaviour of light when the wavelength is very small when compared to the physical dimensions of the irregularities of the surfaces it interacts with. Physical optics is directly based on electromagnetic theory and accounts for the wavelike nature of light, such as diffusion and interference [48].
    ${ }^{\S}$ Unless explicitly shown dependance upon wavelength is implied.

[^5]:    *Convention has this vector pointing outwards from the interior of the frustum, but this leads to negative form factors! I've never figured out why I disagree with the world on this point, but the figure shows the world's view.

[^6]:    ${ }^{\dagger}$ Reichman's view-independent model is an extension of the Kubelka-Munk model, and it assumes that flux striking a scat-

[^7]:    ${ }^{\dagger}$ Richard Hunter began studying gloss in the early 1930s when he developed a glossimeter intended to measure the ability of materials to reflect light specularly at an angle of $45^{\circ}$, and has devoted much of his career to studying the appearance of objects.

[^8]:    ${ }^{\dagger}$ To determine the 'standard' surface reflectance, Gershon draws upon previous work in colour constancy, the area of vision research that addresses the phenomena that humans tend to perceive objects as having a constant colour regardless of the colour of the illuminant. A problem faced by colour constancy researchers was to determine whether it is possible to accurately describe the reflectance spectra of many real objects though a linear combination of a fixed number of basis functions, and if so, how many are needed and which ones. In an attempt to answer these questions Cohen [11] performed a characteristic vector analysis on 150 randomly chosen Munsell colour chips and found that using the first three characteristic vectors as basis vectors accounted for 0.992 of the variance in the spectral data of the chips (This value is the coefficient of multiple determination $R^{2}$ which equals the ratio of the reduction in the sum of squares error in a sample about the proposed model to the total sum of squares error in that sample about the sample's mean, $0 \leq R^{2} \leq 1$ [13]. A value near 1 indicates that the proposed model fits the data nearly identically). Maloney [44] extended Cohen's analysis to 462 Munsell chips and to the reflectance spectra of 337 naturally occurring materials collected by Krinov [36] and found that using Cohen's basis vectors accounted for $0.993,0.979$, and 0.998 of the variance for the median,

[^9]:    first and third quartile fits in Krinov's data set, respectively. Based on these results Gershon defines the standard surface reflectance as the mean of the Krinov data set as approximated with Cohen's first three characteristic vectors.
    ${ }^{\dagger}$ The orientation and extent of the optimally fitting ellipsoid of a dataset is described by the eigenvectors and eigenvalues of the covariance matrix of the dataset [23]. The orthogonal eigenvectors define the axes of the ellipsoid, and their corresponding eigenvalues indicate the amount of variance in the data that is accounted for by that vector. These eigenvectors are the principal components, and representing a point set in terms of its components is equivalent to rotating the axes of the original coordinate system. A dataset can be classified as "linear" or "planar" if its variance is well described by one or two vectors. Note that for a planar cluster the principal components do not define the matte and highlight lines, since these need not be orthogonal.

[^10]:    $V$ denotes all points in the scene
    $P$ pixel value
    $\mathbf{x}_{\ell} \quad$ light emitting point
    $\mathbf{x}_{r} \quad$ point receiving light energy
    $\mathbf{x}_{t} \quad$ target of reflected energy from $\mathbf{x}_{r}$ on image plane

[^11]:    ${ }^{\dagger}$ and everything is defined in terms of the variables $\left(\theta_{c}, \phi_{c}, \phi, \Theta, \vartheta\right)$ as

    $$
    \begin{gathered}
    \Gamma_{i, y}=\left(\frac{\vec{V}_{i} \times \vec{V}_{i+1}}{\left\|\vec{V}_{i} \times \vec{V}_{i+1}\right\|}\right)_{y} \cos ^{-1}\left(\frac{\vec{V}_{i} \cdot \vec{V}_{i+1}}{\left\|\vec{V}_{i}\right\|\left\|\vec{V}_{i+1}\right\|}\right) \\
    \vec{V}_{i}=\vec{C} \pm \vec{R}_{l} \pm \vec{R}_{w} \\
    \vec{C}=D\left(\cos \phi_{c} \sin \theta_{c}, \cos \theta_{c}, \sin \phi_{c} \sin \theta_{c}\right) \\
    \vec{R}_{l}=l(\cos \phi \cos \Theta,-\sin \Theta, \sin \phi \cos \Theta) \\
    \vec{R}_{w}=w(\cos \phi \sin \Theta \sin \vartheta-\sin \phi \cos \vartheta, \cos \Theta \sin \vartheta, \sin \phi \sin \Theta \sin \vartheta+\cos \phi \cos \vartheta)
    \end{gathered}
    $$

    $\vec{V}_{i}$ is the vector from vertex $v_{i}$ of the source to $\mathbf{X}_{r}, \vec{R}_{l}$ is the half-length of the rectangle and $\vec{R}_{w}$ is the half-width.

[^12]:    ${ }^{\dagger}$ Like so:

