Image Analysis using the Physics of Light Scattering

PETER NILLIUS

Doctoral Thesis
Stockholm, Sweden 2004

Akademisk avhandling som med tillstånd av Kungl Tekniska högskolan framlägges till offentlig granskning för avläggande av teknologie doktorsexamen fredagen den 4 juni 2004 kl 10.00 i sal D1, Lindstedtsvägen 17, Kungliga Tekniska Högskolan, Stockholm.

Abstract

Any generic computer vision algorithm must be able to cope with the variations in appearance of objects due to different illumination conditions. While these variations in the shading of a surface may seem a nuisance, they in fact contain information about the world. This thesis tries to provide an understanding what information can be extracted from the shading in a single image and how to achieve this. One of the challenges lies in finding accurate models for the wide variety of conditions that can occur.

Frequency space representations are powerful tools for analyzing shading theoretically. Surfaces act as low-pass filters on the illumination making the reflected light band-limited. Hence, it can be represented by a finite number of components in the Fourier domain, despite having arbitrary illumination. This thesis derives a basis for shading by representing the illumination in spherical harmonics and the BRDF in a basis for isotropic reflectance. By analyzing the contributing variance of this basis it is shown how to create finite dimensional representations for any surface with isotropic reflectance.

The finite representation is used to analytically derive a principal component analysis (PCA) basis of the set of images due to the variations in the illumination and BRDF. The PCA is performed model-based so that the variations in the images are described by the variations in the illumination and the BRDF. This has a number of advantages. The PCA can be performed over a wide variety of conditions, more than would be practically possible if the images were captured or rendered. Also, there is an explicit mapping between the principal components and the illumination and BRDF so that the PCA basis can be used as a physical model.

By combining a database of captured illumination and a database of captured BRDFs a general basis for shading is created. This basis is used to investigate material classification from a single image with known geometry but arbitrary unknown illumination. An image is classified by estimating the coefficients in this basis and comparing them to a database. Experiments on synthetic data show that material classification from reflectance properties is hard. There are mis-classifications and the materials seem to cluster into groups. The materials are grouped using a greedy algorithm. Experiments on real images show promising results.

Keywords: computer vision, shading, illumination, reflectance, image irradiance, frequency space representations, spherical harmonics, analytic PCA, model-based PCA, material classification, illumination estimation

Sammanfattning

Föremål ser olika ut vid olika ljusförhållanden. För att en datorseendealgoritm ska vara generell måste den kunna hantera dessa variationer. Det är lätt att uppfatta den som ett problem, men faktum är att den s.k. skuggningen av en yta eller ett föremål innehåller information om omvärlden. Avhandling innebär ett försök att förstå vilken information som kan utvinnas ur skuggningen från en enskild bild och hur detta kan omsättas i algoritmisk form. En av svårigheterna ligger i att finna modeller som är tillräckligt noggranna och också kan hantera alla de förhållanden som kan förkomma.

Representationer i frekvensrymden är kraftfulla verktyg för att analysera skuggning i teoretisk mening. Ljus som träffar en yta lågpassfiltreras. Det gör det reflekterade ljuset bandbegränsat och därför möjligt att representeras i Fourier-domänen med ett begränsat antal parametrar, trots att ljuskällan kan vara godtyckligt utformad. I avhandlingen härleds en bas för skuggning genom att representera belysningen med klotytsfunktioner och ytreflektansen (BRDF:en) i en bas för isotrop reflektans. Genom att analysera variansen av bidraget från basfunktionerna visas hur en noggrann representation med ett ändligt antal parametrar kan skapas.

Denna ändliga representation används för att analytiskt härleda en principalkomponentsbas för den mängd av bilder som skapas av variationer i belysningen och ytreflektansen. Principalkomponentsanalysen (PCA) görs modellbaserat vilket innebär att variationerna i bilderna beskrivs genom variationerna i belysningen och BRDF:en. Detta medför många fördelar. Principalkomponentanalysen kan genomföras över en stor mängd förhållanden, många fler än vad som är praktiskt möjligt om analysen skulle göras med fotograferade eller genererade bilder. Dessutom finns en explicit avbildning från principalkomponenterna till belysningen och BRDF:en vilket gör att PCA-basen kan användas som fysikalisk modell.

Genom att kombinera en databas med insamlad belysing med en databas av insamlade BRDF:er så skapas en generell bas för skuggning. Denna bas används för att undersöka hur material i en enskild bild med känd scengeometri men okänd godtycklig belysning kan klassificeras. Klassificeringen sker genom att skatta bildens koefficienter i basen och jämföra dessa med en databas. Experiment med syntetiska data visar att det är ett svårt problem. Det förekommer felklassificeringar och materialen tycks bilda grupper inom vilka ett material ofta felklassificeras som ett annat material i samma grupp. En s.k. girig algoritm används för att hitta dessa grupper. Experiment på bilder av verkliga objekt visar lovande resultat.

Sökord: datorseende, skuggning, belysning, reflektans, bildbestrålning, frekvensrymdsrepresentationer, klotytsfunktioner, analytisk PCA, modell-baserad PCA, materialklassificering, belysningsskattning

Acknowledgments

First, I would like to thank my supervisor **Prof Jan-Olof Eklundh** for introducing me to Computer Vision and giving me the opportunity to explore this captivating field. Thank you for your encouragement and sound advice regarding both research and other matters. Your experience has been invaluable at times.

Many other people at CVAP have contributed to this thesis by creating a stimulating and pleasant working atmosphere with discussions and friendship. Thank you **Daniel** for many interesting discussions on research, mathematics and life in general. **Lars B**, what would life be without a good game of boule? Thanks for many good times. **Danny** and **Ivan**, thanks for the time sharing an office. **Lars P**, always up for a coffee (when he's not on the other side of the planet). **Eric**, a man you can count on. Thanks for many excellent dinners! **Johan**, always fun to talk with you. Äta bör man annars dör man and preferably so with friends. Thanks to all the others in the lunch crew, **Stefan**, **Elin**. **Gareth**, come back! The service at Sophiahemmet has not been the same in your absence. Thank you **Barbara**, for creating a little Italy in Gamla Stan. **Mårten**, thanks for help on many matters. There are many more, former and current members of CVAP, that have influenced this work in various ways. Thank you **all**.

Thank you **Josephine**, without your support and help I would probably still be writing this thesis. And thanks for excellent proof reading!

Finally, I would like to express my gratitude to **Pappa**, my sister **Karin** and my brother **Erik**, for their love and support.

Contents

Chapter 1

Introduction

None of our senses provides us with as rich information about the world around us as the visual system. Using our eyes we can navigate through the world and interact with our environment. We detect and recognize objects and people. On a more detailed level our vision system gives us information about the properties of the surfaces in our environment, such as their shape, the material they are made of and their relative position.

The goal of computer vision is to make computers "see". Although it is not necessary for computers to emulate humans in this task, the human and other biological visual systems are a great inspiration to the field. They are after all the only working systems we know of.

Seeing involves acquiring information about the three-dimensional world from two-dimensional images. Information cues include motion fields, stereo disparity, edges, texture, color and shading. Motion and stereo are maybe the most important for our understanding of the structure of the world, but even in the total absence of these cues, the human visual system can acquire a lot of information from an image. This becomes obvious when we think of the information we can deduce from a photo, or a painting such as van Eyck's *The Arnolfini Portrait* in Figure 1.1. Apart from the many objects we detect in the painting, we also get a feeling for the local shape of many of the surfaces in the scene, e.g. the folds in the textile and the cylindrical shape of the brass candelabrum. Furthermore, we classify the different materials the surfaces are made of, such as textile, velvet, fur, brass, etc. We also estimate the nature of the illumination, which in this case is dim indirect lighting probably from the window.

A majority of the work in computer vision is based on edges and corners, but these features cannot explain the perception of local shape or the recognition of materials. This inference is based on the shading of the surfaces. Shading is the gradual changes in brightness due to changes in the surface normal relative to the light source and the image sensor. How the brightness changes depends on the reflectance properties of the surface and the illumination.

Figure 1.1: *The Arnolfini Portrait* by van Eyck exhibiting numerous examples of shading from a variety of materials.

To understand what information can be extracted from shading and how to obtain this, are the motivating factors behind this thesis.

1.1 Physics-Based Vision

Shading is special in that the physics that create the shading is well known. The brightness of a point in the image depends on three components, the surface normal, the surface reflectance properties and the illumination. Simply put, the shading is a product of three factors: the shape of the surface, the surface' reflectance properties and the illumination,

$$\text{Image} = f(\text{Shape}, \text{Material}, \text{Illumination}). \qquad (1.1)$$

The work concerned with trying to recover the information about the scene using the physics of image formation is referred to as physics-based vision, (Wolff, Shafer, and Healey 1992).

Although the physics that generate the shading is well known, inferring information from it is far from easy. The problem is highly under-constrained. In order to do any inference on the properties of the scene we need to add assumptions. The brightness of a pixel is determined by equation (1.1). It is a function of the surface normal, the surface reflectance properties and the illumination *at that point*. First we must find several points in the image where the surface reflectance does not change, i.e. when a surface is made from a homogeneous material such as on one of the textiles in the van Eyck painting. If the surface reflectance properties change we cannot in general do inference based on the shading. The most obvious example is a photo or a painting, where the brightness and color changes in such a way that it appears like a three dimensional scene, something that Magritte pointed out, see Figure 1.2.

Another common assumption is that the illumination is distant, relative to the points in the image, so that the illumination for a number of points in the image can be assumed constant. This assumption is broken when part of the segment is shadowed, when there are strong local inter-reflections or when the light source is near the surface.

To do inference based on the shading in the image, it is necessary to take all three components of image formation into account. If all three components are unknown one has to estimate all three. Usually it is assumed that one or more of the parts are given. In the pioneering work of Horn (1970, 1975) it is assumed that the surface has Lambertian reflectance and the illumination is a point light source at a known position. From this the shape of the surface is recovered. This is what is known as shape-from-shading (SFS). A lot of work has been devoted to SFS and it is not our intention to review it here since it is not the main topic of the thesis. A collection of relevant papers can be found in the book by Horn and Brooks (1989). A more recent survey can be found in (Zhang et al. 1999).

Figure 1.2: *Ceci n'est pas une pipe* by Magritte. To be able to do inference from shading it is necessary that the reflectance properties of the surface do not change. Even a flat surface can be made to look like a three dimensional object if painted appropriately.

The goal of this thesis has been to develop methods that can be used in realistic conditions, i.e. non-Lambertian surface reflectance and complex illumination. The main contributions are on models and representations of shading that can be used under such conditions.

1.1.1 Models and Representation

Physics-based vision is inherently model-based. In order to do inference it is necessary to model the components of the image formation process.

1.1.1.1 Surface Reflectance Functions

Given an opaque surface, light that hits the surface is scattered in different directions. How the surface scatters the light is described by the bidirectional reflectance distribution function (BRDF) (Nicodemus, Richmond, Hsia, Ginsberg, and Limperis 1977), which tells how much light is reflected in a direction due to the incident light from another direction.

There are several models of the BRDF for natural surfaces. The most commonly used is Lambert's law. A Lambertian surface is an ideal diffuser. It scatters the light equally in all directions and hence it appears equally bright from all viewpoints. The reason it is used is probably due to its simplicity, rather than how well it represents the world. Very few materials follow Lambert's law perfectly.

Many materials exhibit a shine, also called specularity. Part of the light does not penetrate the surface and is instead reflected in the mirrored incident direction. The specularity can be increasingly diffuse depending on the surface roughness. Torrance and Sparrow (1967) developed a model for this phenomenon, which can be combined with a diffuse (Lambertian) part (Cook and Torrance 1981).

Oren and Nayar (1995) develop a more realistic model for diffuse reflectance by taking into account self-shadowing and interreflections in the microstructure of rough surfaces.

1.1.1.2 Modeling the Illumination

One of the hardest factors to model has been the illumination. Most work on shading have used a simplified model such as a point light source. A point light source might work well as an approximation to direct sunlight, but it is inadequate for most other situations.

The illumination in its fullest description is a function on a sphere, returning the intensity of the incident light from every direction. However, this is not a useful model for inference since we can never expect to recover this function from the shading alone. The problem needs to be regularized.

An important notion is that the surface acts as a low-pass filter on the incident illumination (Cabral, Max, and Springmeyer 1987; D'Zmura 1991). This makes the reflected light band-limited, hence suitable to represent in the Fourier domain. D'Zmura expresses the shading equations in spherical harmonics, which are the sphere's analog to the Fourier series on a line, and discusses the ambiguities that arise in the appearance of a surface. There are several illumination-reflectance pairs that result in to the same shading.

Recent work uses spherical harmonics to analyze the reflected light from a Lambertian surface. Basri and Jacobs (2001a, 2003) and Ramamoorthi and Hanrahan (2001a) simultaneously published work that analyze the properties of the Lambertian reflectance function in spherical harmonics. They showed that the reflected light can be approximated as a linear combination of the first nine spherical harmonic components of the illumination and this under more or less any illumination. This means that, in the case of Lambertian surfaces we can use the first nine spherical harmonic coefficients as a model for the illumination.

Surfaces with other reflectance properties may also be analyzed in the frequency domain. The shinier a surface is the more the appearance of that surface depends on higher frequencies of the illumination. Hence, more harmonic coefficients are needed to represent the illumination. Ramamoorthi and Hanrahan (2001b) write the surface reflectance function and illumination in spherical harmonics and examines the properties of a number of reflectance models used in computer graphics and vision, such as the Phong model and the Torrance and Sparrow model (1967).

1.1.1.3 Modeling Shading

Given the illumination, the surface reflectance and the shape of the surface, the shading is given by the image irradiance equation (Horn 1986). One can also examine shading from an appearance point of view. What is the general appearance of a surface? What is the space of images of a surface under varying illumination?

When considering the images of a surface, the variability of the images depends on the surface's shape. For example, a planar surface will appear equally bright everywhere (assuming distant illumination and no cast shadows), hence the images can be represented using only a single parameter. A spherical surface has all the visible surface normals represented in the image, making the images of a sphere lie in a higher dimensional space. In fact Belhumeur and Kriegman (1998) show that the dimensionality of this space, even in the case of Lambertian surfaces, has the same dimensionality as the number of unique surface normals in the image. However, empirically it has been shown that the images due to variation in the illumination lie in a much lower dimensional sub-space. Hallinan (1994) and Epstein et al. (1995) perform principal component analysis (PCA) on sets of images of faces and objects with a varying light source and find that around five components are enough to capture most of the variance in the images.

The spherical harmonic representation in (Basri and Jacobs 2003; Ramamoorthi and Hanrahan 2001a) can be used to represent shading. Ramamoorthi (2002) uses this representation to analytically derive the principal components. He studies the space of images of a convex Lambertian surface illuminated by a distant point light source from all directions. The assumption of a convex object is often made to avoid the problems with cast shadows. He computes the PCA bases for several shapes including a sphere and a human face and gives a theoretical explanation to the empirical results that the images can be represented by a low-dimensional basis with about five basis functions.

The spherical harmonic representations provide a powerful framework to analyze shading theoretically. Since they allow arbitrary illumination, they are also useful in practice and have successfully been applied to a number of computer vision problems, such as photometric stereo (Basri and Jacobs 2001b), object recognition (Osadchy, Jacobs, and Ramamoorthi 2003), shape from motion (Simakov, Frolova, and Basri 2003) and face recognition (Zhang and Samaras 2003).

1.2   The Present Work

Motivating this work is an attempt to understand what information can be extracted from shading and how to do it. I have focused on analyzing single images, because of the fact that humans have no problems in interpreting such images, but also because I believe that many of the challenges in analyzing shading are not substantially helped by using more images.

When analyzing single images it is important to take into account the ambiguities that occur. Some of these ambiguities are easily identified in the frequency domain. The frequency space based representation also allows modeling of more or less any illumination situation, which is a major advance from the previous point light source models.

The thesis is divided in two parts. The first part deals with representations of shading. It presents a framework on how to compute principal component bases for

the appearance of surfaces. Using the frequency space representation, the image irradiance can be written as a finite dimensional linear model, despite having arbitrary illumination. This finite dimensional model makes it possible to analytically derive the principal components of the image irradiance.

By analytically deriving the principal components it is possible to combine the advantages of appearance based models and the physics based models. There is an explicit relation between the principal components, capturing the appearance, and the coefficients of the illumination and surface reflectance.

The principal components are constructed so that they maximize the variance in the image. The first component has the highest variance. The subsequent components variances are maximized while being orthogonal to all previous components. The property of having high variance is of great importance when estimating the coefficients of the components. A component with high variance has a high signal-to-noise ratio (SNR), hence is robust to estimate. In other words, the principal component representation allows us to analyze how many components that can be robustly estimated. Furthermore, because the basis is analytically derived and there is an explicit relation to the illumination and surface reflectance, we can interpret the estimated principal coefficients with regards to the illumination and surface reflectance and the different uncertainties and ambiguities that occur.

The second part of the thesis presents algorithms that extract data from shading in images. The first algorithm is a demonstration of the principal component framework presented in part I. In this algorithm we investigate the feasibility of classifying the material of the surface, when the illumination is unknown, but the shape of the surface is known. Knowing the surface shape may seem as an unrealistic assumption, but the algorithm is not intended as a final solution but rather for use in a probabilistic framework, where in the end, the information is extracted when none of the properties of the scene are known beforehand. Note that we do not need to know the shape to use such an algorithm, we only need an estimate of how likely the different shapes are given an image. This can be seen by looking at the factorization of the posterior distribution for the illumination L, the material M and the shape S given the image I.

$$P(L, M, S|I) = P(L, M|S, I)P(S|I) \qquad (1.2)$$

The classification is done by building a low-dimensional basis for the appearance of the shape. This basis can now represent the appearance of this shape with a wide variety of surface reflectance properties and under more or less any illumination. Given a material of the surface the images form a manifold in the space of the PCA basis. We learn these manifolds using captured real world illumination and real world surface reflectance properties. An image is classified by finding the manifold closest to the point representing the image.

The second algorithm is about illumination estimation. This work was completed when the frequency space representations of shading were largely unknown to the computer vision community and therefore the algorithm uses conventional

models such as a point light source and Lambertian surface reflectance. The goal of this work is to completely automatically extract information about the light source. The algorithm uses a model for the shading near occluding contours. The occluding contour is where an object occludes itself like the earth at the horizon. At the occluding contour the shape of the object can easily be determined because the surface normal is perpendicular to the viewing vector and can be determined by the image edge direction. The algorithm works by extracting a number of candidate contours using color and edge information. Now, if the light source is distant, the shading at the occluding contours should have a regularity since they share the same light source. We exploit this regularity and simultaneously estimate the light source direction and determine which contours are occluding ones, using Bayesian network inference.

1.3   Contributions

This section summarizes the contributions of the thesis. The frequency space representations of shading provide a powerful framework for analyzing shading theoretically. This thesis derives a new basis for shading by representing the illumination in spherical harmonics and the BRDF in the basis of Koenderink and van Doorn (1998) for isotropic surfaces. This basis is analyzed and it is shown how to compute which frequencies of the illumination that contribute to the shading for any isotropic material. Furthermore, the analysis leads to a finite dimensional representation of shading under arbitrary illumination.

Moreover, the thesis presents a novel framework for creating low-dimensional generative models of shading. The models are created using PCA. The PCA is derived analytically and in a model-based fashion so that an explicit relation between the principal components and the physical properties of the surfaces can be found. Some insights regarding how the PCA should be performed in terms of centering of the dataset is provided and a modification is suggested. A number of PCA bases are computed and analyzed.

The model-based PCA framework is used to create a basis for a database of materials. Using this basis the material is classified from a single image when the shape of the object is known and under arbitrary unknown illumination conditions. Experiments are performed on real world images taken under varying natural illumination conditions.

Finally a new algorithm that automatically estimates the slant direction to the light source is presented. By combining simple perceptual grouping, Bayesian network inference and the physics of light scattering, the algorithm classifies contours as occluding or not and simultaneously estimates the slant direction to the light source.

1.4 Thesis Overview

The thesis is divided in two parts. Part one deals with models and representations of image irradiance. Chapter 2 describes the basics of image formation theory, necessary to understand the following chapters. Chapter 3 presents and analyzes the frequency space representation for the image irradiance used in the thesis. Chapter 4 describes how to analytically derive principal components in the frequency space representations when there are variations in both the illumination and the surface reflectance properties. In Chapter 5 a number of bases for Lambertian surfaces are computed and compared. In Chapter 6 a basis for a group of materials is computed and analyzed.

The second part contains examples of algorithms to extract information from images using the physics of light scattering. Chapter 7 contains experiments on material classification of an object of known shape, with unknown natural illumination. Chapter 8 presents an algorithm that automatically estimates the light source slant direction using the shading near occluding contours.

Finally, Chapter 9 concludes the thesis.

1.5 List of Publications

Most of the work presented in the thesis has previously appeared in the following publications:

- Nillius, P. and Eklundh, J.O., "Automatic Estimation of the Projected Light Source Direction", In *Proc. IEEE Computer Vision and Pattern Recognition*, December 2001, pp. I:1076-1083

- Nillius, P. and Eklundh, J.O., "Low-Dimensional Representations of Shaded Surfaces under Varying Illumination", In *Proc. IEEE Computer Vision and Pattern Recognition*, June 2003, pp. II:185-192

- Nillius, P. and Eklundh, J.O., "Phenomenological Eigenfunctions for Image Irradiance", In *Proc. International Conference on Computer Vision*, October 2003, pp. 568-575

- Nillius, P. and Eklundh, J.O., "Classifying Materials from their Reflectance Properties", In *Proc. European Conference on Computer Vision*, May 2004

Other articles by the author that are not directly related to the topic of the thesis are:

- Nillius, P. and Eklundh, J.-O., "Fast Block Matching with Normalized Cross-Correlation using Walsh Transforms", Technical Report ISRN KTH/NA/P--02/11--SE, Sept. 2002.

- Roobaert, D., Nillius, P. and Eklundh, J.-O., "Comparison of learning approaches to appearance-based 3D object recognition with and without cluttered background", In *Proc. 4th Asian Conference on Computer Vision*, January 8-11, 2000, Taipei, Taiwan, pp. 443-448

Part I

Representations of Image Irradiance

Chapter 2

Image Formation

In order to analyze images using the physics of light scattering, we need to know how the pixel values in the image are related to the physics of the world. This chapter presents the basics of image formation necessary for the analysis performed in the following chapters.

There are two fundamental aspects to image formation: a geometric and a radiometric. The geometry determines where a point in the scene is projected in the image. The radiometry determines how bright that point will be. This chapter describes the radiometric aspects of image formation.

2.1   Radiometry

Radiometry is the field of measuring light. There are two important concepts in radiometry: *radiance* and *irradiance*.

Radiance is the power of light emitted from a surface in a particular direction, per unit foreshortened area and per unit solid angle. It is measured in watts per square meter per steradian ($Wm^{-2}sr^{-1}$). This somewhat awkward definition stems from the fact that radiance is related to how bright a surface appears.

Irradiance is the power of the light falling onto a surface. It is expressed in watts per square meter (Wm^{-2}). The irradiance is important because it is what the image sensor measures.

2.2   Image Formation

The brightness of a pixel is determined by the *image irradiance* on the corresponding element of the image sensor. But what does this correspond to in the scene? Using a single lens camera model Horn (1986) shows that the image irradiance, E
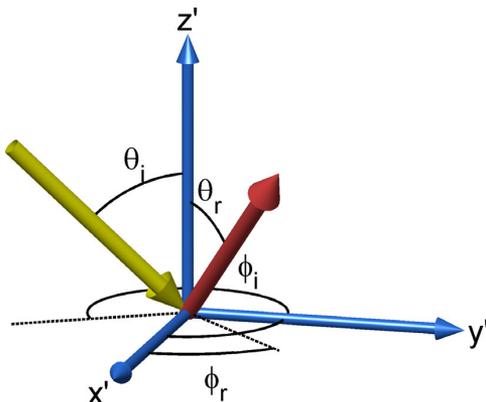
Figure 2.1: The incident and reflected directions.

is approximately proportional to the scene radiance, L.

$$E \propto \cos^4 \alpha \, L, \tag{2.1}$$

where α is the angle between the ray from the scene to the image, and the optical axis. When the camera-object distance is much greater than the object size the $\cos^4 \alpha$ factor is negligible.

Another factor that influences this relationship is vignetting. A camera with multiple lenses will in effect have multiple apertures which screen off light coming in from an angle, resulting in darker pixels around the edges of the image.

Vignetting and the $\cos^4 \alpha$ factor are both deterministic and can be compensated for. In this thesis we ignore their effects and assume that the image irradiance is proportional to the scene radiance, a common assumption in most computer vision algorithms.

2.3   Surface Scattering

The image irradiance is proportional to the scene radiance, but what is the scene radiance? This leads us to surface scattering.

Light that hits a surface is reflected, scattered or absorbed. How the light is scattered depends on the incoming angle of the light and the reflectance properties of the surface. The reflectance properties of the surface are described by its bidirectional reflectance distribution function (BRDF), (Nicodemus et al. 1977).

The BRDF describes the ratio of radiance, dL_r in the reflected direction (θ_r, ϕ_r) to the irradiance, dE_i due to the incident light from direction (θ_i, ϕ_i), see Figure 2.1. Thus, it is a function of two directions.

$$f_r(\theta_i, \phi_i, \theta_r, \phi_r) = \frac{dL_r(\theta_r, \phi_r)}{dE_i(\theta_i, \phi_i)} \tag{2.2}$$

2.3.1   Helmholtz Reciprocity

An important and very general symmetry of the BRDF is *Helmholtz's reciprocity*. Helmholtz's reciprocity states that the ratio of incident irradiance and reflected radiance is the same if you swap the incident and reflected directions, i.e.

$$f_r(\theta_i, \phi_i; \theta_r, \phi_r) = f_r(\theta_r, \phi_r; \theta_i, \phi_i). \tag{2.3}$$

This symmetry holds for all materials, except in special cases such as fluorescence (Minnaert 1941). Therefore, without loss of generality we can incorporate it in our representations. The number of parameters now needed is halved, (Westin, Arvo, and Torrance 1992; Koenderink and van Doorn 1998).

2.3.2   Isotropy

Another important property of the BRDF is isotropy. Surfaces with isotropic reflectance have, in a statistical sense, no direction on the surface. The distribution of the microstructure of the surface constituting the BRDF is independent of the direction on the surface. In such cases the BRDF's only azimuthal dependence is on the absolute difference of the incident and reflected directions,

$$f_r(\theta_i, \phi_i; \theta_r, \phi_r) = f_r(\theta_r, \theta_i, \Delta\phi_{ir}). \tag{2.4}$$

where $\Delta\phi_{ir} = |\phi_i - \phi_r|$, (Koenderink and van Doorn 1998).

Not all materials have isotropic surface reflectance, but many do and in some cases the surface reflectance can be approximated as isotropic, (Dana et al. 1999). Examples of anisotropic materials are surface that have been brushed in a particular direction or certain textiles such as corduroy.

2.4   Scene Radiance

The scene radiance is the sum of all light emitted in the direction towards the camera. The incident light, or the illumination at a point is a function on a sphere returning the incident radiance in each direction, $L_i(\theta_i, \phi_i)$. This incident radiance gives rise to irradiance, $dE_i(\theta_i, \phi_i)$, depending on the incident polar angle (due to the foreshortening effects on light coming in from an angle), $dE_i(\theta_i, \phi_i) = L_i(\theta_i, \phi_i) \cos\theta_i d\omega_i$. This irradiance on the other hand gives rise to scene radiance according to the BRDF, $dL_r = f_r(\theta_i, \phi_i; \theta_r, \phi_r) L_i(\theta_i, \phi_i) \cos\theta_i d\omega_i$. The total scene radiance is computed by integrating this radiance over the hemisphere of possible incident directions, $\mathbf{H_{S^2}}$,

$$L_r(\theta_r, \phi_r) = \int_{\mathbf{H_{S^2}}} L_i(\theta_i, \phi_i) f_r(\theta_i, \phi_i; \theta_r, \phi_r) \cos\theta_i d\omega_i. \tag{2.5}$$

$d\omega_i$ is the infinitesimal solid angle equal to $d\omega_i = \sin\theta_i d\theta_i d\phi_i$.

2.5   Image Irradiance

The image irradiance, which in the end determines the pixel value, is assumed to be
proportional to the scene radiance. Because there always will be an unknown scale
factor in the illumination, we let the illumination take care of the scene radiance-
image irradiance proportionality factor and simply write the image irradiance as

$$E(\theta_r, \phi_r) = \int_{\mathbf{H_{S^2}}} L_i(\theta_i, \phi_i) f(\theta_i, \phi_i, \theta_r, \phi_r) \cos \theta_i d\omega_i. \tag{2.6}$$

2.5.1   Distant Light Source Assumption

When analyzing local parts of the image, so that the distance to the light source(s)
is large relative to the size of the analyzed area, the incident light field $L_i(\theta, \phi)$ can
be assumed constant in that area. This is a common assumption of many computer
vision algorithms.

   In (2.6) the light field is in a local coordinate frame, set by the surface normal.
To compare image irradiance from points with different surface normals under the
same illumination we need to define the light field w.r.t. a global coordinate frame.
To achive this, the incident direction in the image irradiance equation (2.6) is
rotated from the local to the global frame.

   There are several ways of parameterizing 3D rotations. In this thesis we use the
ZYZ Euler angle parameterization, which corresponds to three consecutive coun-
terclockwise rotations around the z-, y- and z-axis respectively (Craig 1989).

$$R_{\alpha,\beta,\gamma} = R_Z(\beta) R_Y(\alpha) R_Z(\gamma) \tag{2.7}$$

This parameterization has an intuitive interpretation. Consider two coordinate
frames arranged as in Figure 2.2b. (α, β) are the spherical coordinates of z' in
the non-primed system. γ is the rotation of the $\{x', y', z'\}$ system around the z'.
Spherical coordinates in these two systems are related as follows.

$$
\begin{aligned}
(\theta, \phi) &= R_{\alpha,\beta,\gamma}(\theta', \phi') & (2.8) \\
(\theta', \phi') &= R_{\alpha,\beta,\gamma}^{-1}(\theta, \phi) & (2.9)
\end{aligned}
$$

 Now, let the global coordinate frame be positioned with its z-axis along the optical
axis pointing towards the camera and the x- and y-axes in the image plane. For
a point with surface normal located (α, β, γ) relative the global frame the image
irradiance can be written as

$$E(\alpha, \beta, \gamma; \theta_r, \phi_r) = \int_{\mathbf{H_{S^2}}} L(R_{\alpha,\beta,\gamma}(\theta_i, \phi_i)) f_r(\theta_i, \phi_i; \theta_r, \phi_r) \cos \theta_i d\omega_i \tag{2.10}$$

where $L(\theta, \phi)$ is the incident light in the global frame. The parameter γ defines the
rotation of the local frame around the surface normal. It should be set so that the
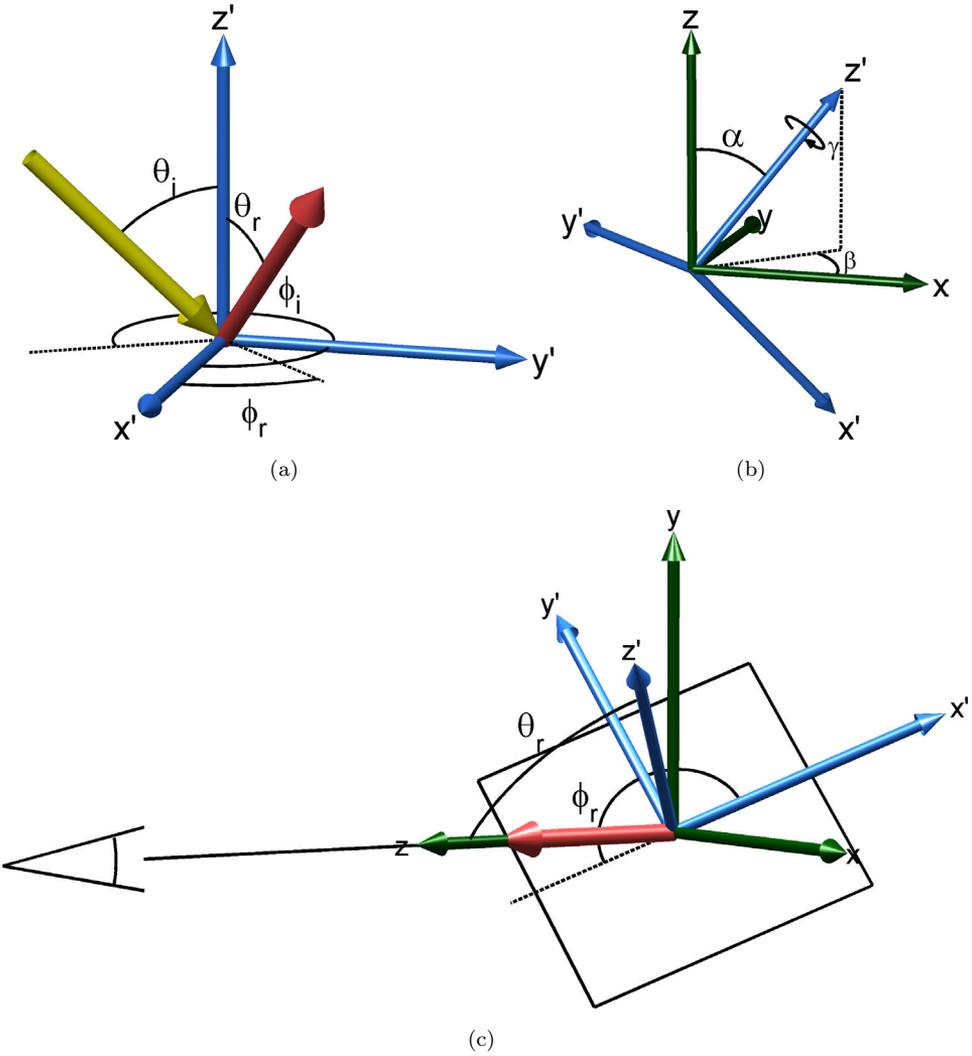
Figure 2.2: Geometry of reflectance. a) The incident and reflected beam in the local coordinate frame. b) The relative position of the global and local coordinate frames. c) Assuming orthographic projection, the reflected beam aligns with the optical axis leading to $\theta_r = \alpha$ and $\phi_r = \pi$.

local frame aligns with the "direction" of the material on the surface (to correspond with the parameterization of the BRDF). Isotropic materials have no direction. In that case, γ can be selected arbitrarily. Hence, we can set it to $\gamma = 0$.

2.5.2   Orthographic Projection

If the camera projection can be approximated as orthographic, i.e. the perspective effects are negligible, the viewing geometry has further implications for the parameters of (2.10). As can be seen in Figure 2.2c the reflected beam aligns with the optical axis (the z-axis in the global frame) which means that $\theta_r = \alpha$. Furthermore, setting $\gamma = 0$ leads to $\phi_r = \pi$. So, for isotropic materials we can eliminate θ_r and ϕ_r from the equation and the image irradiance is uniquely determined by the surface normal (α, β),

$$E(\alpha, \beta) = \int_{\mathbf{H_{S^2}}} L(R_{\alpha,\beta,0}(\theta_i, \phi_i)) f(\theta_i, \phi_i, \alpha, \pi) \cos \theta_i d\omega_i. \tag{2.11}$$

This is what Horn calls the reflectance map, (Horn 1977; Horn and Sjoberg 1979). For Lambertian surfaces, which appear equally bright from all directions, the reflectance map always exists. However, for general BRDFs we need to assume orthographic projection. In perspective projection the image irradiance also depends on the viewing geometry. Then the five (which reduces to four for isotropic surfaces) parameter function in (2.10) is needed to describe the image irradiance.

Chapter 3

Frequency Space Representations of the Image Irradiance

In order to do inference from the image irradiance it is necessary to have a parametric model of the illumination and surface reflectance. The frequency space representations look very promising for this purpose. Although, there are a lot variation in the reflectance functions many of them are smooth and can therefore be represented in frequency space with a limited number of parameters. Also, since surfaces act as low-pass filters on the illumination, the reflected light is band-limited and can be represented in frequency space with a finite number of parameters despite the infinite variations the can occur in the illumination.

This chapter derives a frequency space basis for the image irradiance. This is done by expressing the illumination and the BRDF by their frequency space representations and inserting them into the image irradiance equation (2.11). The resulting basis can represent the image irradiance from any isotropic surface under any illumination. Of course, the representation in the general case is an infinite sum, but because many surfaces act as low-pass filters on the illumination this sum can be truncated and still be an accurate representation of the image irradiance.

3.1   Illumination in Spherical Harmonics

The illumination in the general case can be represented as a function on the sphere. This function is sometimes referred to as the light field. The frequency space basis for functions defined on a sphere are spherical harmonics. Spherical harmonics on the sphere are the analog to the Fourier series on the line. They form an orthonormal basis for functions on the sphere.
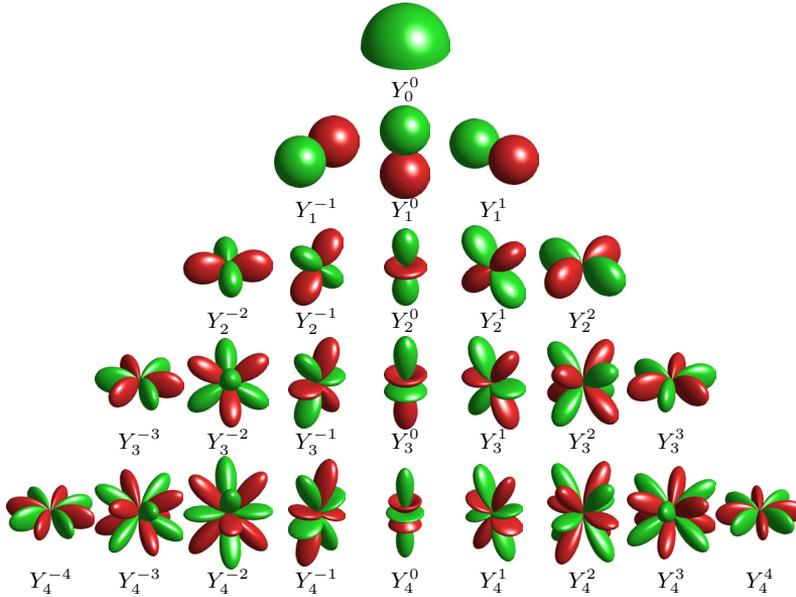
Figure 3.1:   The spherical harmonic basis functions. The radius is equal to the absolute value of the functions. Green (bright) depicts a positive value and red (dark) a negative value.

The light field written in spherical harmonics is

$$L(\theta, \phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} L_l^m Y_l^m(\theta, \phi), \tag{3.1}$$

where L_l^m are the spherical harmonic coefficients of the light field and $Y_l^m(\theta, \phi)$ are the spherical harmonic basis functions, see Figure 3.1. Since the basis functions are orthonormal the coefficients can be calculated as[1]

$$L_l^m = \int_{S^2} L(\theta, \phi) Y_l^m(\theta, \phi) d\omega = \int_0^{2\pi} \int_0^{\pi} L(\theta, \phi) Y_l^m(\theta, \phi) \sin \theta d\theta d\phi \tag{3.2}$$

One important property of spherical harmonics is that they behave nicely under rotation. If a basis function Y_l^m is rotated, the rotated version can be described as a linear combination of the basis function of the same polar order l. The sets of basis functions with same polar order l form sub-groups under 3D rotations. How to linearly combine the basis functions to form 3D rotations can be computed

---

[1]Note that we are using real valued spherical harmonics and therefore we don't need complex conjugate on Y_l^m.

directly from the ZYZ-Euler angles. In particular the light field in (2.11) can be written as

$$L(R_{\alpha,\beta,\gamma}(\theta_i, \phi_i)) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \sum_{n=-l}^{l} D_l^{mn}(\alpha, \beta, \gamma) L_l^m Y_l^n(\theta_i, \phi_i), \tag{3.3}$$

where $D_l^{mn}(\alpha, \beta, \gamma)$ are the rotation reparameterization functions, sometimes referred to as the Wigner D-functions from quantum mechanics, (Chirikjian and Kyatkin 2001). Spherical harmonics are usually defined as complex functions. This thesis uses their real versions. How these are defined and rotated is described in Appendix A.

3.2 BRDF Representation

The BRDF also has a frequency space representation. In this case the BRDF is represented as a sum of coefficients and basis functions of increasing frequency. As the BRDF in many cases is a smoothly varying function it can be well represented with a finite number of basis functions.

Although, in principle, physical models are preferred when representing the BRDF, such models are not always available. For instance many of the materials in the CUReT database (Dana et al. 1996) are not modeled well using a combination of the Oren-Nayar (1995) model for diffuse reflectance and the Torrance-Sparrow model (1967) for specular reflectance. While an eighth order frequency space model of Koenderink and van Doorn (1998) can capture the reflectance properties of most of the different materials in this database.

Even though the Frequency space bases are not derived from physically motivated constraints, they can be constructed to fulfill the general Helmholtz's reciprocity and the quite common isotropy property.

There are several ways of creating frequency space representations for the BRDF. The BRDF is a function from two hemispheres (incident and reflected direction) to $\mathbb{R}$. The natural way to construct a representation is to use a basis for the hemisphere. Since such a basis has not been available spherical harmonics (which are the Fourier basis for the sphere) have been used instead. Westin, Arvo, and Torrance (1992) construct a basis for the hemisphere, using spherical harmonics, by mirroring the function on the lower hemisphere, thereby reducing the number of basis functions by a factor of two. Ramamoorthi and Hanrahan (2001b) instead let the BRDF be a function from two spheres, but fit the model to data so that the lower hemispheres are zero.

Koenderink and van Doorn (1998) use the Zernike polynomials to create a basis for the BRDF. The Zernike polynomials can be used to create a basis for the unit disk, frequently used in optics. By mapping the basis for the unit disc onto the hemisphere Koenderink and van Doorn (1998) create an orthonormal basis for the hemisphere, see Figure 3.2a.

(a)

(b)
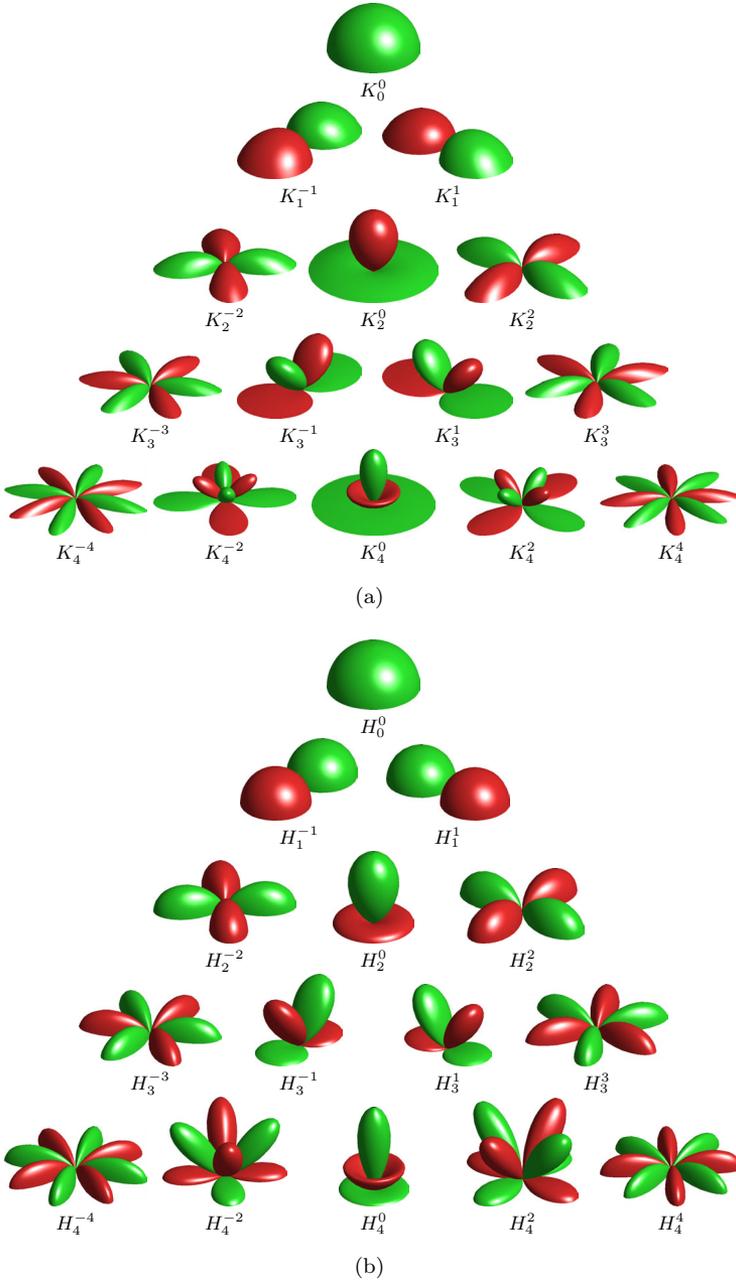
Figure 3.2: Orthogonal bases for the hemisphere based on a) the Zernike polynomials and b) spherical harmonics.

Given a basis for the hemisphere, a basis for the BRDF can be created by multiplying the basis functions of two independent hemispheres. The degrees of freedom of this basis can then be reduced by enforcing the various symmetries of the classes of BRDFs you are interested in representing. For instance, enforcing Helmholtz's reciprocity reduces the number of basis functions with a factor two. Incorporating isotropy drastically reduces the number of functions, since the azimuthal order of the two hemispheres then must be the same. For more on how to create bases for BRDFs see (Westin, Arvo, and Torrance 1992; Koenderink and van Doorn 1998; Ramamoorthi and Hanrahan 2001b).

Which of the frequency representations that is superior is too early to say. Figure 3.2 displays two orthogonal bases for the hemisphere based on the Zernike polynomials and spherical harmonics. Qualitatively they look very similar, but which one that best fits the real world data has not yet been shown.

We use the isotropic BRDF representation of Koenderink and Van Doorn (1998) constructed with the Zernike polynomials. It offers a compact orthonormal representation of the BRDF while incorporating Helmholtz reciprocity and isotropy. The BRDF is represented as

$$f(\theta_i, \theta_r, \Delta\phi_{ir}) = \sum_{opq} b_{op}^q I_{op}^q(\theta_i, \theta_r, \Delta\phi_{ir}), \tag{3.4}$$

where

$$
\begin{aligned}
I_{op}^q(\theta_i, \theta_r, \Delta\phi_{ir}) = {} & \frac{1}{2\pi}\sqrt{\frac{(o+1)(p+1)}{(1+\delta_{op})(1+\delta_{q0})}} \\
& \times \left(R_o^q(\sqrt{2}\sin\frac{\theta_i}{2})R_p^q(\sqrt{2}\sin\frac{\theta_r}{2}) + R_p^q(\sqrt{2}\sin\frac{\theta_i}{2})R_o^q(\sqrt{2}\sin\frac{\theta_r}{2}) \right) \\
& \times \cos q\Delta\phi_{ir}
\end{aligned}
\tag{3.5}
$$

$R_o^q(\rho)$ are the Zernike polynomials. The Kronecker delta is used for a more compact notation of the normalization factor, $\delta_{ij} = 1$ if $i = j$ and 0 otherwise. The restrictions on the indices are $o \geq p \geq q \geq 0$ and $(o - q)$ and $(p - q)$ are even.

3.3   Image Irradiance Representation

Inserting the representations (3.3) and (3.4) into the irradiance expression (2.11) and moving everything not depending on (θ_i, ϕ_i) out of the integral results in

$$E(\alpha, \beta) = \sum_{\substack{lmn \\ opq}} L_l^m b_{op}^q D_l^{mn}(\alpha, \beta, 0)$$

$$\times \int_{\mathbf{H_{S2}}} Y_l^n(\theta_i, \phi_i) I_{op}^q(\theta_i, \alpha, \phi_i - \pi)\cos\theta_i d\omega_i. \tag{3.6}$$

Now let us study the integral in equation (3.6). Inserting the expressions for the spherical harmonic (see (A.6) in the appendix) and the BRDF (3.5) basis functions, we see that it can be separated into an azimuthal integral over ϕ_i and a polar integral over θ_i. The azimuthal integral is easily solved and eliminates all terms in the sum (3.6) except when $n = q$.

$$\int_0^{2\pi} \Phi_n(\phi_i) \cos q(\phi_i - \pi) d\phi_i = \delta_{nq} \pi (-1)^q \sqrt{2(1 + \delta_{q0})} \tag{3.7}$$

The polar integral can also be calculated analytically but the expression is rather complicated and for now we denote it

$$C_{lo}^q = \int_0^{\frac{\pi}{2}} P_l^q(\cos \theta_i) R_o^q(\sqrt{2} \sin \frac{\theta_i}{2}) \cos \theta_i \sin \theta_i d\theta_i. \tag{3.8}$$

The derivation and analytic expression for C_{lo}^q can be found in Appendix B.

Inserting these results in (3.6) we get an expression for the irradiance

$$E(\alpha, \beta) = \sum_{\substack{lm \\ opq}} L_l^m b_{op}^q (-1)^q N_l^q \sqrt{\frac{(o+1)(p+1)}{2(1+\delta_{op})}} D_l^{mq}(\alpha, \beta, 0)$$

$$\times \left(C_{lo}^q R_p^q(\sqrt{2} \sin \frac{\alpha}{2}) + C_{lp}^q R_o^q(\sqrt{2} \sin \frac{\alpha}{2}) \right). \tag{3.9}$$

Let

$$E_{lop}^{mq}(\alpha, \beta) = (-1)^q N_l^q \sqrt{\frac{(o+1)(p+1)}{2(1+\delta_{op})}} D_l^{mq}(\alpha, \beta, 0)$$

$$\times \left(C_{lo}^q R_p^q(\sqrt{2} \sin \frac{\alpha}{2}) + C_{lp}^q R_o^q(\sqrt{2} \sin \frac{\alpha}{2}) \right). \tag{3.10}$$

Then

$$E(\alpha, \beta) = \sum_{\substack{lm \\ opq}} L_l^m b_{op}^q E_{lop}^{mq}(\alpha, \beta). \tag{3.11}$$

The restrictions of the indices are $o \geq p \geq q \geq 0$, $(o - q)$ and $(p - q)$ are even, $l \geq q$ and $m = -l, \ldots, l$.

We have derived a basis for the image irradiance from isotropic materials. The image irradiance is represented as a sum of the basis functions, E_{lop}^{mq} multiplied by the coefficients of the light source, L_l^m and the coefficients of the BRDF of the material, b_{op}^q. This basis can represent the image irradiance from any isotropic surface under any illumination. For the general case an infinite number of basis functions is needed, but in many cases a finite number is sufficient. For instance, a perfect mirror under some illumination will require as many components as the illumination. However, most natural materials are not perfect mirrors. Their BRDF is band-limited and will in effect low-pass filter the illumination. In that case the series (3.11) can be truncated to a finite series and still be an accurate representation of the image irradiance.

3.3.1 Separability of the basis functions

The basis functions $E_{lop}^{mq}(\alpha, \beta)$ can be factored into a product of two functions. One depending only on the polar angle of the surface normal, α and the other depending only on the azimuth angle, β.

The expression for the rotation functions, $D_l^{mq}(\alpha, \beta, \gamma)$ are given in (A.17) in Appendix A. Because $q \geq 0$ and $\gamma = 0$ this expression can be simplified to

$$D_l^{mq}(\alpha, \beta, 0) = \frac{P_l^{mq}(\cos \alpha) + (-1)^q P_l^{m,-q}(\cos \alpha)}{\sqrt{2(1 + \delta_{q0})}} \Phi_m(\beta), \qquad (3.12)$$

which leads to a factorization of the basis functions,

$$E_{lop}^{mq}(\alpha, \beta) = \Theta_{lop}^{mq}(\alpha)\Phi_m(\beta), \qquad (3.13)$$

where

$$\begin{aligned}
\Theta_{lop}^{mq}(\alpha) = N_l^q \sqrt{\frac{(o+1)(p+1)}{4(1+\delta_{op})(1+\delta_{q0})}} \left((-1)^q P_l^{mq}(\cos\alpha) + P_l^{m,-q}(\cos\alpha)\right) \\
\times \left(C_{lo}^q R_p^q(\sqrt{2}\sin\frac{\alpha}{2}) + C_{lp}^q R_o^q(\sqrt{2}\sin\frac{\alpha}{2})\right).
\end{aligned} \qquad (3.14)$$

$\Phi_m(\beta)$ is defined in (A.7).

3.4 Approximating the Image Irradiance

A surface acts as a low-pass filter on the illumination. Thus the higher frequencies of illumination do not affect the appearance of the surface. This means in our case that the series (3.11) can be truncated at some illumination order l. If the BRDF can be expressed, or approximated, by a finite sum of BRDF basis functions then the image irradiance can also be expressed as a finite sum of the basis functions, E_{lop}^{mq}.

Basri and Jacobs (2003) and Ramamoorthi and Hanrahan (2001a) showed that the reflected light from a Lambertian surface can be approximated using only the first three orders of the illumination. Now follows an equivalent analysis for BRDFs represented in the basis (3.4) of Koenderink and van Doorn (1998). We consider the variance of the image irradiance, the variance in the image, which is what the camera registers and our eyes see. The variance, V, of the image irradiance is computed over all visible surface normals,

$$\begin{aligned}
V &= \int_{\mathbf{H_{S^2}}} \left(\sum_{\substack{lm \\ opq}} L_l^m b_{op}^q E_{lop}^{mq}(\alpha, \beta) \right)^2 d\omega \\
&= \int_{\mathbf{H_{S^2}}} \sum_{\substack{lm \\ opq}} \sum_{\substack{l'm' \\ o'p'q'}} L_l^m L_{l'}^{m'} b_{op}^q b_{o'p'}^{q'} E_{lop}^{mq}(\alpha, \beta) E_{l'o'p'}^{m'q'}(\alpha, \beta) d\omega.
\end{aligned} \qquad (3.15)$$

The variance depends both on the BRDF (the b_{op}^q's) and the illumination (the L_l^m's). To be independent of the position of the light source, we consider the average variance as the light source undergoes all 3D rotations, $SO(3)$. Consider a light source with coefficients $\tilde{L}_l^m$. The coefficients of that light source after a rotation $R \in SO(3)$ are $L_l^m = \sum_{n=-l}^l D_l^{nm}(R)\tilde{L}_l^n$. In (3.15) the only part depending on this rotation is $L_l^m L_{l'}^{m'}$ so we can compute the average $\int_{SO(3)} L_l^m L_{l'}^{m'} dR$ separately. Now, the D_l^{mn} are orthogonal over the rotation group,

$$\int_{SO(3)} D_l^{nm}(R)D_{l'}^{n'm'}(R)dR = \frac{\delta_{l,l'}\delta_{m,m'}\delta_{n,n'}}{2l+1},\tag{3.16}$$

so the average simply becomes

$$\int_{SO(3)} L_l^m L_{l'}^{m'} dR = \int_{SO(3)} \sum_{n,n'} \tilde{L}_l^n \tilde{L}_{l'}^{n'} D_l^{nm}(R)D_{l'}^{n'm'}(R)dR$$

$$= \delta_{ll'}\delta_{mm'}\frac{1}{2l+1}\sum_{n=-l}^l (\tilde{L}_l^n)^2.\tag{3.17}$$

Let

$$V_L(l) = \frac{1}{2l+1}\sum_{n=-l}^l (\tilde{L}_l^n)^2.\tag{3.18}$$

V_L is the average signal variance of each mode of the illumination. It is the same for all modes of the same order, hence only depends on the order l.

The expression for the average variance is obtained by inserting (3.17) into (3.15),

$$V_{ave} = \sum_l V_L(l) \sum_{\substack{opq \\ o'p'q'}} b_{op}^q b_{o'p'}^{q'} \sum_m \int_{\mathbf{H_{S2}}} E_{lop}^{mq}(\alpha,\beta)E_{lo'p'}^{mq'}(\alpha,\beta)d\omega.\tag{3.19}$$

Let

$$V_B(l) = \sum_{\substack{opq \\ o'p'q'}} b_{op}^q b_{o'p'}^{q'} \sum_m \int_{\mathbf{H_{S2}}} E_{lop}^{mq}(\alpha,\beta)E_{lo'p'}^{mq'}(\alpha,\beta)d\omega.\tag{3.20}$$

Then

$$V_{ave} = \sum_l V_L(l)V_B(l).\tag{3.21}$$

By studying the *average* variance the expression has been greatly simplified. Since the basis functions E_{lop}^{mq} are not orthogonal their covariances contribute to the variance, but in the average variance all covariances between different illumination orders have vanished. In this way the average variance can be written as a sum over l where each term is a product of V_L, which depends only on the illumination and

V_B, which depends only on the BRDF. The V_L's, again, are the average variances of the illumination modes. The V_B's can be seen as the attenuation factors of the BRDF on the illumination. They determine which frequencies of the illumination that are let through the BRDF.

We now return to the question: At which l can the series be truncated and still achieve a desired accuracy. Define the accuracy to be the variance of the truncated series.

$$\text{Accuracy} = \frac{\sum_{l=0}^{l_{max}} V_L(l)V_B(l)}{\sum_{l=0}^{\infty} V_L(l)V_B(l)} \tag{3.22}$$

3.4.1 Frequency Properties of Natural Illumination

To determine the accuracy we first have to take into account the frequency properties of the illumination. Consider first a point light source. The spherical harmonic coefficients of a light source are computed by integrating the light source with each harmonic basis function. Since a point light source can be described as a delta function on the sphere, its coefficients are the basis functions sampled at the position of the light source. If the light source is positioned at the north pole ($(0,0)$ in spherical coordinates) its coefficients are $\tilde{L}_l^m = Y_l^m(0,0)$, where

$$Y_l^m(0,0) = \begin{cases} \sqrt{\frac{2l+1}{4\pi}} & m = 0 \\ 0 & m \neq 0 \end{cases} \tag{3.23}$$

Consequently

$$V_{L,pointlight}(l) = \frac{1}{4\pi}. \tag{3.24}$$

To illustrate the properties of natural illumination we computed $V_L(l)$ for a number of captured real world illuminations. These illumination maps are captured by photographing a mirror sphere from several directions and at several exposures. The images are calibrated and merged to form a high dynamic range image of the full view sphere, (Debevec 1998). See Figure 3.3 for some examples. Figure 3.4 shows the $V_L(l)$ for a point light source and some of the illumination maps created by Debevec. While the point light source has a constant profile, the natural illumination rapidly decrease for higher l.

In the coming sections, when we compute the accuracy for a number of different BRDFs, we use the illumination properties of a point light source. From Figure 3.4 it can be seen that this gives a conservative estimate of the accuracy.

Figure 3.3:   Three of the nine illumination maps created by Debevec (1998). From left to right: an eucalyptus groove, St Peter's basilica and a breezeway. The captured illumination is an image on a sphere and mapped here onto a disc. The radius and angle of the circular coordinates of the disc correspond respectively to the polar and azimuthal angle of the spherical coordinates.
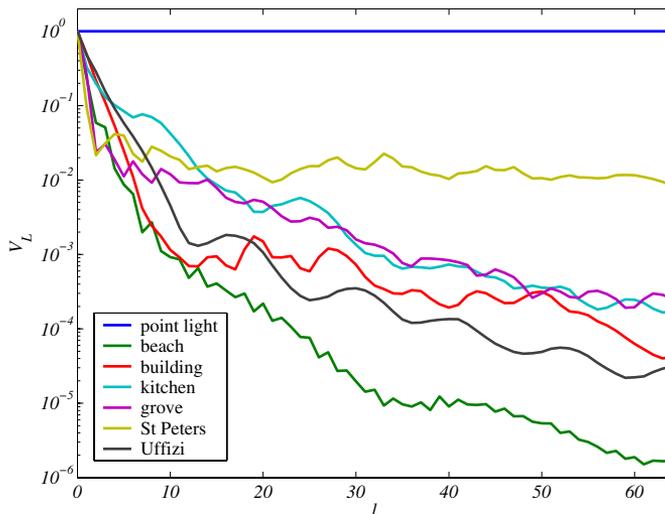
Figure 3.4:   Frequency properties of natural illumination. The plot shows the variance (log scale) of each order l of a point light source and a number of captured real world illumination maps (Debevec). While the variances over frequency stay constant for a point light, they rapidly decrease at high frequencies (high order l) for natural illumination.
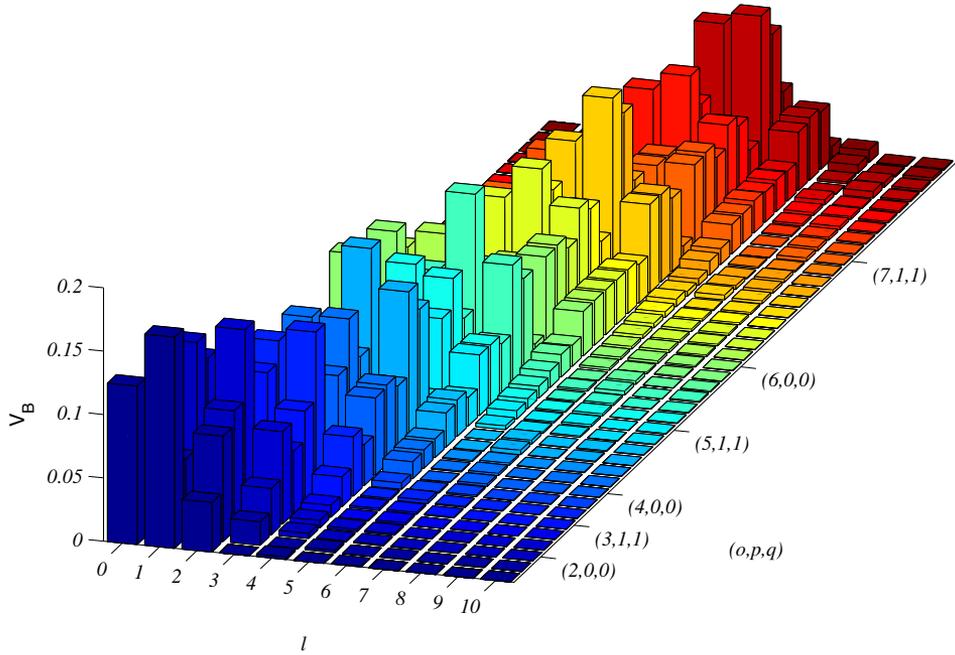
Figure 3.5: $V_B(l)$ for each reflectance mode, I_{op}^q.

3.4.2 Single Reflectance Mode

First, let's consider the reflected light from a single reflectance mode, I_{op}^q. For a single reflectance mode the analytical expression $V_B(l)$ becomes extremely simple.

$$
\begin{aligned}
V_{B,I_{op}^q}(l) &= \sum_m \int_{\mathbf{H_{S^2}}} (E_{lop}^{mq}(\alpha,\beta))^2 d\omega \\
&= \pi(N_l^q)^2 \left((o+1)(C_{lo}^q)^2 + (p+1)(C_{lp}^q)^2 \right)
\end{aligned}
\tag{3.25}
$$

The derivation can be found in Appendix C.

Figure 3.5 shows V_B for all the reflectance modes up to order $o = 7$. The coefficients have a maximum at around the same illumination order l as the material order o. For higher l it rapidly tends to zero.

We have not been able to prove this, but it seems as if $V_{B,I_{op}^q}(l)$ asymptotically tends to

$$
\lim_{l \to \infty} V_{B,I_{op}^q}(l) = \begin{cases} k_1 \, l^{-4} & l - q \text{ even} \\ k_2 \, l^{-6} & l - q \text{ odd} \end{cases}
\tag{3.26}
$$

If the series follows the asymptote at $l = l_{asym}$, we can estimate the variance of the "tail" of the series by

$$V_{tail} = \sum_{l=l_{asym}+1}^{\infty} V_L(l)V_B(l) \leq \sum_{l=l_{asym}+1}^{\infty} k_1 \, l^{-4} = k_1(\frac{\pi^4}{90} - \sum_{l=1}^{l_{asym}} k_1 \, l^{-4}) \quad (3.27)$$

since $\sum_{l=1}^{\infty} l^{-4} = \pi^4/90$. k_1 is estimated by $k_1 = V_L(l_{asym})V_B(l_{asym})l_{asym}^4$. In all computations below, the series are computed up to $l_{asym} = 200$. To ensure accuracy when computing the sums for such high orders we use infinite precision rational numbers.

Once we have an estimate of the tail variance we can compute the accuracy of the representation for a given l. Figure 3.6 shows the required l_{max} for each reflectance mode up to $o = 7$ to achieve three different levels of accuracies. Mode I_{00}^0 corresponds to a Lambertian BRDF. It's enough to use only the three first illumination order (up to $l = 2$) as was shown before by Basri and Jacobs (2003) and Ramamoorthi and Hanrahan (2001a). For the higher order reflectance modes more illumination modes are required. For instance, roughly an illumination order of $l = o + 2$ is required to achieve a 98% accuracy in most cases.

3.4.3   Composite BRDFs

Real BRDFs are linear combinations of the reflectance modes I_{op}^q. To compute the accuracy we compute $V_B(l)$ as in (3.20) which takes into account the covariances of the reflectance modes. It can also be calculated analytically. The expression can be found in Appendix C.

The asymptotic behaviour of $V_B(l)$ is similar to that of the single mode case. $V_B(l)$ is a linear combination of series that have asymptotes of $O(l^{-4})$ or $O(l^{-6})$. Therefore $V_B(l)$ has an asymptote of $O(l^{-n})$, where n lies somewhere between 4 and 6. Since it is still bounded by the $O(l^{-4})$ asymptote we compute the tail variance as in the single mode case.

As an illustration we have performed a variance analysis of the materials in the CUReT database (Dana et al. 1999). The CUReT database contains captured reflectance and textures of 61 materials. Examples are felt, leather, velvet, plaster, concrete and wood. The data have been captured by photographing the materials at different viewing and lighting angles. There are 205 samples of viewing and lighting direction combinations per material.

We have only used the reflectance part of the database. The samples of the BRDF were fitted to a Koenderink and van Doorn model of order $o_{max} = 7$. This was the highest order that could be used before overfitting occured. The $V_B(l)$ was computed and the tail variance was estimated and from that the accuracy was determined. Figure 3.7 shows $V_B(l)$ and the accuracy for four of the materials in the database. For a matte material, such as plaster, most of the variance depends only on illumination orders up to $l = 2$ and over 99% accuracy is achieved using only
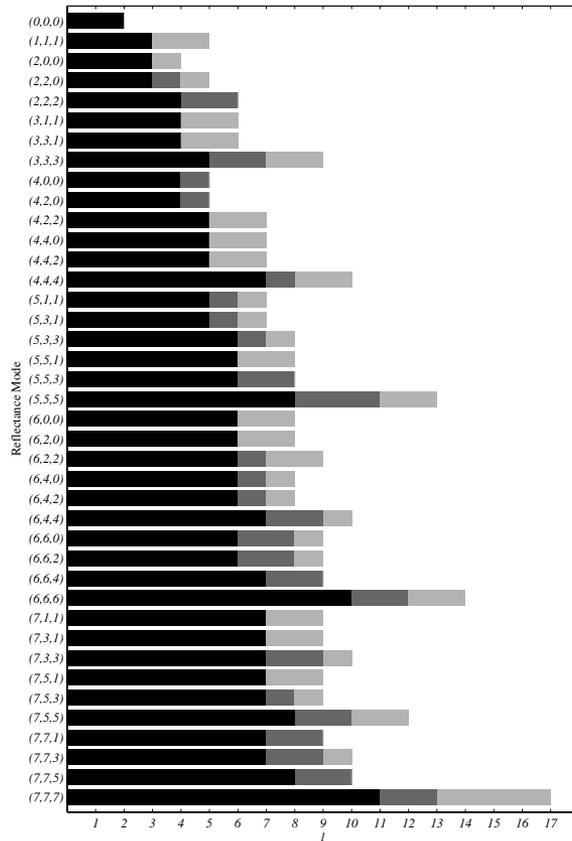
Figure 3.6: Required minimum illumination order for each reflectance mode. The black, dark gray and light gray bars show the required order to achieve 95%, 98% and 99% accuracy respectively.

those orders. Glossy and shiny materials, such as the leather and brick, depend on higher frequencies of the illumination which creates the specular reflection. This is shown as a "bump" in Figure 3.7a around $l = 7$. Velvet which has unusual reflectance properties (Lu, Koenderink, and Kappers 1998) passes through higher frequencies than plaster, but less so than leather and brick.

Graphs for all the materials in the CUReT database can be found in Figures C.1-C.6 in Appendix C. A summary of these results is shown in Figure 3.8 which shows the required minimum order, for all the materials, for three different accuracies. Notably quite a few materials can be accurately represented using only illumination orders up to $l = 2$. The specular materials are easily spotted as they require higher orders, up to $l = 10$.
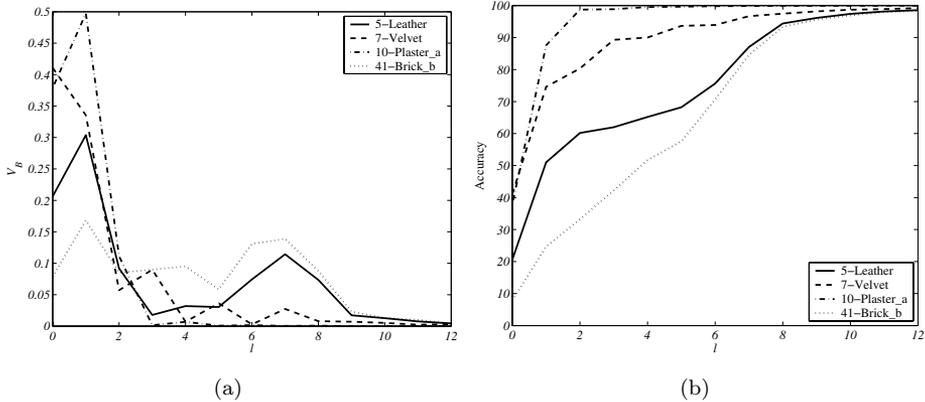
Figure 3.7:   Properties of some of the materials in the CUReT database. a) Attenuations factors V_B. b) Accuracy of image irradiance representation when using illumination orders up to l. Curves for all the materials in the CUReT database can be found in Figures C.1-C.6 in Appendix C.

3.5   Summary

BRDFs in the real world are in many cases smooth functions. This property makes them suitable to model by frequency space representations. Frequency space representations can account for BRDFs where other models have failed. Also, since frequency space models are general a single model can be used for a wide range of materials.

The smoothness of the BRDFs also accounts for why the BRDF acts as a low-pass filter on the illumination. This has the effect that only the lower frequencies of the illumination contribute to the appearance of the surface. Hence, only those low-frequencies should be taken into account for when analyzing shading. This suggests that also the illumination should be represented in the Fourier domain.

In this chapter, we have written the illumination in spherical harmonics and the BRDF in the basis of Koenderink and van Doorn, and derived a basis for the image irradiance. This basis consists of functions which are the contributions due to an illumination mode reflected by an reflectance mode.

In this basis the image irradiance from any isotropic BRDF under any illumination can be represented. An infinite number of basis functions is required in the worst case, but for smooth BRDFs which can be represented with a finite number of reflectance modes, the series can be truncated and still be an accurate representation.

By studying the variance of the image irradiance we show how to compute at which illumination order the series can be truncated. When analyzing the reflectance properties of the materials in the CUReT database, we see that matte
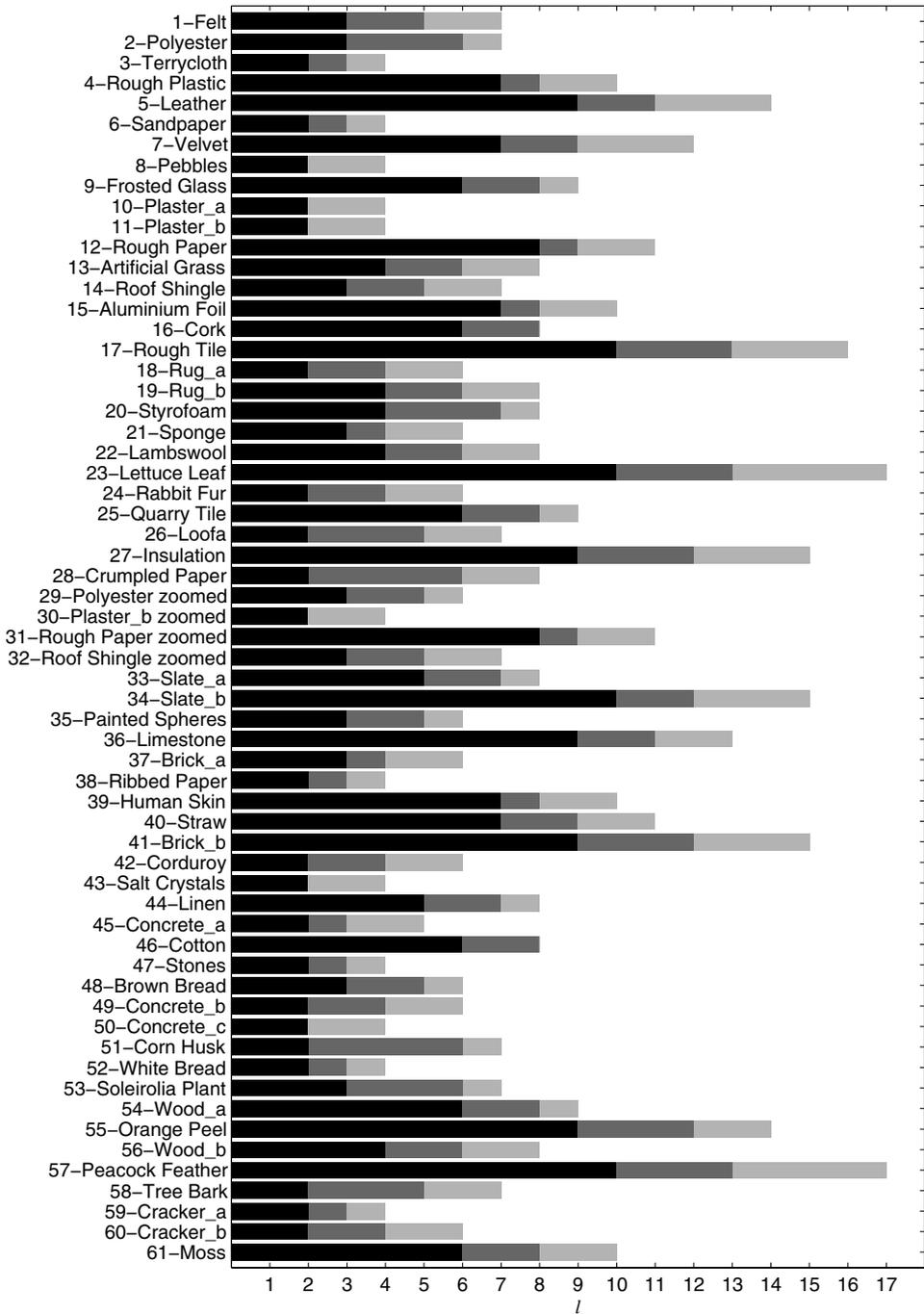
Figure 3.8: Required minimum illumination order for the materials in the CUReT database. The black, dark gray and light gray bars show the required order to achieve 95%, 98% and 99% accuracy respectively.

materials require illumination order up to $l = 2$, while the shinier materials require order up to around $l = 7$.

One important result is that we get a general finite dimensional representation for the image irradiance. In the following chapter we use principal component analysis to further reduce the dimensionality of the representation. The finite representation allows us to analytically derive the principal components, something which will prove very useful.

Chapter 4

Phenomenological Eigenfunctions for Image Irradiance

This thesis is primarily concerned with answering the question: What information can be extracted from the shading in a single image? For this purpose the model derived in the previous chapter, although finite dimensional, has far more parameters than can be estimated from a single image, in practice. Shading has been shown to lie in a much lower dimensional space.

Hallinan (1994) and Epstein, Hallinan, and Yuille (1995) have shown that images of faces and other objects lie in a low-dimensional subspace. They captured images of an object while varying the position of the light source. After performing principal component analysis on the images they found that the images could be recreated using only a few (5 ± 2) principal components.

Recently, Ramamoorthi (2002) analytically derived the principal components of a convex Lambertian object illuminated by a varying point light source. Since the principal components are derived analytically every possible position of the light source can be taken into account. This corresponds to an infinite number of images if one were to do the experiments empirically. His results confirm the empirical results, both qualitatively and quantitatively.

When estimating the parameters describing the shading the principal components are ideal. They are constructed to maximize the variance of the component they represent, which means they have a high signal-to-noise ratio (SNR). The problem is that in general it is hard to relate the principal components to the physical properties of the underlying model, in our case the illumination and the surface reflectance.

The solution is to do, what can be called, model-based PCA. In this chapter we construct the principal components by applying PCA directly on the image irradiance representation. Instead of capturing a large number of images, the variations in the images are described through the variations in the illumination and

the BRDF. As in (Ramamoorthi 2002) the rotation of the light source can be done analytically, which reduces the amount of data that needs to be acquired. But most importantly, by performing the PCA model-based we obtain an explicit relation between the principal components and the surface reflectance properties and the illumination. This enables the basis to be used as a physical model.

Unlike previous work, our derivation also takes into account variations in the surface reflectance properties. As a result of this analysis we obtain a framework for creating low-dimensional bases that can represent the images of a surface of a wide variety of reflectance properties and under practically all illumination conditions.

4.1 Principal Component Analysis

Principal component analysis amounts to finding a coordinate transformation such that the covariance matrix of the random variable (in our case the image) is diagonalized. A geometric interpretation of PCA is that the first principal component, U_0, is the direction which has the highest variability, i.e. U_0 maximizes the variance of the scalar product of itself and the images, or in mathematical terms:

$$U_0 = \underset{||U||=1}{\operatorname{argmax}} \operatorname{Var}\{I \bullet U\}, \tag{4.1}$$

where the image, I, is the random variable. The following principal components are constructed iteratively by maximizing the same variance with the added condition that they are orthogonal to all the previous ones, (Johnson and Wichern 1998).

4.1.1 Image-Centered PCA

The criteria in (4.1) maximizes the variance over the image distribution. For the purpose of estimating the components it is important that the variance *in the image* is maximized. Components with a high variance in the image have a high SNR. This changes how the images should be centered.

The objective is to find a basis that decomposes the image into a linear combination of coefficients and orthonormal basis functions, $I = \sum_i d_i U_i$. Typically images are analyzed by estimating the coefficients d_i. The higher the variance of the component $d_i U_i$ the higher SNR the component will have. Since the basis functions are orthonormal, d_i is given by $d_i = I \bullet U_i$ and the variance of the component is $d_i^2 = (I \bullet U_i)^2$.

Moreover, the basis should contain the constant function. The illumination frequently contains an ambient component which can vary. Including the constant function in the basis makes the remaining functions of the basis independent of the ambient component. The constant function is included in the basis by subtracting from each image its mean. This forces the basis functions to be orthogonal to the constant function. The constant function is then added to the basis at a later stage.
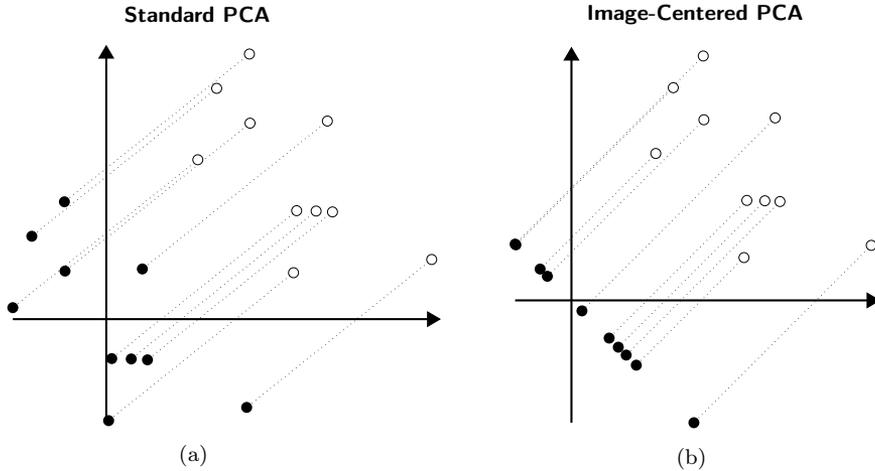
Figure 4.1: Variants of centering of the dataset. a) In standard PCA the dataset is centered by subtracting the mean image. b) In image-centered PCA each pixel in an image is subtracted by the mean of that particular image. This corresponds to projecting the datapoints, each representing an image, onto the hyperplane with normal $(1, 1, ..., 1)^T$.

Since the variance in the image depends on the illumination and reflectance properties, which vary, we maximize the variance *on average*. Taking this into account as well as subtracting the image mean, we arrive at the following criterion.

$$U_0 = \underset{||U||=1}{\operatorname{argmax}} \operatorname{E}\{((I - \mu_I \mathbf{1}) \bullet U)^2\} \tag{4.2}$$

where μ_I is the mean of image I.

To conclude, the standard PCA criterion (4.1) maximizes the variance over illumination and material changes while criterion (4.2) maximizes the variance in the image. In more detail, the difference between the two criteria is how the data is centered. Standard PCA subtracts the mean image from the dataset. Criterion (4.2) suggests that this should not be done. Instead, each image should be centered individually by subtracting the individual mean. To distinguish the latter method from standard PCA we call the method *Image-Centered PCA*. See Figure 4.1 for an illustration on how the centering differs in the two methods. Others have noticed that care should be taken when centering the dataset in PCA. One of the earlier references is (Noy-Meir, Walker, and Williams 1975).

4.2 The Image Set

The set of images to be analyzed are the images created from the variations in the illumination and surface reflectance. The surface shape and pose is fixed.

As in the previous chapter we approximate the projection as orthographic, which implies that the image irradiance is uniquely determined by the surface normal. Furthermore, we assume that the light field is the same for all the points on the surface. This assumptions is true if the light source is distant. It is also necessary that there are no cast shadows or local inter-reflections, which is true when the surface is convex. For non-convex surface this model is an approximation. The quality of the approximation depends on the concavities and the material-illumination conditions. For instance, bright objects will have stronger inter-reflections than dark objects.

Given the surface normal (α_i, β_i) at every pixel i, the spherical harmonic coefficients of the illumination and the coefficients of the BRDF, the image can be computed from the reflectance map represented in (3.10). This means that the variations of the images can be described through the variations in the illumination and the BRDF. Let the illumination distribution be described as a distribution of the spherical harmonic coefficients, $p_L(\mathbf{L})$, where $\mathbf{L}$ is the vector containing the coefficients L_l^m. Furthermore, let the variations in the surface reflectance properties be described as a distribution of the coefficients of the BRDF representation (3.4), $p_b(\mathbf{b})$, where $\mathbf{b}$ is a vector containing the elements b_{op}^q. These two distributions and the surface normals are enough to characterize the set of images of the surface.

4.2.1 Convexity

Because light is non-negative and light sources can be superposed, the images due to variations in the illuminations only, can be shown to lie in an convex cone in $\mathbb{R}^n$, where n is number pixels in the image, (Belhumeur and Kriegman 1998). Belhumeur and Kriegman call this set of images the *Illumination Cone*.

When there are variations in the BRDF as well, the set is not necessarily convex anymore. However, BRDFs can often be written as a linear combination of reflectance modes, e.g. diffuse reflection, specular reflection and backscatter reflection. If all linear combinations of BRDFs are valid (e.g. non-negative) the set of images is still convex.

Convexity is important. It ensures that PCA can be used to find a low-dimensional representation that will characterize the set well.

4.3   Deriving Model-Based PCA

In this section we derive the principal components of the image set by applying the PCA maximization criteria on the image irradiance expression (3.11). The standard PCA basis as well as the Image-Center basis is constructed.

For convenience and clarity we order the basis functions, E_{lop}^{mq} so that they can be index with a single variable. In single indexed notation the image irradiance in

equation (3.11) can be written as

$$E(\alpha, \beta) = \sum_s c_s E_s(\alpha, \beta), \tag{4.3}$$

where $c_s = L_l^m b_{op}^q$ and l, m, o, p and q are given by s from the ordering of the functions.

Begin by writing the eigenfunction as a sum of the basis functions of the image irradiance. This ensures that the symmetries of the BRDF are preserved in the eigenfunctions.

$$U(\alpha, \beta) = \sum_s u_s E_s(\alpha, \beta) \tag{4.4}$$

It is convenient to define the scalar product as the sum of the products of the images' pixels divided by the number of pixels, N. The scalar product of I and U is

$$I \bullet U = \frac{1}{N} \sum_{i=0}^{N} \sum_{s,s'} c_s u_{s'} E_s(\alpha_i, \beta_i) E_{s'}(\alpha_i, \beta_i) = \sum_{s,s'} c_s u_{s'} m_{ss'} \tag{4.5}$$

where $m_{ss'}$ are the scalar products of the irradiance mode images,

$$m_{ss'} = \frac{1}{N} \sum_i E_s(\alpha_i, \beta_i) E_{s'}(\alpha_i, \beta_i) \tag{4.6}$$

We can rewrite (4.5) in matrix form,

$$I \bullet U = \mathbf{u}^T \mathbf{M} \mathbf{c}, \tag{4.7}$$

where $\mathbf{u}$ is a vector containing the coefficients for U, $\mathbf{M}$ is a matrix that contains the elements $m_{ss'}$ and $\mathbf{c}$ is the vector containing the coefficients c_s.

Now, the only random variable in (4.7) is $\mathbf{c}$. Let $Covar\{\mathbf{c}\} = \boldsymbol{\Sigma_c}$. Then

$$\mathrm{Var}\{I \bullet U\} = \mathrm{Var}\{\mathbf{u}^T \mathbf{M} \mathbf{c}\} = \mathbf{u}^T \mathbf{M} \boldsymbol{\Sigma_c} \mathbf{M} \mathbf{u}. \tag{4.8}$$

The transpose on $\mathbf{M}$ has been dropped since $\mathbf{M}$ is symmetric. (4.8) should be maximized subject to the condition that U is normalized. We obtain the following constraint on the coefficients of U

$$U \bullet U = \frac{1}{N} \sum_i \sum_{s,s'} u_s u_{s'} E_s(\alpha_i, \beta_i) E_{s'}(\alpha_i, \beta_i) = \sum_{s,s'} u_s u_{s'} m_{ss'} = \mathbf{u}^T \mathbf{M} \mathbf{u} = 1. \tag{4.9}$$

The maximization problem can now be written in terms of U's coefficients.

$$\mathbf{u_0} = \operatorname*{argmax}_{\mathbf{u}^T \mathbf{M} \mathbf{u} = 1} \mathbf{u}^T \mathbf{M} \boldsymbol{\Sigma_c} \mathbf{M} \mathbf{u} \tag{4.10}$$

Applying the coordinate transform

$$\mathbf{v} = \mathbf{M}^{1/2}\mathbf{u}, \tag{4.11}$$

we obtain

$$\mathbf{v_0} = \underset{\mathbf{v}^T\mathbf{v}=1}{\operatorname{argmax}} \mathbf{v}^T\mathbf{M}^{1/2}\mathbf{\Sigma_c}\mathbf{M}^{1/2}\mathbf{v}. \tag{4.12}$$

This is a quadratic expression of $\mathbf{v}$ which should be maximized under the condition that $\mathbf{v}$ is normalized. It is well known that the maximum occurs when $\mathbf{v}$ is eigenvector of $\mathbf{M}^{1/2}\mathbf{\Sigma_c}\mathbf{M}^{1/2}$ with the largest eigenvalue. The subsequent eigenvectors maximize the expression while being orthogonal to the previous ones. We have in fact performed the PCA. To be sure we can convince ourselves that orthogonality of the vectors $\mathbf{v}$ corresponds to orthogonality in the image space.

$$U_i \bullet U_j = \mathbf{u_i}^T\mathbf{M}\mathbf{u_j} = \mathbf{v_i}^T\mathbf{v_j} \tag{4.13}$$

The coefficients for the eigenimages U_i are computed as

$$\mathbf{u_i} = \mathbf{M}^{-1/2}\mathbf{v_i}. \tag{4.14}$$

4.3.1   Image-Centered PCA

When constructing the basis according to Image-Centered PCA (4.2) the image mean should be subtracted from the image. The image mean is

$$\mu_I = \frac{1}{N} \sum_i \sum_s c_s E_s(\alpha_i, \beta_i) = \sum_s c_s e_s, \tag{4.15}$$

where e_s is the mean of E_s over the surface normal distribution.

$$e_s = \frac{1}{N} \sum_i E_s(\alpha_i, \beta_i) \tag{4.16}$$

Now,

$$\begin{aligned}
(I - \mu_I \mathbf{1}) \bullet U &= \frac{1}{N} \sum_i \sum_{s,s'} c_s u_{s'}(E_s(\alpha_i, \beta_i) - e_s)E_{s'}(\alpha_i, \beta_i) \\
&= \sum_{s,s'} c_s u_{s'}(m_{ss'} - e_s e_{s'}) \\
&= \mathbf{u}^T(\mathbf{M} - \mathbf{e}\mathbf{e}^T)\mathbf{c}
\end{aligned} \tag{4.17}$$

Furthermore

$$\mathrm{E}\{((I - \mu_I \mathbf{1}) \bullet U)^2\} = \mathbf{u}^T(\mathbf{M} - \mathbf{e}\mathbf{e}^T)\mathbf{V_c}(\mathbf{M} - \mathbf{e}\mathbf{e}^T)\mathbf{u} \tag{4.18}$$

where $\mathbf{V_c} = \mathrm{E}\{\mathbf{cc}^T\}$ which corresponds to the covariance matrix $\boldsymbol{\Sigma_c}$ calculated without subtracting the mean (or the second moment matrix).

Applying the same coordinate change as previously we obtain the $\mathbf{v}$ vectors as the eigenvectors of the matrix

$$\mathbf{M}^{-1/2}(\mathbf{M} - \mathbf{ee}^T)\mathbf{V_c}(\mathbf{M} - \mathbf{ee}^T)\mathbf{M}^{-1/2} \tag{4.19}$$

and finally the coefficients of the eigenimages by $\mathbf{u_i} = \mathbf{M}^{-1/2}\mathbf{v_i}$ as before.

4.3.2   Continuous Images

If the images are continuous the PCA bases can be computed in the same way. The only thing that changes are the elements of $\mathbf{M}$ and $\mathbf{e}$. For continuous images

$$m_{ss'} = \int_{\mathbf{H^2}} p_{\hat{n}}(\alpha, \beta) E_s(\alpha, \beta) E_{s'}(\alpha, \beta) d\omega \tag{4.20}$$

and

$$e_s = \int_{\mathbf{H^2}} p_{\hat{n}}(\alpha, \beta) E_s(\alpha, \beta) d\omega \tag{4.21}$$

where $p_{\hat{n}}(\alpha, \beta)$ is the surface normal distribution in the image. As an example, in the image of a sphere the surface normal distribution is proportional to the foreshortening factor, $p_{\hat{n}}(\alpha, \beta) = 2\cos\alpha$.

4.3.3   Calculating the Covariances

It is worth making some comments about calculating the covariance matrices $\boldsymbol{\Sigma_c}$ and $\mathbf{V_c}$. The two matrices are related by

$$\boldsymbol{\Sigma_c} = \mathbf{V_c} - \mu_{\mathbf{c}}\mu_{\mathbf{c}}^T \tag{4.22}$$

where $\mu_{\mathbf{c}}$ is the mean vector of $\mathbf{c}$. Since we can assume that the illumination and the surface reflectance are independent[1] the calculation of the elements of $\mu_{\mathbf{c}}$ and $\mathbf{V_c}$ can be partitioned into

$$\mu_{c_s} = \mathrm{E}\{L_l^m\}\mathrm{E}\{b_{op}^q\} \tag{4.23}$$

$$\mathrm{E}\{c_s c_{s'}\} = \mathrm{E}\{L_l^m L_{l'}^{m'}\}\mathrm{E}\{b_{op}^q b_{o'p'}^{q'}\} \tag{4.24}$$

The indices l, m, o, p and q are given by s from the ordering of the basis functions.

---

[1] Though one could argue that e.g. the illumination from an overcast sky correlates with shiny materials in the form of raincoats.

4.3.3.1   Rotated Light Sources

The position of a light source has a big impact on the appearance of a surface. To take into account all possible positions of the light source one can rotate the light source through all 3D rotations. By doing the rotation of the light source analytically it is possible to take into account all possible light source configurations. This corresponds to taking an infinite number of images if one were to do the experiments empirically.

Consider a light source, defined by its spherical harmonic coefficients $\tilde{L}_l^m$. If the variation in the illumination is characterized by the rotation of this light source, the second moments of the illumination are the product of two rotated coefficients integrated over the rotation group, $SO(3)$. The coefficients of the light source after a 3D rotation $R \in SO(3)$ is

$$L_l^m = \sum_{n=-l}^{l} D_l^{nm}(R)\tilde{L}_l^n,$$ (4.25)

Since the D_l^{mn}s are orthogonal,

$$\int_{SO(3)} D_l^{nm}(R) D_{l'}^{n'm'}(R) dR = \frac{\delta_{l,l'}\delta_{m,m'}\delta_{n,n'}}{2l+1}$$ (4.26)

the second moments $\mathrm{E}\{L_l^m L_{l'}^{m'}\}$ will be non-zero only when $l = l'$ and $m = m'$.

$$\mathrm{E}\{L_l^m L_{l'}^{m'}\} = \begin{cases} \frac{1}{2l+1}\sum_{n=-l}^{l}(\tilde{L}_l^n)^2 & l = l' \text{ and } m = m' \\ 0 & \text{otherwise} \end{cases}$$ (4.27)

As $D_0^{00}(R) = 1$ and due to the orthogonality relation, the means are zero for all L_l^m except L_0^0.

$$\mathrm{E}\{L_l^m\} = \begin{cases} \tilde{L}_0^0 & l = 0 \text{ and } m = 0 \\ 0 & \text{otherwise} \end{cases}$$ (4.28)

4.4   Relating the Principal Components to the Image Formation Model

An important feature of model-based PCA is that we can relate the principal component coefficients to the parameters of the illumination and the surface reflectance.

Say that we have found the principal component decomposition of image I,

$$I = \sum_j d_j U_j.$$ (4.29)

The eigenimages U_j are represented in the image irradiance basis functions E_i.

$$U_j = \sum_i u_{ij} E_i$$ (4.30)

Inserting (4.30) in (4.29) results in

$$I = \sum_j d_j \sum_i u_{ij} E_i.$$

(4.31)

The image can also be written in the image irradiance representation as

$$I = \sum_i c_i E_i.$$

(4.32)

where $c_i = L_l^m b_{op}^q$. Now, we can identify the coefficients c_i as

$$c_i = \sum_j u_{ij} d_j$$

(4.33)

In matrix form, all the illumination-BRDF coefficients are given by

$$\mathbf{c} = \mathbf{Ud}$$

(4.34)

where $\mathbf{c}$ is a vector containing the elements, c_i, $\mathbf{U}$ is a matrix containing the coefficients of the principal components and $\mathbf{d}$ is a vector containing the coefficients representing the image in the PCA basis.

(4.34) is not a one-to-one mapping. The matrix U is rank deficient, reflecting the ambiguities that exists when analyzing shading.

For e.g. rendering it is useful to do the reverse transformation

$$\mathbf{d} = \mathbf{U}^T \mathbf{Mc}$$

(4.35)

as $\mathbf{U}^{-1} = \mathbf{U}^T \mathbf{M}$.

4.5 Conclusions

We have derived the principal components for the set of images from a surface of a fixed shape under varying illumination and with varying BRDF. The PCA is performed in a model-based way so that it is not necessary to acquire or render images. Instead the variations are described by the variations in the illumination and BRDF.

Using the finite dimensional frequency space representation, the solution of the PCA is transformed from being the eigensystem of the image covariance matrix into an eigenvalue problem of the covariances of the frequency space coefficients. This separates the covariance matrix into a product of a matrix depending only on the surface shape and another matrix depending only on the variations in the illumination and the BRDF. Furthermore, the illumination-BRDF covariance matrix is separable on an element level. In effect, using this framework it is possible to rapidly compute bases for a wide variety of conditions and different surface shapes

by combining different surface, illumination and BRDF covariances. Since the rotation of the light source can be done analytically, the need for captured data is reduced dramatically.

Another advantage of model-based PCA is that there is an explicit relation between the principal components and the coefficients of the illumination and the BRDF. A PCA basis created from empirical data can capture the appearance of the images well, but it is very hard to relate the basis to the underlying properties of the data. With model-based PCA we are able to create a basis that takes into account the appearance ambiguities and at the same time can be used as a physics model.

Care has also been taken in forming the principal components. For image analysis it is important that the components to be estimated have a high signal-to-noise ratio. The basis should also be invariant to variations in the ambient component of the illumination. These factors suggest that the image should be centered by subtracting the individual image mean rather the mean image.

All these things together create a powerful framework for creating low-dimensional generative models of images of a surface under a wide variety of conditions. In the subsequent chapters we compute bases for a number of different illumination distributions and surface reflectances. The two variants of PCA will also be compared. In Chapter 7 the framework will be used for material classification.

Chapter 5

Bases for Lambertian Surfaces

In this chapter we compute PCA bases for Lambertian surfaces. Bases for a number of different illumination conditions are computed using both standard PCA and Image-Centered PCA. These are tested and compared on a set of rendered images.

5.1 Computing the Bases

The bases are computed as described in Section 4.3. First, the image irradiance representation in (3.11) is truncated to a finite sum, a necessary task for the computation of the PCA basis. The Lambertian BRDF is represented in the isotropic basis by a single basis function. All coefficients b_{op}^q are zero except b_{00}^0. As already stated only nine basis functions are needed to represent the image irradiance from a Lambertian surface. They are E_{l00}^{m0}, $l = 0, 1, 2$ and $m = -l, ...l$. If these functions are simplified they turn out to be proportional to the nine first spherical harmonic basis functions (since $D_l^{m0}(\alpha, \beta, 0) \propto Y_l^m(\alpha, \beta)$). It is the same representation as in (Basri and Jacobs 2003; Ramamoorthi and Hanrahan 2001a).

These nine functions form the basis for the image irradiance. We order and rename them as E_s, $s = 0, ..., 8$. Later, in the tables listing the bases, the eigenfunctions are written in the spherical harmonic basis functions to enable comparison with the results of Ramamoorthi (2002).

To compute the basis we need to compute the matrices $\mathbf{M}$ and $\mathbf{\Sigma_c}$. For Image-Centered PCA we also need to compute the $\mathbf{e}$ and $\mathbf{V_c}$. $\mathbf{M}$ and $\mathbf{e}$ are determined by the shape of the surface. We have chosen the shape of a sphere since it has all the possible visible surface normals represented in the image. Also we analyze continuous images to consider all surface normals.

The matrix $\mathbf{M}$ contains the covariance of the basis functions over the surface normals. For continuous images of a sphere the percentage of points with surface normal (α, β) is proportional to the foreshortening factor $\cos \alpha$, so the elements of

M are

$$m_{ss'} = \frac{1}{\pi} \int_0^{2\pi} \int_0^{\pi/2} E_s(\alpha, \beta) E_{s'}(\alpha, \beta) \cos \alpha \sin \alpha d\alpha d\beta \tag{5.1}$$

and the elements of **e** are

$$e_s = \frac{1}{\pi} \int_0^{2\pi} \int_0^{\pi/2} E_s(\alpha, \beta) \cos \alpha \sin \alpha d\alpha d\beta \tag{5.2}$$

$\mathbf{\Sigma_c}$ and $\mathbf{V_c}$ are related, see equation (4.22), so we only need to compute $\mathbf{V_c}$ and $\mu_\mathbf{c}$, the second moment matrix and the mean vector of the coefficients **c**. There are no variations in the surface reflectance as b_{00}^0 is constant. This implies that $\mathbf{V_c}$ and $\mu_\mathbf{c}$ are determined by means and covariances of the illumination. In the next section we compute the means and covariances for a number of illumination configurations of interest.

5.2   Illumination Distributions

The illumination distributions we consider always include all possible 3D rotations of the light source. This means we can use the formulas derived in Section 4.3.3.1 for rotated light sources. In such cases the covariances are all zero and only the variances $E\{(L_l^m)^2\}$ need to be calculated. Moreover, the means are all zero except $E\{L_0^0\}$.

5.2.1   Point Light Source

Now, consider a single point light source rotated to illuminate the surface from all possible directions. To compute the covariance matrix for this illumination distribution we simply find the spherical harmonic coefficients of one position of the light source and use formulas (4.27) and (4.28) to compute the necessary values for the light source when it undergoes all 3D rotations.

To find the spherical harmonic coefficients for a light source we integrate the product of the light source and each spherical harmonic function. A point light source is a delta function on the sphere which means that its coefficients are samples of the spherical harmonic basis functions at the position of the light source. With the point light at $(0,0)$, $\tilde{L}_l^m = Y_l^m(0,0)$ and

$$Y_l^m(0,0) = \begin{cases} \sqrt{\frac{2l+1}{4\pi}} & m = 0 \\ 0 & m \neq 0 \end{cases} \tag{5.3}$$

Hence

$$E\{(L_l^m)^2\} = \frac{2l+1}{4\pi(2l+1)} = \frac{1}{4\pi} \tag{5.4}$$

$$E\{L_0^0\} = \frac{1}{2\sqrt{\pi}} \tag{5.5}$$
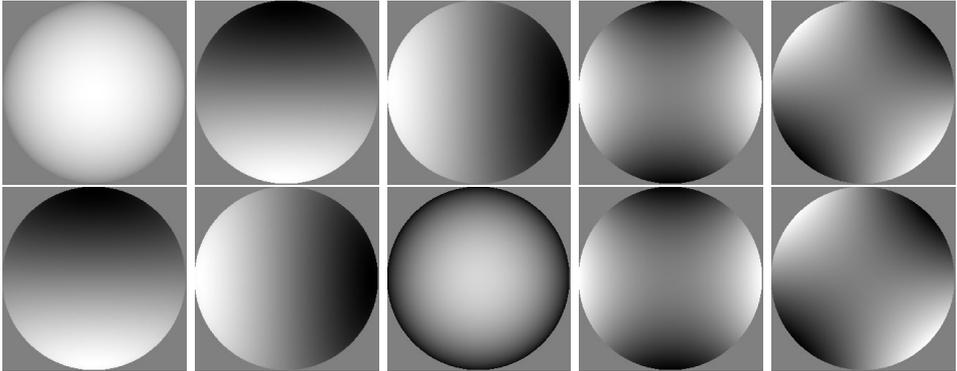
Figure 5.1: The first five eigenimages of a Lambertian sphere under a point light source. The first row shows images created with Standard PCA and the second row eigenimages from Image-Centered PCA. The Image-Centered basis is orthogonal to the constant function and needs the constant function to be complete. The Standard PCA basis is complete, but before an image is fitted to the basis, the mean image needs to be subtracted from the image.

From the above computations we can compute all necessary matrices needed to compute the basis. Table 5.1 lists the eigenfunctions with their eigenvalues and the cumulative sum of eigenvalues, corresponding to the percentage of the variance accounted for (VAF). Figure 5.1 shows the first five eigenimages generated from these vectors.

The differences between the bases are discussed later.

5.2.2   Point Light and Ambient Source

A more interesting case is the when there is a point light source and an ambient component. As before we position the point light at $(0, 0)$ and use equation (4.27) to take all rotations into account. The ambient component only contributes to $\tilde{L}_0^0$. The other $\tilde{L}_l^m$'s remain the same. This means that the top left element of the covariance matrix differs from the previous section. We also allow the level of the ambiance, a to vary from zero to a_{max}.

$$E\{(L_0^0)^2\} = \frac{1}{a_{max}} \int_0^{a_{max}} (\frac{1}{2\sqrt{\pi}} + a)^2 da \tag{5.6}$$

and

$$E\{(L_l^m)^2\} = \begin{cases} \frac{1}{4\pi} + \frac{a_{max}}{2\sqrt{\pi}} + \frac{a_{max}^2}{3} & l = m = 0 \\ \frac{1}{4\pi} & otherwise \end{cases} \tag{5.7}$$

$$E\{L_0^0\} = \frac{1}{2\sqrt{\pi}} + \frac{a_{max}}{2} \tag{5.8}$$

The eigenfunctions, for $a_{max} = 1$, are shown in Table 5.2.

5.2.3 Two Point Light Sources

We can construct the distribution of all configurations of two point light sources. Position the first light source at $(0,0)$ and the second at $(\psi, 0)$. Let ψ vary between 0 and π and let equation (4.27) take care of the rotations. The coefficients before rotation are $\tilde{L}_l^m = Y_l^m(0,0) + Y_l^m(\psi, 0)$ and

$$E\{(L_l^m)^2\} \quad = \quad \sum_{n=-l}^{l} \frac{\int_0^\pi (Y_l^n(0,0) + Y_l^n(\psi,0))^2 d\psi}{\pi(2l+1)} \tag{5.9}$$

$$E\{L_0^0\} \quad = \quad \frac{1}{\pi} \int_0^\pi (Y_0^0(0,0) + Y_0^0(\psi,0)) d\psi \tag{5.10}$$

Table 5.3 lists the eigenfunctions.

5.2.4 Set of Captured Illumination Maps

Of greater interest is to perform PCA over more realistic illuminations. In this section we derive the variances for a set of captured illumination maps, each which is rotated with all possible rotations. To compute the actual basis we use the nine illumination maps of Debevec (1998).

Consider n_L light sources. Each light source $k = 1, \ldots, n_L$ has spherical harmonic coefficients $\tilde{L}_{l,k}^m$ and a probability of $P(k)$. Again, starting at (4.27) and summing over all light sources results in

$$E\{(L_l^m)^2\} \quad = \quad \sum_{k=1}^{n_L} \sum_{n=-l}^{l} \frac{(\tilde{L}_{l,k}^m)^2 P(k)}{(2l+1)} \tag{5.11}$$

$$E\{L_0^0\} \quad = \quad \sum_{k=1}^{n_L} \tilde{L}_{l,k}^m P(k) \tag{5.12}$$

With each illumination map being equally probable $P(k) = \frac{1}{n_L}$. The eigenfunctions are in Table 5.4.

5.3 Discussion

The main difference between the PCA variants is how they behave under variations in the ambient component of the illumination. When the ambient component varies, such as in Table 5.2 and 5.4, this gets encoded in the standard basis. The Image-Centered basis on the other hand is not affected at all by this. For instance, the Image-Centered bases created from a point light and a point plus ambient light, are identical.

The Image-Centered bases are not affected by variations in the ambient component. The changes in the Image-Centered bases for different illumination distributions are small in general. The basis functions remain in the same order. Some

Table 5.1: Eigenfunctions of Sphere under Point Light Source

Standard PCA			Image-Centered PCA		
eigenfunction	λ	VAF	eigenfunction	λ	VAF
$0.99Y_1^0 + 0.1Y_2^0$	0.43	0.43	$0.99Y_1^{-1} + 0.17Y_2^{-1}$	0.39	0.39
$0.99Y_1^{-1} + 0.17Y_2^{-1}$	0.24	0.68	$0.99Y_1^1 + 0.17Y_2^1$	0.39	0.78
$0.99Y_1^1 + 0.17Y_2^1$	0.24	0.92	$-0.79Y_0^0 + 0.59Y_1^0 + 0.2Y_2^0$	0.13	0.9
Y_2^2	0.024	0.95	Y_2^2	0.038	0.94
Y_2^{-2}	0.024	0.97	Y_2^{-2}	0.038	0.98
$-0.59Y_1^0 + 0.81Y_2^0$	0.019	0.99	$-0.78Y_1^1 + 0.63Y_2^1$	0.009	0.99
$-0.78Y_1^1 + 0.63Y_2^1$	0.005	0.99	$0.78Y_1^{-1} - 0.63Y_2^{-1}$	0.009	1
$-0.78Y_1^{-1} + 0.63Y_2^{-1}$	0.005	1	$0.65Y_0^0 - 0.7Y_1^0 + 0.3Y_2^0$	0.001	1
$0.65Y_0^0 - 0.7Y_1^0 + 0.29Y_2^0$	0	1	Y_0^0	0	1

Table 5.2: Eigenfunctions of Sphere under Point and Ambient Light Source

Standard PCA			Image-Centered PCA		
eigenfunction	λ	VAF	eigenfunction	λ	VAF
$0.89Y_0^0 + 0.46Y_1^0 + 0.037Y_2^0$	0.63	0.63	$0.99Y_1^{-1} + 0.17Y_2^{-1}$	0.39	0.39
$0.99Y_1^{-1} + 0.17Y_2^{-1}$	0.15	0.78	$0.99Y_1^1 + 0.17Y_2^1$	0.39	0.78
$0.99Y_1^1 + 0.17Y_2^1$	0.15	0.93	$-0.79Y_0^0 + 0.59Y_1^0 + 0.2Y_2^0$	0.13	0.9
$-0.82Y_0^0 + 0.51Y_1^0 + 0.28Y_2^0$	0.035	0.96	Y_2^2	0.038	0.94
Y_2^2	0.015	0.98	Y_2^{-2}	0.038	0.98
Y_2^{-2}	0.015	0.99	$0.78Y_1^1 - 0.63Y_2^1$	0.009	0.99
$-0.78Y_1^1 + 0.63Y_2^1$	0.003	1	$0.78Y_1^{-1} - 0.63Y_2^{-1}$	0.009	1
$0.78Y_1^{-1} - 0.63Y_2^{-1}$	0.003	1	$0.65Y_0^0 - 0.7Y_1^0 + 0.3Y_2^0$	0.001	1
$0.65Y_0^0 - 0.7Y_1^0 + 0.3Y_2^0$	0.000	1	Y_0^0	0	1

Table 5.3: Eigenfunctions of Sphere under Two Point Light Sources

Standard PCA			Image-Centered PCA		
eigenfunction	λ	VAF	eigenfunction	λ	VAF
$0.99Y_1^0 + 0.13Y_2^0$	0.42	0.42	$0.98Y_1^{-1} + 0.21Y_2^{-1}$	0.38	0.38
$0.98Y_1^{-1} + 0.21Y_2^{-1}$	0.24	0.66	$0.98Y_1^1 + 0.21Y_2^1$	0.38	0.76
$0.98Y_1^1 + 0.21Y_2^1$	0.24	0.91	$-0.79Y_0^0 + 0.57Y_1^0 + 0.24Y_2^0$	0.13	0.89
Y_2^2	0.028	0.94	Y_2^2	0.044	0.93
Y_2^{-2}	0.028	0.96	Y_2^{-2}	0.044	0.98
$-0.59Y_1^0 + 0.81Y_2^0$	0.022	0.99	$0.78Y_1^1 - 0.63Y_2^1$	0.01	0.99
$-0.78Y_1^1 + 0.63Y_2^1$	0.006	0.99	$-0.78Y_1^{-1} + 0.63Y_2^{-1}$	0.01	1
$-0.78Y_1^{-1} + 0.63Y_2^{-1}$	0.006	1	$0.65Y_0^0 - 0.7Y_1^0 + 0.3Y_2^0$	0.001	1
$0.65Y_0^0 - 0.7Y_1^0 + 0.29Y_2^0$	0	1	Y_0^0	0	1

Table 5.4: Eigenfunctions of Sphere under Set of Illumination Maps

Standard PCA			Image-Centered PCA		
eigenfunction	λ	VAF	eigenfunction	λ	VAF
$0.86Y_0^0 + 0.51Y_1^0 + 0.024Y_2^0$	0.64	0.64	$1Y_1^1 + 0.097Y_2^1$	0.41	0.41
$1Y_1^{-1} + 0.097Y_2^{-1}$	0.15	0.79	$1Y_1^{-1} + 0.097Y_2^{-1}$	0.41	0.82
$1Y_1^1 + 0.097Y_2^1$	0.15	0.95	$-0.78Y_0^0 + 0.62Y_1^0 + 0.12Y_2^0$	0.12	0.94
$-0.81Y_0^0 + 0.56Y_1^0 + 0.17Y_2^0$	0.029	0.98	Y_2^2	0.024	0.96
Y_2^2	0.009	0.99	Y_2^{-2}	0.024	0.99
Y_2^{-2}	0.009	0.99	$-0.77Y_1^1 + 0.64Y_2^1$	0.006	0.99
$0.77Y_1^{-1} - 0.64Y_2^{-1}$	0.002	1	$0.77Y_1^{-1} - 0.64Y_2^{-1}$	0.006	1
$-0.77Y_1^1 + 0.64Y_2^1$	0.002	1	$0.65Y_0^0 - 0.7Y_1^0 + 0.3Y_2^0$	0.001	1
$0.65Y_0^0 - 0.7Y_1^0 + 0.3Y_2^0$	0.000	1	Y_0^0	0	1

Figure 5.2: Spheres rendered with groove, beach and St Peter's illumination maps of Debevec. The background is actually the visible part of the illumination map.

basis functions have minor changes. It seems that when changes in the ambient component is accounted for, the illumination distribution has little effect on the basis. This indicates that the Image-Centered basis could be an adequate basis under a wide range of illumination distributions.

5.4   Testing the Bases

In order to compare the different bases, we rendered images of spheres using the illumination maps of Debevec. Naturally, it is more desirable to use real images. However, by using synthetic images we rule out many potential sources of errors. More importantly it gives us access to the geometry of the scene, needed for the experiments. The rendering was implemented in a manner similar to (Ramamoorthi and Hanrahan 2002) although we use the Koenderink, van Doorn basis for the BRDF. The rendering involves summing the contributions from each illumination and material mode. By calculating the variance of each component we get an estimate of the error and can render to a very high degree of accuracy. By using real world illumination maps we get images with the full complexity of natural lighting. Four views from each of the nine illumination maps were rendered resulting in 36 images. Figure 5.2 shows three of the rendered images.

Each basis was tested by calculating how well it represents the rendered images. The image was fitted to the basis and the variance of the residual was used as a measure. For each basis the residual was calculated using from one to all of the basis functions. Figure 5.3a shows the residual errors for bases of the two PCA variants constructed from the illumination map distribution. The Standard PCA basis is slightly worse for low numbers of basis function but quickly catches up with the Image-Centered basis.

When comparing bases constructed from different illumination distributions the differences are smaller. Figure 5.3b shows the relative residual errors. The basis constructed with a point light plus ambient source was not included since it is identical to the basis constructed with only a point light. The illumination map
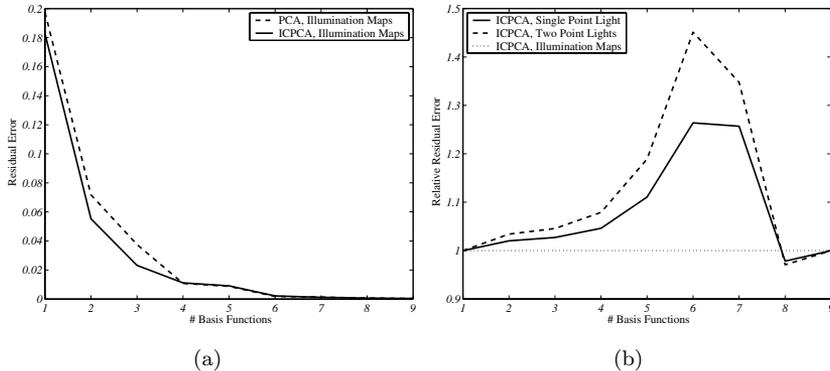
Figure 5.3: Residual error of images reconstructed with PCA bases with an increasing number of basis functions. a) PCA bases constructed from the illumination map distribution. b) Relative error of Image-Centered bases from different illumination distributions.

basis was constructed in a cross-validation sense so that the map used to render the image never was in the "training" set. Still it shows the best representability for the most part.

5.5 Light Field Estimation from a Single Image and Known Geometry

As a demonstration of the potential applications of the PCA bases, the light field is estimated from a single (rendered) image.

Light field estimation from several view points and known geometry has already been demonstrated using the spherical harmonic representation, (Ramamoorthi and Hanrahan 2001a). When estimating the light field from a single image we need to take into account that the image lies in a subspace of lower dimensionality than the 9D space required by the spherical harmonic representation. The problem should be regularized. The natural way to do this is to estimate only the components with the highest variance in the image. Components with a low variance will "drown" in the noise, making the estimates error prone. The principal components are created to have as high variance as possible and are therefore ideal in this situation.

The procedure is as follows. Using least squares the principal component coefficients are estimated so that the image is written as a linear combination of the principal components.

$$I = \sum_{i=0}^{N} d_i U_i \qquad (5.13)$$
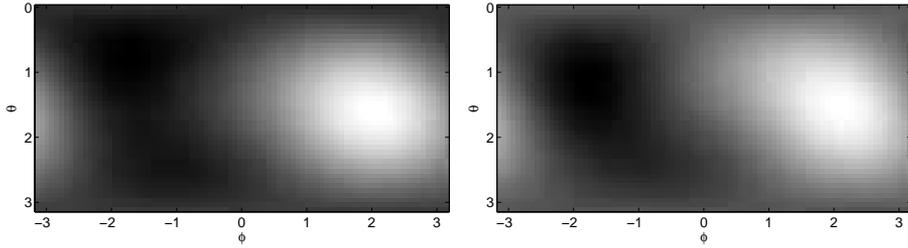
Figure 5.4: Low-pass filtered version of the groove light field (left) and the estimated light field (right). It was estimated using an Image-Centered PCA basis with six basis functions, including the constant function.

Now, let $\mathbf{d}$ be a vector containing the estimated coefficients and $\mathbf{U}$ the matrix containing the column vectors of spherical harmonics coefficients of $U_i, i = 0, ..., N$. Then the light field coefficients are given by $\mathbf{L} = \mathbf{Ud}$, according to equation (4.34).

Since the surface acts as low-pass filter on the illumination, only the lower frequencies of the light field are recoverable. Figure 5.4 shows the low-pass filtered light field of the groove illumination map and its estimated counterpart. The light field was estimated using an Image-Centered basis of five principal components plus the constant function (Y_0^0).

Chapter 6

A Basis for a Group of Materials

The apperance of a surface depends on the material of the surface. It would be very useful if we could have a single basis to represent the images from a set of materials. In this chapter we construct a basis using a database of real world surface reflectances, the CURET database (Dana et al. 1999), and a database of captured illumination, (Debevec 1998). The goal is to create a basis that can represent the images of a surface for a wide variety of surface reflectance properties and under most illumination conditions.

6.1   Computing the Basis

The difference from the bases for Lambertian surfaces is that we are now considering variations in the surface reflectance properties. Hence, it is necessary to compute the covariances of the coefficients of the surface reflectance functions.

Also, since many of the BRDFs in the databases are non-Lambertian we need to find a new finite approximation of the frequency space image irradiance representation. Nine basis functions are no longer sufficient. To represent the BRDFs in the CUReT database we use material modes up to $o = 7$. Now, according to Chapter 3 illumination orders up to $l = 12$ are enough to obtain a decent approximation of the image irradiance. This limits the number of image irradiance functions to 3704. Still many of those functions contribute little to the image irradiance. To reduce the number of functions further we choose only the 1000 basis functions that have the highest variance in the image on average (over the illumination and BRDF changes).

As in the case of the Lambertian surfaces we compute the basis for continuous images of a sphere. The matrix $\mathbf{M}$ and $\mathbf{e}$ are computed in the same way, with the difference that we have to compute the covariances of the 1000 functions in the selected image irradiance representation. The actual computation of the elements of $\mathbf{M}$ and $\mathbf{e}$ was achieved by solving the integrals in (5.1) and (5.2) analytically.
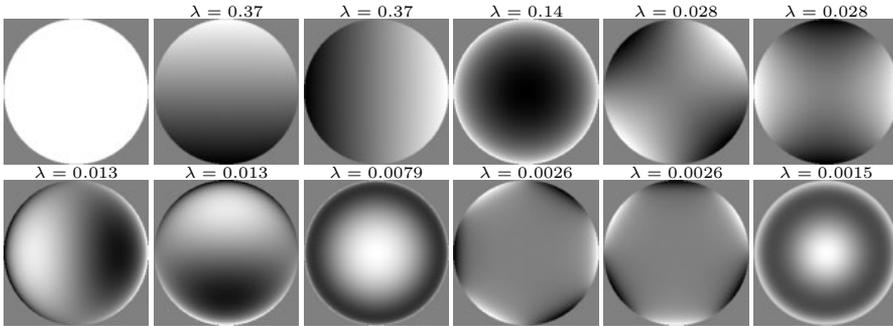
Figure 6.1: The first 12 basis functions for a sphere created from set of real world materials and illuminations using Image-Centered PCA. The top left image is the constant basis function added to the basis after the PCA is performed.

The result is a quadruple sum of various factorials. To ensure accurate computation of this sum, it was computed using infinite precision rational numbers.

The illumination distribution is the nine illumination maps undergoing all rotations and the illumination covariances are computed as in Section 5.2.4.

The variations in the surface reflectance are described by the set of $n_b = 61$ materials in the database. With each material k having BRDF coefficients $b_{op,k}^q$, the covariances and means are computed as

$$\mathrm{E}\{b_{op}^q b_{o'p'}^{q'}\} \quad = \quad \frac{1}{n_b} \sum_{k=1}^{n_b} b_{op,k}^q b_{o'p',k}^{q'} \tag{6.1}$$

$$\mathrm{E}\{b_{op}^q\} \quad = \quad \frac{1}{n_b} \sum_{k=1}^{n_b} b_{op,k}^q \tag{6.2}$$

From the means and covariances of the illumination and the BRDFs and the matrix $\mathbf{M}$, the basis is computed as described in Section 4.3. Figure 6.1 shows the 12 first basis functions of the Image-Centered basis.

6.2   Required number of components

When determining how many principal component that are needed to achieve a given accuracy, the normal procedure is to sum up the eigenvalues to see at which principal component the required percentage of the variance is accounted for. In this case this is not possible since the variance of the constant function is not included in the eigenvalues.

However, model-based PCA allows the variance of the constant function to be computed analytically. It also allows for the computation of the variances for the images of a single material and not only for the whole set.

The images are decomposed in the principal component basis, including the constant function, as $I = \sum_i d_i U_i$. The coefficients d_i can be computed from the illumination and BRDF by $\mathbf{d} = \mathbf{Ac}$, where $\mathbf{A} = \mathbf{U}^T \mathbf{M}$ and $\mathbf{c}$ is the vector with the products of the illumination and BRDF coefficients (as described in Section 4.4). The variance of each component is d_i^2. For a single material the expected variance of each component can be computed by

$$\mathrm{E}\{d_i^2\} = \sum_{j,k} a_{ij} a_{ik} b_{o(j),p(j)}^{q(j)} b_{o(k),p(k)}^{q(k)} \, \mathrm{E}\{L_{l(j)}^{m(j)} L_{l(k)}^{m(k)}\}, \tag{6.3}$$

where a_{ji} are the elements of $\mathbf{A}$.

Using the above formula we can compute the required number of basis functions to achieve a given accuracy. Figure 6.2 shows the required number of basis functions for the materials in the CUReT database under a point light source. As can be seen, many of the materials require only around five basis functions to capture most of the variance in the images. These are the matte materials. Materials with specular reflectance, such as Leather, Rough Tile and Slate a and b, require up to 40 basis functions.

6.3 Testing the Basis

To test the basis we rendered images of a sphere of each material under each illumination map. For each illumination map four views of each materials were rendered, resulting in a total of 2196 images. The rendering was performed in the same manner as in the previous chapter. Figure 6.3 shows three of the rendered images.

The testing was done by fitting the basis to the images and computing the average error of the dataset. Both the Standard PCA and the Image-Centered PCA basis were tested. The number of basis function used varied between 1 and 16. The error was chosen to be the variance of the residual divided by the total variance of the image. Figure 6.4 shows the results obtained. The Standard PCA and Image-Centered PCA bases show similar results. Note that fitting the Standard PCA basis requires subtracting the mean image before fitting the basis. This is not required for the Image-Centered basis.

As in the previous section we can compute a theoretical value of the error of the dataset. This predicted error is also plotted in Figure 6.4. The predicted error follows the same trend as the empirical error, but underestimates the error. There are a number of possible reasons for this. For instance, the images are rendered with perspective projection while the model approximates the projection as orthographic. Also the dataset only contains four rotations of each illumination map, while the theoretical value takes into account every possible rotation.

The number of components required to represent an image depends on the material. This is illustrated by Figure 6.5 which shows the errors for each individual material. Many materials require only a small (around five) number of components,
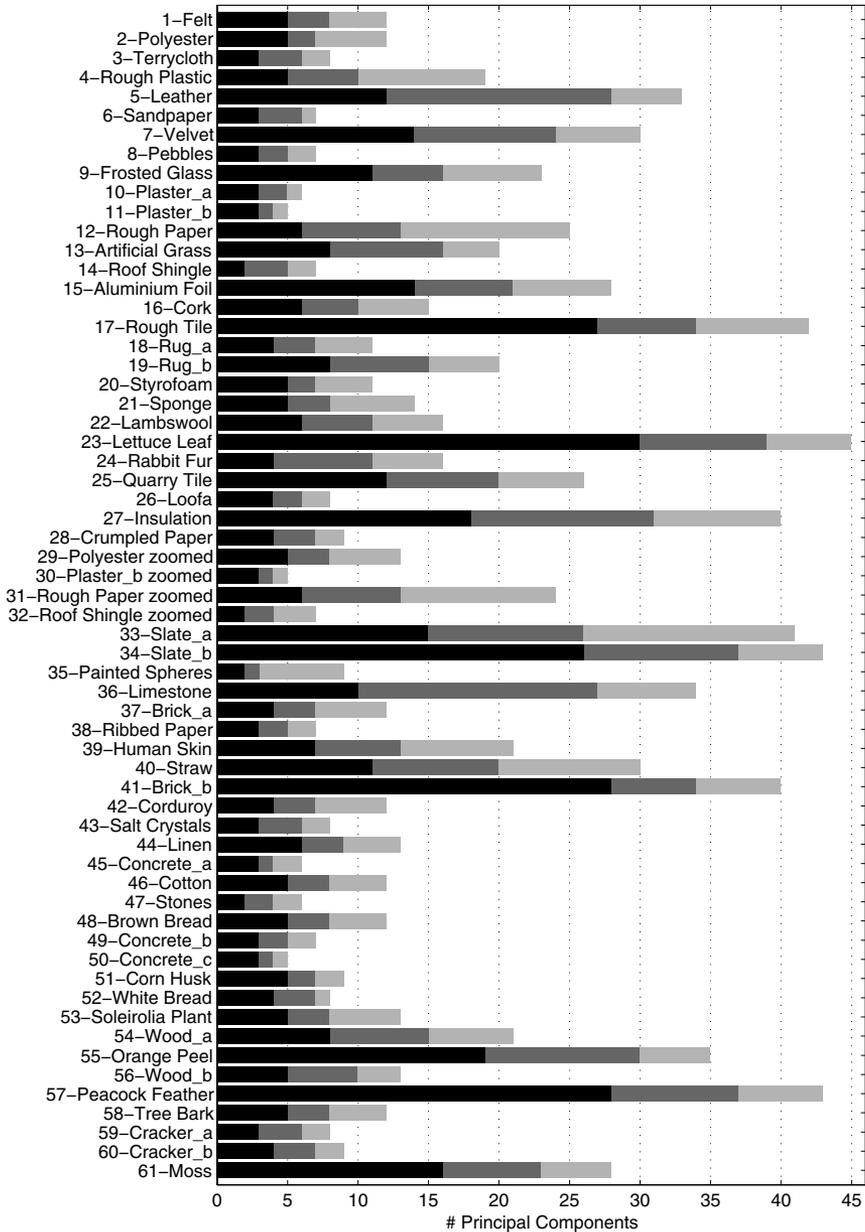
Figure 6.2: Required number of principal components for the materials in the CUReT database. The black, dark gray and light gray bars show the required number of basis functions to achieve 95%, 98% and 99% accuracy respectively. The illumination is a point light source.
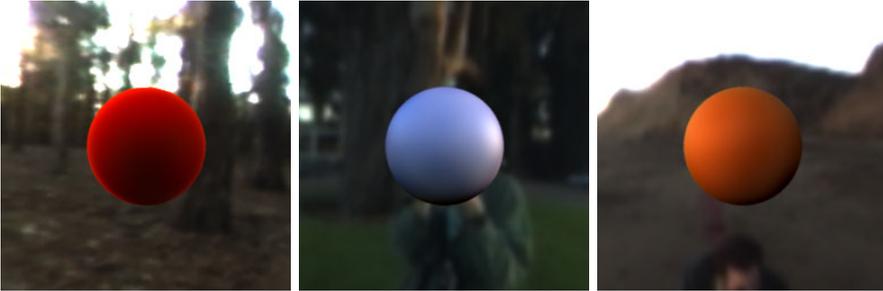
Figure 6.3: Three of the rendered images. From left to right, velvet in a eucalyptus grove, leather on campus and an orange on the beach. Notice the characteristic bright rim of the velvet sphere and the specularity on the leather.
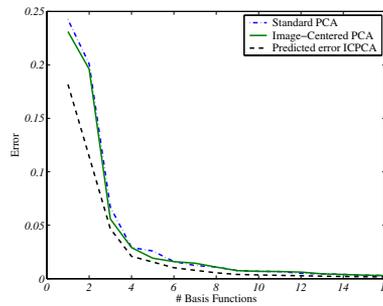
Figure 6.4: Average error for the whole data-set as the number of principal components increases for the basis created with Standard PCA and Image-Centered PCA. The error is normalized so that $Accuracy = 1 - Error$.

such as the matte materials, but also highly non-Lambertian materials like velvet, see Figure 6.6a.

More specular materials, such as the leather in the database, require more components. The diffuse shading is recreated with a few components but to recreate the specularity 30-50 components are required in this case, see Figure 6.6b.

Figure 6.6 also shows that the accuracy, as computed here, does not always correspond to what our visual system is sensitive to. For example, the leather sphere only requires five basis functions to reconstruct 99.1% of its variance, but to achieve a visually realistic leather sphere requires more than 30 basis functions. The velvet sphere on the other hand is more or less indistinguishable from the original image at 4 basis functions giving 99.1% accuracy.

1	Felt	11	Plaster_b	21	Sponge	35	Painted Spheres	50	Concrete_c
2	Polyester	12	Rough Paper	22	Lambswool	36	Limestone	52	White Bread
3	Terrycloth	13	Artificial Grass	23	Lettuce Leaf	37	Brick_a	53	Soleirolia Plant
4	Rough Plastic	14	Roof Shingle	24	Rabbit Fur	39	Human Skin	55	Orange Peel
5	Leather	15	Aluminium Foil	25	Quarry Tile	41	Brick_b	59	Cracker_a
6	Sandpaper	16	Cork	26	Loofa	43	Salt Crystals	60	Cracker_b
7	Velvet	17	Rough Tile	27	Insulation	45	Concrete_a	61	Moss
8	Pebbles	18	Rug_a	28	Crumpled Paper	47	Stones		
9	Frosted Glass	19	Rug_b	33	Slate_a	48	Brown Bread		
10	Plaster_a	20	Styrofoam	34	Slate_b	49	Concrete_b		

Figure 6.5: Error as the number of principal components increases for each material in the CURET database.

Number of principal components (Accuracy)

(a)

Number of principal components (Accuracy)
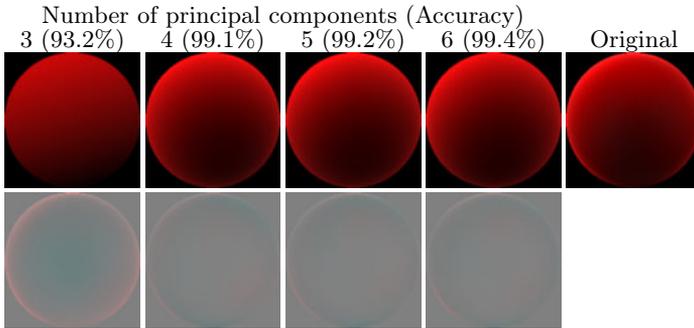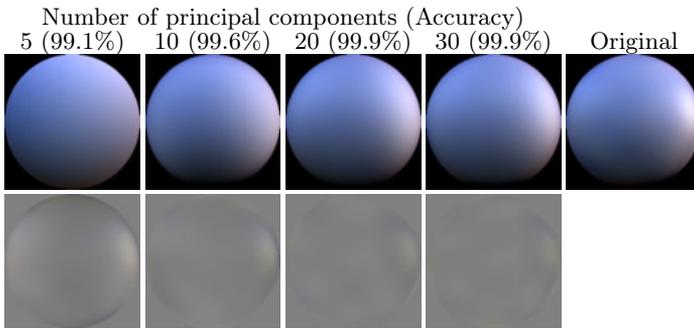
(b)

Figure 6.6: Reconstructed and residual images of a) velvet and b) leather.

6.4   Conclusions

Using databases of captured BRDFs and captured illumination maps it is possible to create a basis that can represent the images of a wide variety of materials under a wide variety of illuminations.

In a basis created from the CUReT databases and the illumination maps of Debevec the images are accurately represented using between 5 and 40 basis function depending on the surface material. Matte materials, including highly non-Lambertian materials such as velvet, are well represented using around five basis functions, while materials with specular reflectance require up to 40 basis functions.

The accuracy of the basis is measured by the percentage of variance the basis accounts for. However, this measure does not always correspond to the sensitivity of the human visual system. The reconstructed image of a matte surface can look indistinguishable from the original image at 99.1% accuracy, while a specular material may require up to 99.9% accuracy to be indistinguishable from the original image. The same observation was made by Epstein, Hallinan, and Yuille (1995).

Part II

Image Analysis

Chapter 7

Classifying Materials from their Reflectance Properties

7.1 Introduction

The appearance of a surface depends on its shape, the illumination and the material of the surface. In a normal vision task none of these properties are known a priori. Despite that, human observers are very good at determining the material of an object, even in the absence of texture. The estimation is based purely on the reflectance properties of the surface. We explore if this can be done computationally when there is no knowledge about the illumination, but the shape of the object is known.

The framework in Chapter 4 allows us to create, for a given shape, a single basis that can represent the images of a surface with varying reflectance properties and under arbitrary illumination.

Assuming knowledge of the surface shape may seem unrealistic. However, the proposed algorithm is not intended as a final solution but rather for use in a probabilistic framework, where in the end, the information is extracted when none of the properties of the scene is known beforehand. Note that knowledge of the object's shape is not required to use such an algorithm. We only need an estimate of how likely the different shapes are, given an image. This can be seen by examining the factorization of the posterior distribution for the illumination L, the material M, the shape S given the image I.

$$P(L, M, S|I) = P(L, M|S, I)P(S|I) \qquad (7.1)$$

This chapter only deals with the first term, $P(L, M|S, I)$, but the low-dimensional basis could be used to estimate $P(S|I)$ as well. Because the basis is a generative model we can compute a goodness of fit of the model, giving an indication of how likely the shape is given the image. It does not however reveal the full answer since
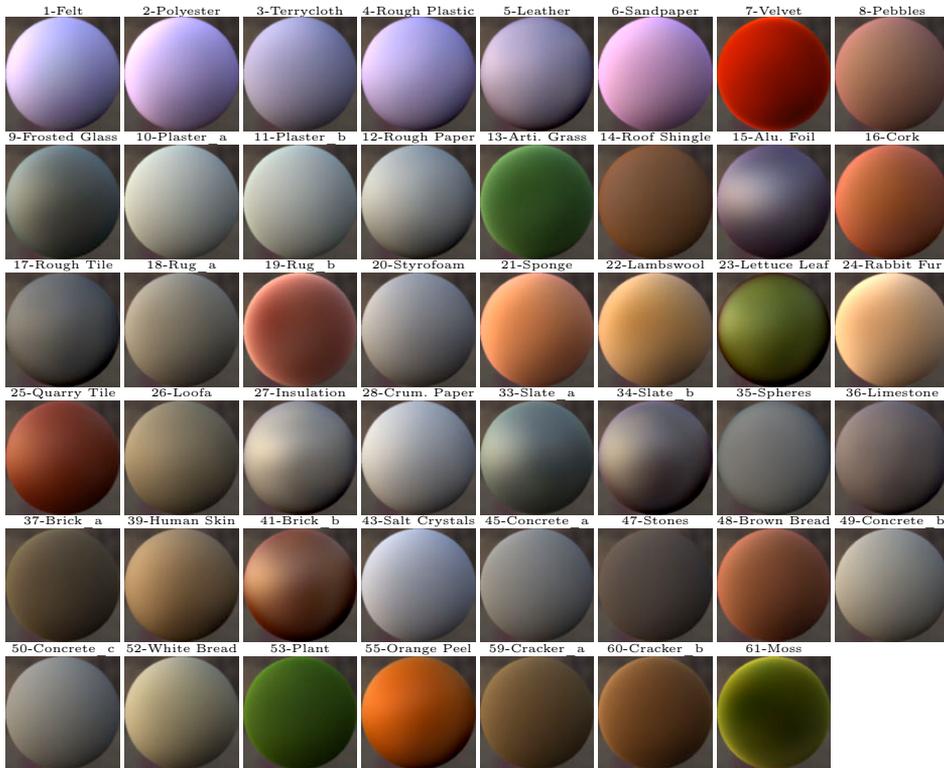
Figure 7.1: Rendered images using BRDFs from the CUReT database. Classifying materials from their reflectance properties can frequently be very hard, especially If you disregard the color information.

there could be several shape-material-illumination combinations that give rise to the same shading pattern.

In this chapter we classify the material of an object of known shape from a single image, when the illumination is unknown. Dror et al. (2001) recognize materials under similar assumptions. They use histograms of filter responses and rely on the structure of the specular reflections to classify the material. Our approach is different in that we represent the images using a generative model, allowing us to discriminate between materials without specular reflections such as felt and velvet.

7.2   Material Recognition

Our approach to material recognition is to represent the images in an Image-Centered PCA basis constructed as in Chapter 4. The image's coefficients in the
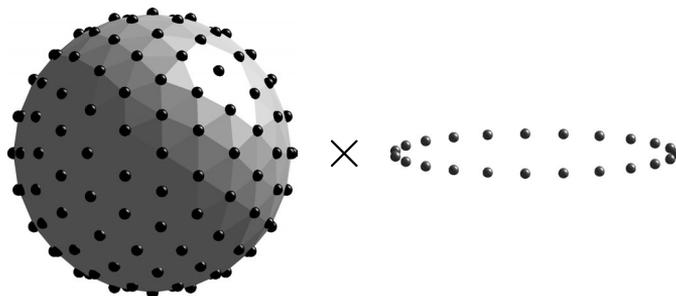
Figure 7.2: Sample density of the 3D rotation group $SO(3)$. $SO(3)$ is sampled by combining a sphere and a circle. The sphere is sampled by subdividing an icosahedron. The circle is sampled in the same density as the sphere. The sampling above is the result of two recursive subdivisions and is what is used to build the appearance manifolds.

basis are estimated and compared to a database of known materials.

Since the illumination is not known we cannot calculate what the corresponding coefficients should be for the materials in the database. We need to take into account all possible illuminations and find the illumination-material pair that best matches the image. For this to be possible it is necessary that the variations in the coefficient space are much smaller than the variations in the illumination (which are infinite). If this is true we can learn the variations in the coefficient space with only a limited amount of training illuminations.

Smooth variations in the illumination result in a manifold of points in the coefficient space. To learn these manifolds we take a set of illumination maps and perform all 3D rotations on them. To store the manifolds we sample the rotation group, $SO(3)$, and calculate the coefficients for each rotation, for every illumination map and material.

The image is classified by finding the manifold which is closest to the point representing the image. The procedure is very much the same as in (Murase and Nayar 1995).

7.2.1 Learning the Manifolds

The manifold for each material is learned from a set of illumination maps that are rotated over the full rotation group. The rotation group is sampled and for each rotation (α, β, γ) the spherical harmonic coefficients of the rotated illumination map are calculated. The point on the manifold is given by equation (4.35).

To sample the rotation group we sample the surface of a sphere and combine it with a circle. The sphere is sampled by starting from an icosahedron inscribed in the sphere. The icosahedron is recursively subdivided by projecting the mid-point of each edge onto the surface of the sphere forming four new triangles for each old triangle, (Ballard and Brown 1982). The circle is sampled at a density as close as
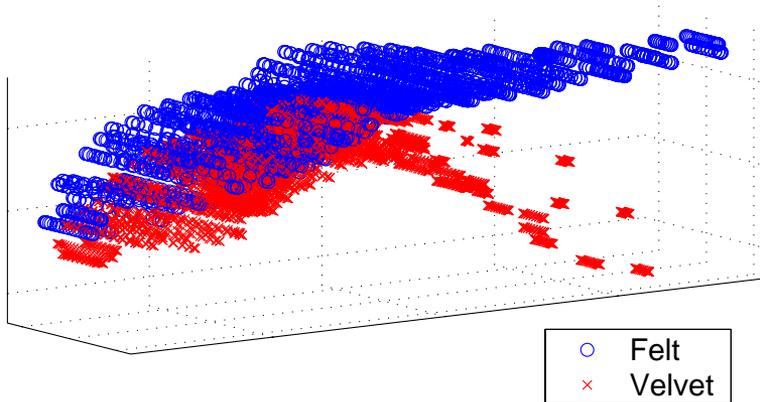
Figure 7.3: Sampled manifolds in the coefficient space of materials 1-Felt (blue rings) and 7-Velvet (red crosses) under one of the illumination maps subject to all 3D rotations, $SO(3)$.

possible to the sampling of the sphere. Figure 7.2 shows the sample density used in the learning of the manifolds.

7.2.2   Finding the Closest Manifold

To find the closest manifold to a point we simply go through all points on each manifold and calculate the distance to the point to be classified. The distance measure is the sum of squared differences in coefficient space.

To aid our algorithm in being illumination invariant we take a number of steps. The first element of the point is discarded. It corresponds to the constant function of the basis and captures the variations in the ambient component of the illumination. By discarding it the algorithm becomes independent of such variations. The remaining elements are normalized to achieve brightness independence. This last step comes at the cost that we will not be able to differentiate between bright and dark materials. Although this could ambiguity could be addressed at a later stage by comparing the signal variances of the images.

7.3   Discrimination of Materials in the CUReT database

Before we move on to real images we need to assess what can be done. How well can materials be discriminated from their reflectance properties alone? Figure 7.1 demonstrates that frequently many materials look similar to the human eye.

To investigate the feasibility of this task we analyze how well the materials in the CUReT database can be discriminated in synthetic images, i.e. when there is

no noise. The illumination is considered to be unknown. The algorithm is tested on images generated from one of the illumination maps, while the other illumination maps are used to build the manifolds for classification. This is repeated for all nine illumination maps (the leave-one-out principle).

There is no actual need to generate any images. Using the low-dimensional basis framework we can directly from the illumination and material coefficients compute the coefficients in the low-dimensional basis of the image. This fact allows for extensive testing. Each of the 47 materials used is tested with nine illumination maps, each under 462 different rotations, summing up to a total of almost 200 000 images used for testing. The 20 first basis functions were used in the classification.

Figure 7.4 shows the classification rates for the different materials. The correct classification rates, which can be seen in the diagonal, range between 5 and 80 percent. Materials with a high classification rate are 7-velvet and 61-moss which have particular reflectance properties. Glossy materials have in general a higher recognition rate than matte materials.

Of particular interest is that the materials seem to cluster into groups. Often a material in a group is systematically mis-classified as one of the other materials in that same group. This becomes apparent when we order the materials in a particular way. Figure 7.5 shows the exact same classification rates as Figure 7.4, but with the materials ordered using a hierarchical grouping algorithm, described in the next section. Blocks become apparent in the diagonal of the matrix. There is a large block of matte materials in the top left corner, formed by the materials 1-Felt, 20-Styrofoam, ..., 24-Rabbit Fur. Following the matte materials is a group of glossy materials, 4-Rough Plastic , ..., 15-Foil. Last comes 7-Velvet and a group of velvet-like appearance (asperity scattering), 13-Artificial Grass, 19-Rug_b and 61-Moss. Finally we have 35-Painted Spheres which forms a group of its own.

7.3.1 Visual Grouping of the Materials

It is clear that we cannot expect to distinguish between some of the materials in the CUReT database. Instead we can try to find groups in which to classify the materials.

Using the matrix containing the classification rates we group the materials. The grouping is done in a greedy fashion. We start with groups of single materials. Then the two groups that maximize the average recognition rate are joined. This is repeated until the desired number of groups is reached. To select the number of groups one can look at the ratio between the recognition rates and the rate of selecting the correct material by chance.

Dividing the CUReT database into 9 groups results in the grouping in Figure 7.6. We have labeled the groups according to the characteristics of their members. All matte materials end up in one group. Materials having specular reflectance are split up into three groups. The last five groups are materials that did not fit into any group. These materials have a high recognition rate on their own.

Figure 7.4: Recognition rates for the CUReT materials. Each row shows the classification rates for a particular material, e.g. the leftmost element in the first row is the rate that material no. 1 is classified as material no. 1, the second element is the rate the material no. 1 is classified as material no. 2. The diagonal is the correct classification rate. These results are discussed in more detail in the text.

1	Felt	11	Plaster_b	21	Sponge	35	Painted Spheres	50	Concrete_c
2	Polyester	12	Rough Paper	22	Lambswool	36	Limestone	52	White Bread
3	Terrycloth	13	Artificial Grass	23	Lettuce Leaf	37	Brick_a	53	Soleirolia Plant
4	Rough Plastic	14	Roof Shingle	24	Quarry Tile	39	Rabbit Fur	55	Orange Peel
5	Leather	15	Aluminium Foil	25	Quarry Tile	41	Brick_b	59	Cracker_a
6	Sandpaper	16	Cork	26	Loofa	43	Salt Crystals	60	Cracker_b
7	Velvet	17	Rough Tile	27	Insulation	45	Concrete_a	61	Moss
8	Pebbles	18	Rug_a	28	Crumpled Paper	47	Stones		
9	Frosted Glass	19	Rug_b	33	Slate_a	48	Brown Bread		
10	Plaster_a	20	Styrofoam	34	Slate_b	49	Concrete_b		

	1	Felt	11	Plaster_b	21	Sponge	35	Painted Spheres	50	Concrete_c
	2	Polyester	12	Rough Paper	22	Lambswool	36	Limestone	52	White Bread
	3	Terrycloth	13	Artificial Grass	23	Lettuce Leaf	37	Brick_a	53	Soleirolia Plant
	4	Rough Plastic	14	Roof Shingle	24	Rabbit Fur	39	Human Skin	55	Orange Peel
	5	Leather	15	Aluminium Foil	25	Quarry Tile	41	Brick_b	59	Cracker_a
	6	Sandpaper	16	Cork	26	Loofa	43	Salt Crystals	60	Cracker_b
	7	Velvet	17	Rough Tile	27	Insulation	45	Concrete_a	61	Moss
	8	Pebbles	18	Rug_a	28	Crumpled Paper	47	Stones		
	9	Frosted Glass	19	Rug_b	33	Slate_a	48	Brown Bread		
	10	Plaster_a	20	Styrofoam	34	Slate_b	49	Concrete_b		

Figure 7.5: When the classification rates from Figure 7.4 are sorted in a particular way a pattern emerges. The materials form groups. Materials within a group are often classified as one of the other materials in the same group. The largest group can be seen as a grey block in the top left corner of the matrix. These are the matte materials, 1-Felt, 20-Styrofoam, ..., 24-Rabbit Fur. After the matte materials comes a group of more glossy materials, 12-Rough Plastic, ..., 36-Limestone. Next comes a group of shiny materials 9-Frosted Glass to 33-Slate_a. Last is a group of materials with asperity type scattering, 7-Velvet, 13-Artificial Grass, 19-Rug_b and 61-Moss.

Group	Members	Label
1	1, 2, 3, 6, 8, 10, ...	Matte
2	4, 5, 12, 17, 36, 55	Glossy
3	9, 23, 25, 27, 34, 41	Shiny
4	15, 33	Shinier
5	7	Velvet
6	13	Art. Grass
7	19	Rug
8	61	Moss
9	35	Spheres
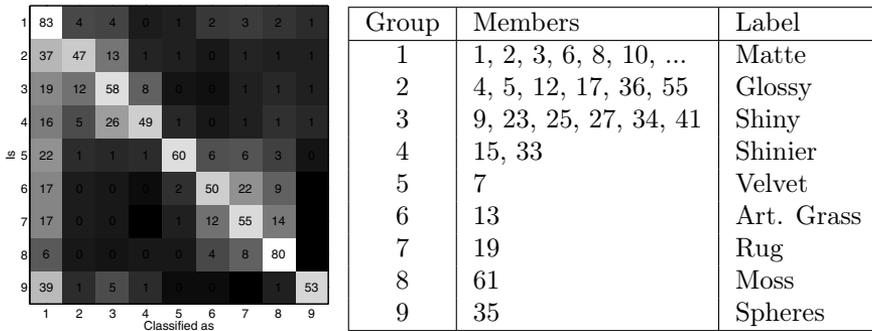
Figure 7.6: Classification rates when the materials are grouped into nine groups. Not all members were listed in the matte group due to space limitations, but this group contains all materials that are not in the other groups.

Most of the groups are sometimes mis-classified as matte materials. This makes sense. In the testing we take all rotations of the illumination into account. This means that sometimes the dominant light source in the scene will be behind the object. Hence, there will be no specularity on the object to differentiate it from a matte material.

7.4   Classifying the Material in Real Images

To test the algorithm we glued five different real materials onto cylinders, see Figure 7.7. Cylinders were chosen due to the difficulty of gluing non-stretchable materials onto a sphere. The cylinders where photographed using a digital camera in different illumination conditions, including outdoor sunny, outdoor cloudy and indoor conditions with indirect light from a window. Before classification the images were radiometricly calibrated, using the method in (Mitsunaga and Nayar 1999). The geometry of the cylinders were estimated by manually marking where in the image the cylinders were.

Using the framework from Chapter 4 we computed a basis for the cylinder. A total of six basis functions were used in the experiments. The coefficients for the image were estimated by projecting the image onto the basis. The image was then classified by finding the closest manifold as described in Section 7.2. The manifolds were this time learned using all nine illumination maps.

Figure 7.8 shows some of the images being classified. Note how well the basis is able to represent the image irradiance in all cases.

A total of 84 images were used in the experiment. Table 7.1 summarizes the results. As predicted by the synthetic experiments only a few of the images where correctly classified on an individual basis. Felt and the two velvets have a recognition rate of 5% to 7.7%, which is still several times greater than chance, which is $1/47 \approx 2.1\%$. When using the grouping in Figure 7.6 the recognition rates are
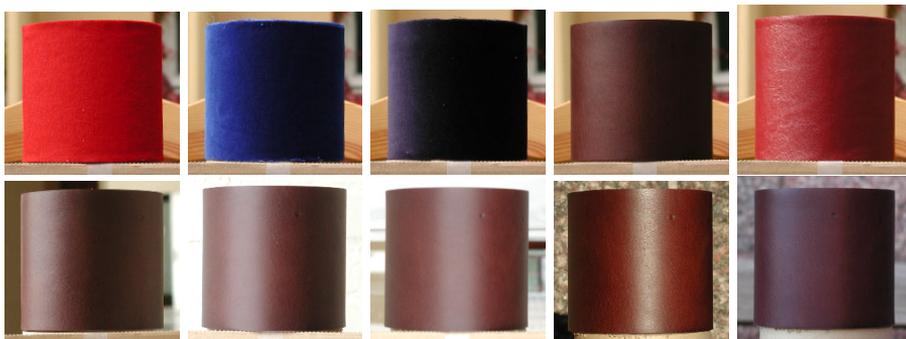
Figure 7.7: The algorithm was tested on images of cylinders with the pieces of five different real materials glued onto them. Top row from left to right: felt, velvet 1, velvet 2, leather and imitation leather. Bottom row: leather in five of the different illumination conditions.

higher. Felt is to a large extent classified as matte. The leather here is classified as Shiny or Shinier, while the leather in the database is categorized as Glossy. This could be because our leather is shiner than the leather in the database. Visually, at least, it appears so. The imitation leather is also mostly classified as Shiny or Shinier.

So far the results match the synthetic results fairly well. The velvet however does not. The synthetic results indicate that velvet should be fairly easy to recognize, but in our experiments the two velvet cylinders are mostly classified as matte. On the other hand, they are also often classified as either one of the groups Grass, Rug or Moss, which have the same type of surface reflectance as Velvet.

7.5 Conclusions

This chapter has demonstrated the use of the low-dimensional basis framework described in Chapter 4. Using this framework a basis to represent the images of the materials in the CUReT database is constructed. The variations in the appearance of the surface of a particular material form manifolds in the space of the low-dimensional basis. These manifolds are learned from the captured illumination maps of Debevec. The material in an image is classified by finding the manifold which is closest the image.

Recognizing the material is hard. Analysis of the recognition rates on synthetic data shows that many materials often are incorrectly classified. The materials seem to form groups where a material in a group often is mis-classified as one of the other materials in the same group. A simple greedy grouping algorithm is able to find sensible groups from a human visual system standpoint.

The results of the experiments on real images of materials to a large extent follow those predicted by the synthetic experiments. They indicate that the method
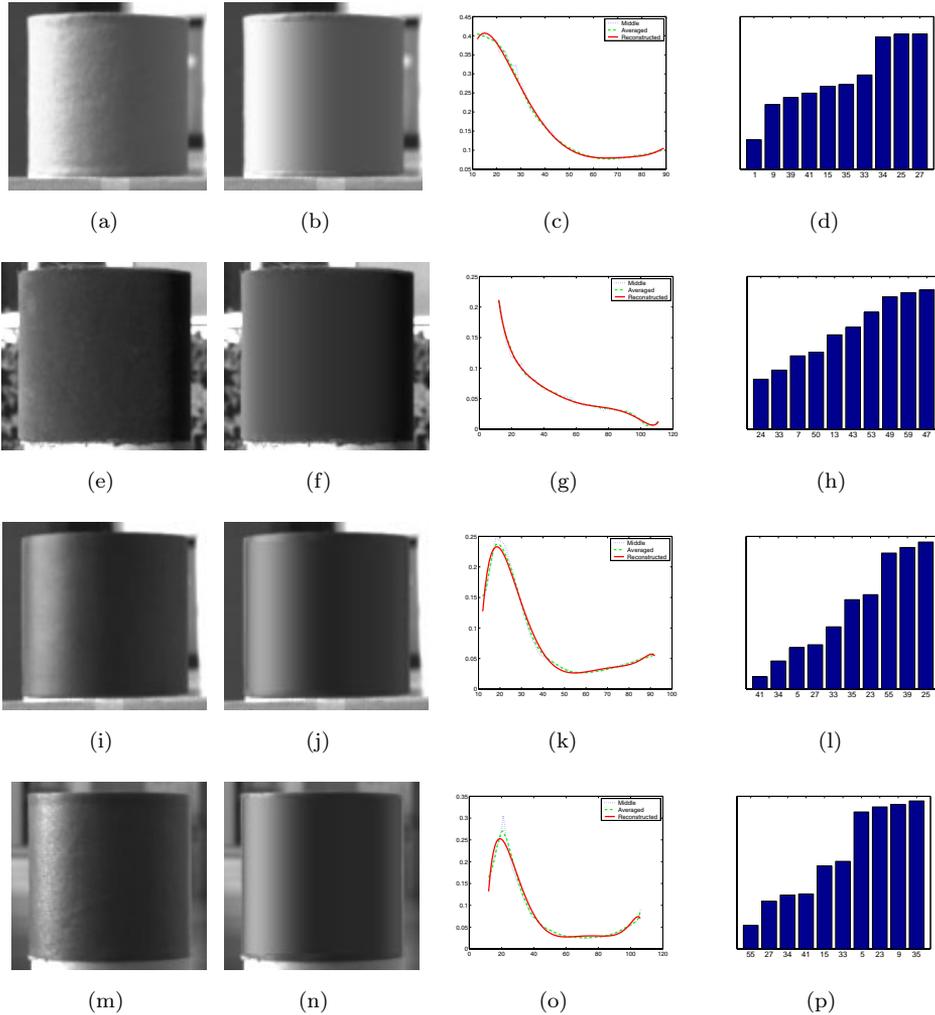
Figure 7.8: Examples of classified images: (a)-(d) images for Felt. (a) calibrated gray image, (b) reconstructed gray image (this is what the algorithm "sees"), (c) image and reconstructed intensity profiles. (d) distances to the ten closest materials. Here the material is correctly classified as felt. (e)-(h) show the same images for Velvet 1. The material is here incorrectly classified as 24-Rabbit Fur, 7-Velvet comes third place. (i)-(l) images for leather which in this case is classified as 41-Brick_b, 5-Leather is the third closest material. (m)-(p) imitation leather: classified as 55-Orange, 5-Leather on seventh place. Notice how well the basis represent the irradiance for the different cases.

Material	Correct	Matte	Glossy	Shiny	Shinier	Velvet	Grass	Rug	Moss	Spheres
Felt	7.7	77	7.7	0	15	0	0	0	0	0
Leather	0	25	6.2	44	19	6.2	0	0	0	0
Im. Leather	0	10	0	40	35	10	0	0	0	5
Velvet 1	5	55	0	5	5	5	10	10	10	0
Velvet 2	6.7	40	0	33	0	6.7	0	13	6.7	0

Table 7.1: Classification Rates for the Cylinder Images.

generalizes to images of new samples of the materials, taken under varying and different illumination conditions.

Chapter 8

Projected Light Source Direction From Occluding Contours

This chapter presents an algorithm that automatically estimates the projected direction (the slant direction) to the light source using shading near occluding contours in the image.

The work was completed when the frequency space representations of shading was largely unknown to the computer vision community and therefore the algorithm uses conventional models such as a point light source and Lambertian surface reflectance.

8.1   Introduction

The occluding contour is where an object occludes itself like the earth at the horizon. At the occluding contour the shape of the object can easily be determined because the surface normal is perpendicular to the viewing vector and can be determined by the image edge direction (Horn 1986).

Many existing algorithms for estimating the light source direction use occluding contours. In (Pentland 1982) it is a requirement that the image is of a convex object bounded by an occluding contour. Vega and Yang (1994) use the same occluding contour assumption to match default shapes to the image and then estimate the light source direction. Both methods require Lambertian surfaces and a segmented image of the object. Yang and Yuille (1991) also exploit the occluding boundary and show that it puts strong constraints on the light source direction.

Other algorithms use known geometry to derive the illumination direction. (Sato et al. 1999a; Sato et al. 1999b) derive the illumination distribution by studying shadows around an object of known geometry. In (Zhang and Yang 2000) multiple light sources are extracted from a sphere of known size.

Previous work depends on segmented images or known geometry. In this chapter we do not make these assumptions, but only determine the projected light source direction. We will present a fully automatic algorithm for recovering this projected direction. First the estimation of the light source direction using the shading near the occluding contour is investigated. Also, the noise distribution is derived. Secondly an algorithm for picking out potential occluding contours using edge and color information is presented. The contours produced by this algorithm are then used in a Bayesian probabilistic framework to estimate the most likely light source direction. Simultaneously, the contours are classified as occluding or not.

8.2   Shading at the Occluding Contour

First we will look at the case of estimating the light source direction from an occluding contour.

Given that the illumination is a single point light source, the model for the image intensity I at a point on a Lambertian surface is

$$I = k(\vec{n} \bullet \vec{l}) + a \tag{8.1}$$

where $\vec{n} = (n_x, n_y, n_z)^T$ is the surface normal at the point, $\vec{l} = (l_x, l_y, l_z)^T$ is the direction to the light source, k is a parameter containing both the surface albedo and the strength of the light source, a is a term representing the contribution of the ambient illumination.

On the occluding contour n_z is equal to zero. This will eliminate l_z from the equation, which is why we will not be able to estimate l_z. Due to the bas-relief ambiguity (Belhumeur, Kriegman, and Yuille 1997) the z-component of the light source direction cannot be estimated using only this model when the surface albedo and light source strength is unknown. Now, the image intensity on the occluding contour will be

$$I = k(\vec{n} \bullet \vec{l}) + a = k(n_x l_x + n_y l_y) + a. \tag{8.2}$$

From the image we can measure $\vec{n}$ and I. $\vec{l}$, a and k are unknown, but are assumed to be constant for each local computation. Since we only are interested in the direction to the light source scaling of the $\vec{l}$ vector doesn't matter. Therefore let $x = kl_x$ and $y = kl_y$. Now we have

$$I = n_x x + n_y y + a = \left(\begin{array}{ccc} n_x & n_y & 1 \end{array}\right) \left(\begin{array}{c} x \\ y \\ a \end{array}\right) \tag{8.3}$$

Equation (8.3) has three unknowns x, y and a and as many equations as the number of points on the occluding contour. Using e.g. least squares we can estimate the light source direction. Let $\mathbf{N}$ be a matrix containing the n_x and n_y for the points

along the contour and **I** a vector containing the intensity values for the same points. Then the least-squares estimate $\hat{L} = (\hat{x}, \hat{y}, \hat{a})^T$ is calculated by

$$\hat{L} = (\mathbf{N}^T \mathbf{N})^{-1} \mathbf{N}^T \mathbf{I}. \tag{8.4}$$

In $\hat{L}$ the estimated angle towards the light source is found as the direction of the vector $\hat{l} = (\hat{x}, \hat{y})^T$, i.e.

$$\hat{\phi} = \arg(\hat{l}). \tag{8.5}$$

8.2.1 Measuring the intensity at the contour

To be able to estimate the light source direction in this way we need to measure the intensity at the occluding contour. This is of course impossible, the intensity cannot be measured at the contour. Also, the Lambertian model becomes inaccurate very close to the contour (Oren and Nayar 1995; Wolff 1994). To overcome this we will look at the intensities some distance away from the contour and extrapolate to get the intensity at the contour.

The extrapolation is done by modeling the image intensities along a line, the u-axis in Figure 8.1a, perpendicular to the edge. By measuring the intensities along the line and using the model, the intensity at the edge can be estimated by extrapolation.

To model the intensities, a model of the shape along the line is needed. We have chosen an ellipse to model the cross-section of the object, see Figure 8.1b. Because of the bas-relief ambiguity the z-axis can be scaled arbitrarily and the shape model can be simplified as a circle, with radius R, without any loss of generality in this case. We have

$$(u - R)^2 + z^2 = R^2 \tag{8.6}$$

from which we can express z as function of u and the normal of the surface as a function of u and z

$$z = \sqrt{2uR - u^2} \tag{8.7}$$

$$N(u) = \frac{1}{R} \begin{pmatrix} u - R \\ z \\ 0 \end{pmatrix}. \tag{8.8}$$

Inserting the expression for the normal into the intensity equation (8.1) gives us the intensity as a function of u,

$$I(u) = \frac{k}{R} \left(l_u(u - R) + l_z \sqrt{2uR - u^2} \right) + a. \tag{8.9}$$

This model is too elaborate to use for extrapolation. The parameters in fact include the direction towards the light source, which is what we want to estimate.

(a)

(b)
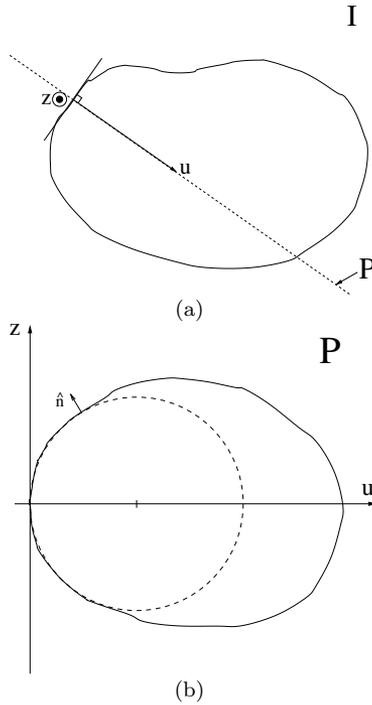
Figure 8.1: a) An image I of an object. To be able to do extrapolation a model of the shape of the object along the u-axis is needed. This is done by modeling the cross-section b) of the object in the plane P with an ellipse.

A power series expansion of the model can however be derived. With the power series we can select the models level of detail by including more or less terms from the series.

The power series of the model (8.9) is of the form:

$$
\begin{aligned}
I(u) &= c_0 + c_1 u^{1/2} + c_2 u + c_3 u^{3/2} + c_4 u^{5/2} \\
&\quad + c_5 u^{7/2} + O(u^{9/2}).
\end{aligned}
\tag{8.10}
$$

The extrapolation is done by fitting the power series polynomial to the measured intensities using least-squares estimation and then calculating the intensity at $u = 0$ using the estimated polynomial, i.e. $I(0) = c_0$.

8.2.2   The Probability Distribution of the Estimates

The estimated vector $\hat{L}$ is a linear combination of image intensities. The extrapolation operation and the estimation of the direction are both linear operations. In general each estimate will be a linear combination of hundreds of image intensi-

ties[1], which means that the estimated vector is approximately normally distributed. Since least-squares is a non-biased estimator, the mean will be $(x, y, a)^T$.

$$\hat{L} \sim N \left(\begin{pmatrix} x \\ y \\ a \end{pmatrix}, \Sigma \right) \quad (8.11)$$

The covariance matrix Σ of a least-squares estimate is

$$Cov(\hat{L}) = (\mathbf{N}^T\mathbf{N})^{-1}\sigma^2. \quad (8.12)$$

The noise variance σ^2 can be estimated with

$$\hat{\sigma^2} = \frac{1}{n-1} \sum_{i=0}^{n} (I_i - (n_x\hat{x} + n_y\hat{y} + \hat{a}))^2. \quad (8.13)$$

8.3   An Automatic Algorithm

The automatic algorithm has three stages. First a heuristic algorithm picks out candidate occluding contours using color and edge information. Secondly, for each of the contours, the light source direction is estimated according to the method presented in the previous section. In the final stage the estimates are fused in a Bayesian network to arrive at the most likely light source direction.

8.3.1   Finding Potential Occluding Contours

The goal of this stage is to, in a heuristic way, pick out contours of which as many as possible are occluding contours. The simple rules used here will naturally not be able to pick out occluding contours perfectly, but they will provide the later stages with good enough candidates.

The estimation of the light source direction described in Section 8.2 only works on occluding contours of uni-colored objects. The algorithm should pick out edge chains that have an uni-colored area perpendicular to the chain direction. Also, for the extrapolation not to be disturbed there should be no edges in that same area. The surface of the object needs to be smoothly curved so a sharp turn of the contour is a strong indication that the contour is not from a useful occluding boundary.

The algorithm works in the following way. First the edges are extracted, using the Canny edge detector, and linked into chains. By following the chains the potential occluding contours are picked out by grouping together consecutive edges if

- The area next to the edge is uni-colored

---

[1]In the experiments the extrapolation uses 7 pixels per edge point and the average contours length is about 60 which means that each estimate is a linear combination of over 400 pixel values.

- The color is the same as previous edges in the chain

- The area next to the edge contains no other edges

- The chains does not make a sharp turn

This is done on both sides of the edges. Very straight contours contain no information for our algorithm and are sifted out in a post-processing stage.

The test whether a surface is uni-colored is done in a similar way as in (Klinker, Shafer, and Kanade 1990). For each piece of a contour segment the pixels next to it are analyzed by a singular value decomposition of the pixel cluster in RGB-space. If the two eigenvectors with the lowest eigenvalues are small enough then the pixels follow the model $\vec{c}_i = k_i \vec{c}$ and are assumed to be on a uni-colored object.

The algorithm is very simple and does an acceptable job for our purposes at this stage, but there are some problems. The main problem is that contours get split up due to edges in the background. What happens is that, in the edge-linking process, the object contour gets linked together with an edge in the background instead of continuing along the contour. This could in the future be solved by using a more clever grouping algorithm incorporating e.g. good continuation. See Figure 8.9 a and d for examples.

For each of the contours picked out by the algorithm, the light source direction is estimated as described in Section 8.2.

8.3.2   Fusing the Estimates

At this stage we have a set of estimates and variances from n edge chains that might or might not be from occluding contours. To separate the correct contours from the incorrect ones it is necessary that the correct contours in general have a smaller variance , i.e. fit the model better, than the incorrect ones. Also, if we have several estimates pointing towards the same light source direction one can draw the conclusion that this is the correct direction even though their variances aren't substantially smaller than those of the the other contours. The final stage of the algorithm fuses the estimates incorporating these conditions.

8.3.2.1   The Probabilistic Model

An estimated vector $\hat{L}_i = (\hat{x}_i, \hat{y}_i, \hat{a}_i)^T$ from contour i depends on the true values $L_i = (x_i, y_i, a_i)^T$ and whether or not the contour is an occluding contour or not, captured in the discrete variable O_i.

Because only the direction of the light source vector is the same for all the contours (the magnitude depends on surface albedo which may be different for different objects), the variables x_i and y_i are represented by their magnitude r_i and direction ϕ. Figure 8.2 shows the causal dependencies between the different variables.
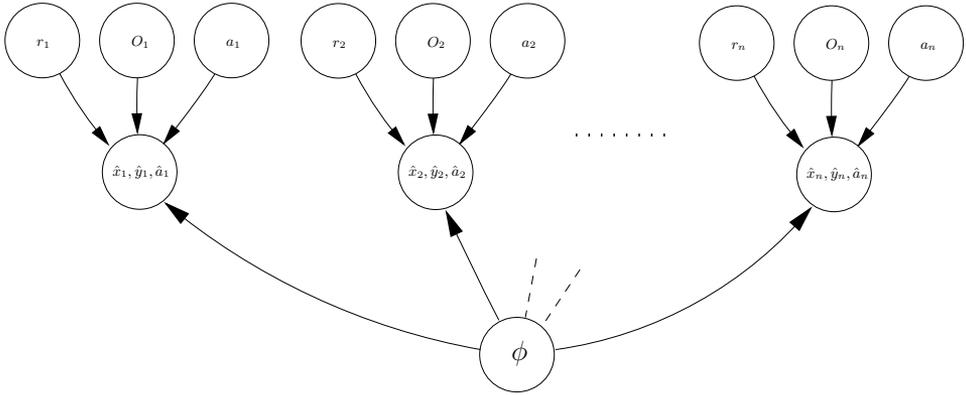
Figure 8.2: Bayesian network model
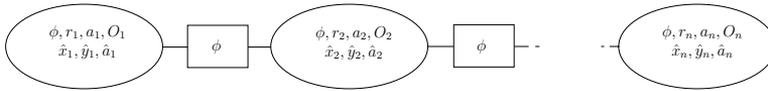
Figure 8.3: Junction tree

8.3.2.2 The Inference Process

What we would like to find is the set of values to the variables that maximizes the probability. Especially we would like to find the ϕ in that set. I.e. if $U = \{\phi, r_1, \ldots, r_n, a_1, \ldots, a_n, O_1, \ldots, O_n, \hat{L}_1, \ldots, \hat{L}_n\}$ we would like to estimate ϕ with

$$\hat{\phi} = \underset{\phi}{\mathrm{argmax}} \max_{U \setminus \phi} P(U) \tag{8.14}$$

Using message propagation this $\hat{\phi}$ can be found with only local computations (Cowell, Dawid, Lauritzen, and Spiegelhalter 1999; Jensen 1996).

By moralizing the graph, i.e. connecting parents with a common observed child, cliques can be identified. The cliques are the smallest sets of variables on which local computations can be done. Figure 8.3 shows the resulting junction tree with cliques $C_i = \{\phi, a_i, r_i, O_i, \hat{x}_i, \hat{y}_i, \hat{a}_i\}, i = 1, \ldots, n$ one for each contour and sepsets $S_j = \{\phi\}, j = 1, \ldots, n-1$ containing the common variables of the neighboring cliques.

The initial distribution for each clique will be

$$
\begin{aligned}
P_0(C_i) &= P_0(r_i, \phi, a_i, O_i, \hat{x}_i, \hat{y}_i, \hat{a}_i) = \\
&= P(r_i)P(\phi)P(a_i)P(O_i)P(\hat{x}_i, \hat{y}_i, \hat{a}_i | r_i, \phi, a_i, O_i), \tag{8.15}
\end{aligned}
$$

where $P(r_i)$, $P(\phi)$, $P(a_i)$ and $P(O_i)$ are prior distributions. $P(\hat{x}_i, \hat{y}_i, \hat{a}_i | r_i, \phi, a, O_i)$ is the distribution of the estimates. How these will be assigned will be discussed Section 8.3.2.4.

Because the junction tree here has a simple structure, the clique probability distribution after the message propagation has a closed form solution, namely

$$P(C_i) = P_0(C_i) \prod_{j \neq i} M_j(\phi). \tag{8.16}$$

where M_j is the max-margin of the initial distribution of clique j, defined as

$$M_j(\phi) = \max_{C_i \backslash \phi} P_0(C_i). \tag{8.17}$$

To find the most likely light source direction we can select an arbitrary clique and find the ϕ giving the maximum probability.

$$\hat{\phi} = \operatorname*{argmax}_{\phi} \max_{C_i \backslash \phi} P(C_i) \tag{8.18}$$

8.3.2.3   Classifying the Contours

From the clique potentials we can also estimate O_i i.e. classify whether the contour is an occluding contour or not. From the clique potential after the message propagation we classify contour i by

$$\hat{O}_i = \operatorname*{argmax}_{O_i} \max_{C_i \backslash O_i} P(C_i). \tag{8.19}$$

8.3.2.4   Priors and Distributions

The prior distributions are selected as follows.

$$
\begin{aligned}
P(\phi) &= \frac{1}{2\pi}, -\pi < \phi \leq \pi & (8.20) \\
P(r_i) &= \frac{1}{r_{max} - r_{min}}, r_{min} \leq r_i \leq r_{max} & (8.21) \\
P(a_i) &= \frac{1}{a_{max} - a_{min}}, a_{min} \leq a_i \leq a_{max} & (8.22) \\
P(O_i) &= (1 - p_{oc}, p_{oc})^T & (8.23)
\end{aligned}
$$

p_{oc} is the prior probability that a a contour is an occluding contour.

Especially the prior for r_i plays an important role, since it will help to sift out a certain class of contours that fit the shading model well but are not occluding contours. This class are contours which have a flat intensity curve. They come typically from shadows or the outsides of object boundaries on planar surfaces. Because of the flat intensity curve their estimated r_i will be close to zero and can then be easily sifted out by setting r_{min} to a value over zero.

The estimates, as derived in Section 8.2.2, are normally distributed around the true values, provided that the contours are occluding contours. The estimates from other contours are also normally distributed, for the same reason as the occluding contours, but around which mean $\vec{m}$ is unknown. The estimates certainly doesn't tell us anything about the variables we are estimating. From (8.4) it can be shown that the mean is

$$\vec{m} = E\left[(\mathbf{N}^T\mathbf{N})^{-1}\mathbf{N}^T\mathbf{I}\right] = \mu_I \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \tag{8.24}$$

where μ_I is the expected value of the image intensities. By calculating a mean over a number of images μ_I was estimated to 0.35. A suitable covariance matrix was roughly estimated to $S = ((0.5, 0, 0)^T(0, 0.5, 0)^T(0, 0, 0.5)^T)$

Hence, the distribution of the estimates is modeled by

$$P(\hat{L}|r_i, \phi, a_i, O_i) = \begin{cases} g(\hat{L}, (r_i\cos\phi, r_i\sin\phi, a_i)^T, \hat{\Sigma}_i) & O_i = oc \\ g(\hat{L}, \vec{m}, S) & O_i = \bar{o}c \end{cases} \tag{8.25}$$

The function $g(\vec{x}, \vec{\mu}, \Sigma)$ is the three-dimensional gaussian p.d.f. with mean $\vec{\mu}$ and covariance Σ.

8.3.2.5 Implementation Issues

The Bayesian network in Figure 8.2 is a hybrid network, meaning it has both continuous and discrete variables. Although we have normal distributions we can not use the developed techniques in e.g. (Cowell, Dawid, Lauritzen, and Spiegelhalter 1999). This is because the argument ϕ is a common variable in the cliques which messes up the marginalizations. Fortunately many of the marginalizations can be solved analytically. Maximizing over r_i and a_i is done by minimizing the quadratic expression in the exponential of the gaussian. With limits on r_i and a_i we need to check the boundaries of the limit region as well.

When maximizing over O_i we need to calculated the probability distributions numerically and therefore it is necessary to discretize ϕ.

8.4 Experiments

All the experiments are done on images captured on an Olympus 3030-Z digital camera. The correct light source direction was measured using the shadow of a small sphere on a piece of wire placed in the scene.

8.4.1 Light Source from a single Occluding Contour

To test the estimation of light source direction from a single occluding contour we used in total 41 contours from three different objects illuminated from six different
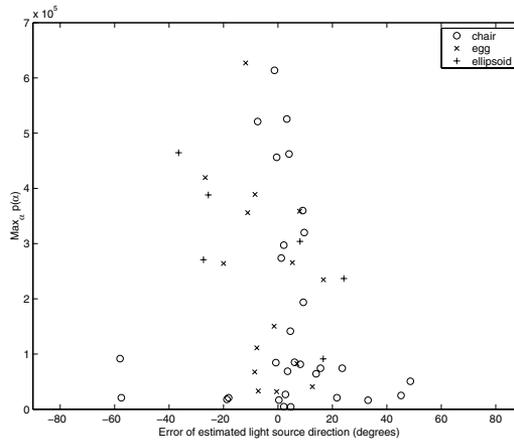
Figure 8.4:   Error of light source direction estimated from occluding contours.
plotted versus the maximum of the probability density.

directions. The contours where extracted by manually selecting parts of contours
extracted by the Canny edge detector followed by edge linking.

In Figure 8.4 the errors of the estimated light source direction is plotted versus
the maximum probability density of the estimated vector. The maximum probabil-
ity is relevant because it plays an important role in the message propagation in the
Bayesian network. Some of the estimates contain non-negligible errors, but it can
be seen that the computed variance reflect the errors, which is of great importance
when fusing them.

8.4.1.1   The Benefits of Extrapolation

What are the benefits of using extrapolation? The alternative is to use the intensity
value closest to the contour. For many objects this would be fine. It can be seen in
the shading model (8.1) that we can include contour points having a normal with
non-zero z-component, as long as the z-component is constant along the contour
for which the points are measured. Since the term $kn_z l_z$ is constant it can be fit
into the simpler model (8.2) where the term a actually is $a + kn_z l_z$.

Contours on objects which have more or less a constant cross-section radius
will therefore not benefit from the extrapolation. Instead the extrapolation will
just amplify the noise. The more complex extrapolation model that is used the
more the noise will be amplified. In our experiments using an object shaped as
an egg or an ellipsoid we will typically not benefit from the extrapolation. The
cross-section will be more or less constant along the occluding contours of these
object. A chair however have a more varied contour and should therefore benefit
from the extrapolation.
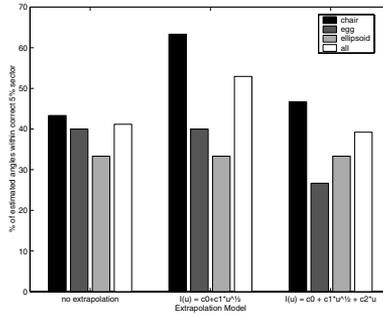
Figure 8.5: The number of estimated angles within the 5% circle sector around the correct angle for different number of terms in the extrapolation model.

This can be seen in Figure 8.5, which is a graph of the percentage of the estimated angles that are within 5% sector around the correct angle, for different extrapolation models. When using $I(u) = c_o + c_1 u^{1/2}$ as opposed to doing no extrapolation[2] we see no increase in the performance of the egg and the ellipsoid. The chair on the other hand has an increase from 43% to about 64%. As more terms are added to the extrapolation model the amplification of the noise affects the results and reduces the performance, see Figure 8.6.

8.4.2 Automatic Light Source Estimation

To test[3] the automatic algorithm two sets of images were used. The first set of images was designed to be easier by having a number of objects with occluding contours in the image. Also, the objects where placed apart so that the grouping algorithm would have less problems finding candidate contours. The set contains 14 images of two scenes illuminated from different directions. For each image the light source direction was estimated using the automatic algorithm. Also a probability measure telling how much the estimate can be trusted was calculated. This was done by summing up the probabilities for all ϕ within $\pm 5°$ of the estimated angle $\hat{\phi}$, thereby making it invariant to the number of discretization levels of ϕ. Note that the message propagation produces the whole distribution for ϕ. Figure 8.7a shows the error of the estimated angles plotted versus the probability measure. As many as half of the estimated angles have less than 5° error. There are also outliers, such as one with 33° error and a very high probability.

In the second set of images the objects sometime occlude each other. The set contains 17 images. Figure 8.7b shows the results. The algorithm is able to estimate light source direction well, with some outliers.

---

[2]When no extrapolation has been used, the image intensity is just measured two pixels from the contour.

[3]The parameters in the priors were in all experiments set to: $a_{min} = 0$, $a_{max} = 0.6$, $r_{min} = 0.1$, $r_{max} = 1$ and $p_{oc} = 0.1$.
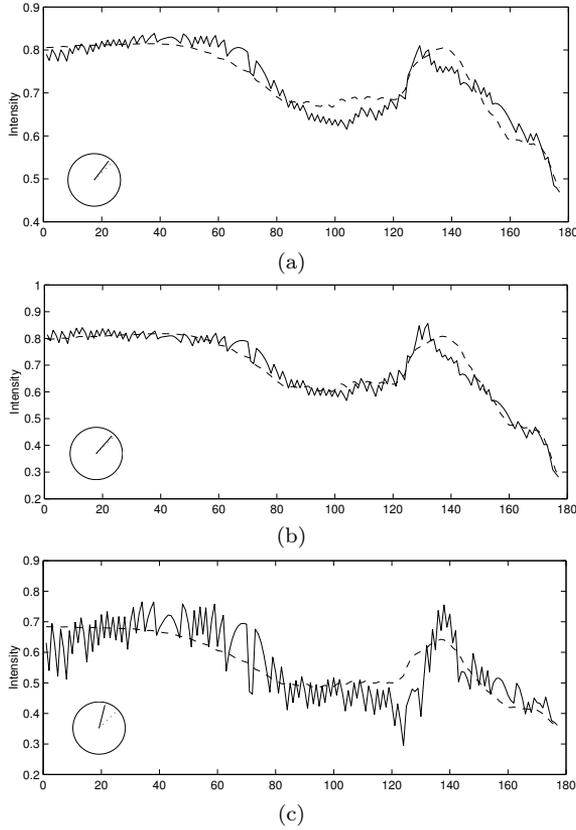
Figure 8.6:   The image intensities along one of the occluding contours in the dataset. The solid lines show the extrapolated intensities and the dashed lines the reconstructed intensities from the estimation, when using extrapolation models a) no extrapolation, b) $I(u) = c_0 + c_1 u^{1/2}$ and c) $I(u) = c_0 + c_1 u^{1/2} + c_2 u$. The inlined circles show the estimated and correct (dotted line) light source directions.

The outliers occur for different reasons. Sometimes a non-occluding contour fits the shading model well. In other cases there are many non-occluding contours accidentally giving similar estimates and thereby reinforcing each other. This mainly happens in the absence of a correct occluding contour giving a good estimate. When the grouping algorithm successfully picks out a good occluding contour, this contour will usually give a better estimate than non-occluding contours.

Note that the probabilities in this case should be compared relatively and not be considered as good estimates of the true probabilities. With better modeling, using e.g. learning, the algorithm should in the future be able to produce probabilities that better reflect the true values.

Figure 8.9 shows two examples of the algorithm estimating the light source
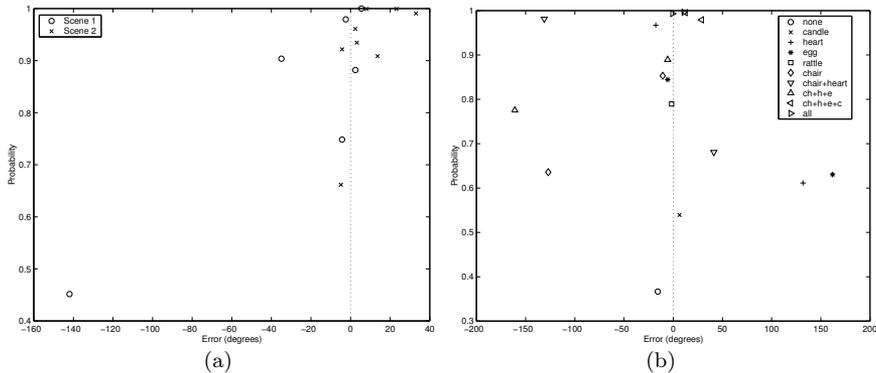
Figure 8.7: Error of estimated light source direction plotted versus probability for a) the first image set and b) the second image set.

direction and classifying the contours. In the first example the direction is correctly estimated, but some of the contours are incorrectly classified as occluding. This can happen when a contour accidentally gives the same estimate as the "winning" light source direction. In the second example the algorithm fails to estimate the correct light source direction and thereby fails to classify the contours.

8.5 Conclusion

We have presented a way to estimate the projected light source direction. The algorithm exploits the occluding boundary. Combined with a heuristic grouping algorithm and Bayesian network inference this is done in a fully automatic way. Contours picked out by the grouping algorithm are also classified as occluding or not.

Our approach to estimate the light source direction from a single occluding contour has been tested on real images with good results. The experiments show that the calculated variance reflects the errors of the estimates, which is crucial when fusing the estimates.

The automatic algorithm shows promising results. A first set of images test the inference stage on scenes suitable for the grouping algorithm. A second image set of more challenging scenes causes the grouping algorithm to fail more often. The algorithm as a whole still works with reasonable robustness.

Future improvements can be achieved in the following respects. The grouping algorithm should incorporate some sort of good continuation to better be able to cope with edges in the background. Moreover, one could model the distribution of estimates from non-occluding contours better, e.g. by learning the distributions. Another improvement would be to use the frequency space representation of shad-
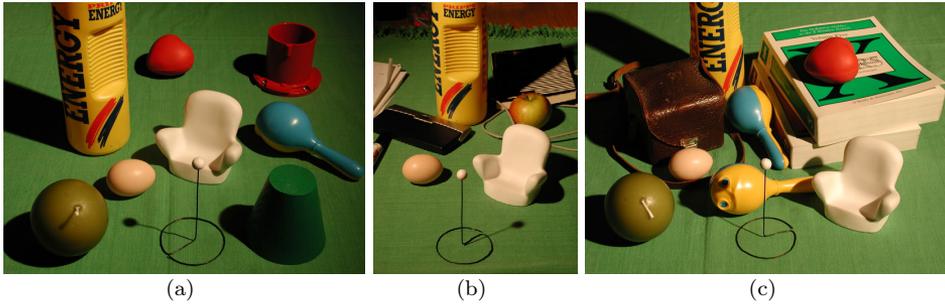
Figure 8.8: Some of the images used in the experiments.
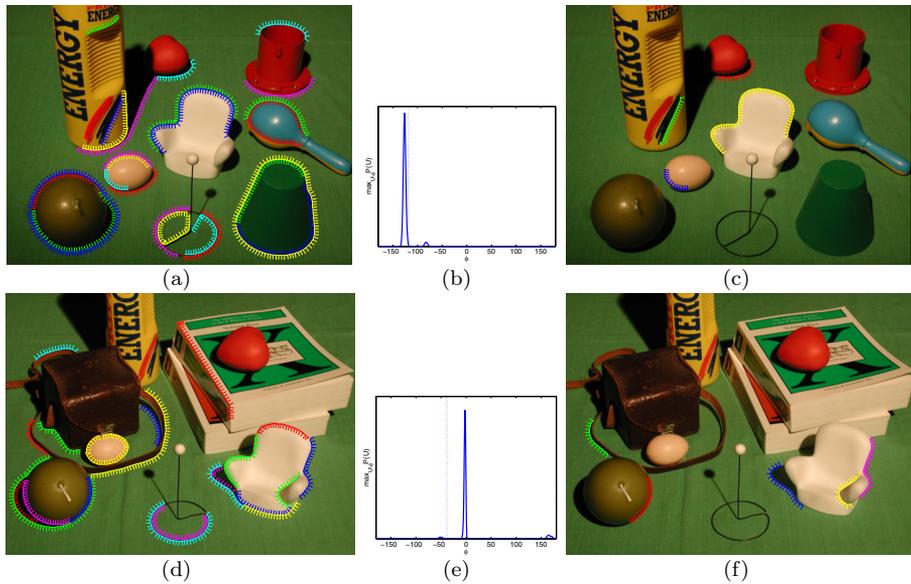
Figure 8.9: Two examples: a, d) Candidate contours from the grouping algorithm. b, e) probability distribution of estimated angle after message propagation. c, f) Contours classified as occluding. The inlined circle shows the estimated angle. Dotted lines show the correct angles.

ing. This could lead to that the algorithm can cope with arbitrary illumination and taking into account attached shadows.

Chapter 9

Discussion

9.1 Representation of Image Irradiance

Accurate models of the image irradiance with a low number of parameters are essential to the practical implementation of physics-based image analysis. To model the image irradiance one needs to model the surface reflectance functions (BRDF) and the illumination.

BRDFs form the macroscale summary of a number of properties on the surface microscale, such as surface roughness and dielectric properties. The variations are substantial and a single physical model has not been able to capture all of them. However, a general phenomenon is that many BRDFs are smooth. This makes them suitable to represent in the Fourier domain. Although these frequency space models are not directly based on the physical properties of surface scattering, they can be constructed to incorporate symmetries in the BRDF, such as Helmholtz's reciprocity and isotropy, reducing the number of parameters tremendously. One such model is created by Koenderink and van Doorn (1998).

The general smoothness of BRDFs has further implications. The light reflected by the surface is low-pass filtered. This suggest that also the illumination should be represented in the Fourier domain. The incident light to a point is a function on a sphere, hence the frequency space representation is in terms of spherical harmonics. Since the reflected light is low-pass filtered the high frequency components of the illumination do not contribute to the reflected light. Consequently, when analyzing image only the low-frequency components of the illumination need to be accounted for. Thus, any illumination can be modeled with a finite number of parameters. In the case of Lambertian surfaces, nine parameters is sufficient to model any illumination (Basri and Jacobs 2003; Ramamoorthi and Hanrahan 2001a). This is a huge step forward. Previously, the majority of the work has used one or more point light sources to model the illumination, something which is inadequate in most natural settings.

Non-Lambertian surfaces may require higher frequencies of the illumination to be modeled. In general, the shinier a surface is the more high frequency components of the illumination are needed. In this thesis we have combined the spherical harmonic representation of the illumination with the frequency space basis for isotropic surface of Koenderink and van Doorn (1998) and derived a basis for the resulting image irradiance. By analyzing the variance in the image we have shown how to compute the number of components of the illumination required to achieve a certain accuracy. Since the BRDF of any isotropic material can be represented in this basis we can compute the exact number of parameters needed to model the illumination for any isotropic material.

When analyzing the materials in the CUReT database we see that many materials, although not Lambertian, still only require up to illumination order $l = 2$ (9 parameters). These are the matte materials. The shinier materials require an illumination order of up to $l = 13$ (196 parameters) to achieve an accuracy greater that 98%.

In the frequency basis the image irradiance is represented as a sum of the product of the illumination coefficients, the coefficients of the BRDF and the image irradiance basis functions, E_{lop}^{mq}. This basis can represent the image irradiance from any isotropic material under any illumination. In the general case an infinite number of basis functions are needed. But for many materials which act as low-pass filters on the illumination, this sum can be truncated and still be an accurate representation.

The BRDF describes how a surface reflects light in all directions. In an image only a fraction of the reflected light is registered. This makes the dimensionality of the signal in the image much lower than that of the reflected light as a whole. This has been discovered empirically by performing principal component analysis (PCA) on a set images of a scene taken with varying illumination (Hallinan 1994; Epstein, Hallinan, and Yuille 1995). Only around five components are sufficient to capture most of the variations in the images. More components are needed for shiny objects. Theoretical work gives the same result. Ramamoorthi (2002) analytically derives the principal components of the images of a Lambertian object under a varying point light source. His results show that six components are enough to capture 98% of the variance.

The dimensionality of the information that can be extracted from the images is the same as the dimensionality of the images. To analyze shading robustly it is necessary to have a model that reflects that. In fact the principal component basis functions are ideal for that purpose. They are constructed by maximizing the variance of the component they represent, hence giving a high signal-to-noise ratio (SNR) when estimating the coefficients of the components.

A general problem with a PCA basis is that it is hard to relate the principal components to the properties of the illumination and the surface reflectance. Moreover, the creating of a basis for a wide variety of conditions requires a very large number of captured training images. For instance, if there are variations in both the illumination and the BRDF then an image is needed for every illumination-BRDF

combination.

A solution to this problem is *Model-Based PCA*. Instead of performing PCA on a set of images, PCA is applied directly to the image irradiance representation. Using the finite dimensional frequency space representation, the solution of the PCA is transformed from being the eigensystem of the image covariance matrix to an eigenvalue problem of the covariances of the frequency space coefficients. This separates the covariance matrix into a product of a matrix depending only on the surface shape and another matrix depending only on the variations in the illumination and the BRDF. Furthermore, the illumination-BRDF covariance matrix is separable on an element level. In effect, using this framework, it is possible to rapidly compute bases for a wide variety of conditions and different surface shapes by combining different surface, illumination and BRDF covariances.

Model-Based PCA allows the variation in the position of the light source to be performed analytically. This is done by rotating the light field representing the illumination. The relative position of the light source plays a major part in the appearance of the surface. By rotating the light source analytically we are able to take into account every possible position of the light source. In contrast, if the PCA is performed empirically the manifold of possible light source positions needs to be sampled.

Another advantage of Model-Based PCA is that there is an explicit relation between the principal components and the coefficients of the illumination and the BRDF. A PCA basis created from empirical data can capture the appearance of the images well, but it is very hard to relate the basis to the underlying properties of the data. With Model-Based PCA we are able to create a basis that takes into account the appearance ambiguities and at the same time can be used as a physics model.

Care has also been taken in forming the principal components. For image analysis it is important that the components to be estimated have a high signal-to-noise ratio. It is also important to make the basis invariant to variations in the ambient component of the illumination. These two factors suggest that the image should be centered by subtracting the individual image mean rather the mean image. Therefore we call the modification *Image-Centered PCA*.

The incorporation of all these ideas result in a powerful framework for creating low-dimensional generative models of images of a surface under a wide variety of conditions. Robust analysis of shading is achieved by estimating only the components with a high signal variance.

9.1.1   Bases for Lambertian Surfaces

The Model-Based PCA framework is demonstrated by computing bases for a spherical Lambertian object under varying illumination. Bases are computed for a number of theoretical illumination distributions such as all configurations of a point light source, a point plus ambient light source and two point light sources. Furthermore, a basis is constructed from a database of captured real world illumination.

For comparison, both standard and Image-Center PCA bases are computed. The main difference between the PCA variants is how they behave under variations in the ambient component of the illumination. When the ambient component varies this is encoded in the standard PCA basis. The Image-Centered basis on the other hand is not affected by these variations.

The changes in the Image-Centered bases for different illumination distributions are small in general. It seems that when changes in the ambient component and all rotations of the light source are accounted for, the illumination distribution has little effect on the basis.

9.1.2   A Basis for the CUReT Database

A basis is also constructed from a database of real world surface reflectances, the CUReT database (Dana et al. 1999), and the same database of captured illumination as before. The goal was to create a basis that can represent the images of a surface of a wide variety of reflectance properties and under more or less any illumination.

In this basis the images are accurately represented using between 5 and 40 basis functions depending on the surface material. Matte materials, including highly non-Lambertian materials such as velvet, are well represented using around five basis functions, while materials with specular reflectance require up to 40 basis functions.

The accuracy of the basis is measured by the percentage of variance the basis accounts for. However, this measure does not always correspond to the sensitivity of the human visual system. The reconstructed image of a matte surface can look indistinguishable from the original image at $99,1\%$ accuracy, while a specular material may require up to 99.9% accuracy to be indistinguishable from the original image. The same observations were made by Epstein et al. (1995).

9.2   Analysis of Shading

An example of analysis of images in this thesis is an algorithm for material classification. Using the model-based PCA framework we constructed a basis to represent the images of the materials in the CUReT database. The variations in the appearance of a surface of a particular material form manifolds in the low-dimensional space of this basis. These manifolds are learned from the captured illumination maps of Debevec. The material in an image is classified by finding the manifold which is closest to the image.

Recognizing the material is hard. Analysis of the recognition rates on synthetic data shows that many materials often are incorrectly classified. The materials seem to cluster into groups in which a material often is mis-classified as one of the other materials in the same group. A simple greedy grouping algorithm is able to find groups which makes intuitive sense from a human visual system standpoint. The

experiments on real images of materials are to a large extent predicted by the synthetic experiments. They indicate that the method generalizes to images of new samples of the materials, taken under varying and different illumination conditions.

9.3 Conclusions

The driving force behind the work in this thesis has been a wish understand what information that can be extracted from shading in images. The focus has been on creating methods and representations that can be used in realistic conditions with complex unknown illumination and non-Lambertian surfaces. In such conditions there are a multitude of ambiguities that need to be taken into account. These ambiguities are easily identified using the Model-Based PCA framework presented. With this framework one can construct low-dimensional models that have parameters that can be robustly estimated, take the appearance ambiguities into account and at the same time function as a physical model.

It is my hope that this framework will make it possible to use shading information in many computer vision applications. The variations in appearance due to changes in the illumination or surface reflectance are actually one of the few things that are relatively simple to model. These regularities should not be seen as a nuisance, but rather be used to our advantage.

Part III

Appendix

Appendix A

Rotating Real Spherical Harmonics

This chapter defines the real valued spherical harmonics used in the thesis. Also the rotation reparameterization functions for the real functions are derived.

A.1   Complex Spherical Harmonics

The spherical harmonics basis functions are defined as

$$y_l^m(\theta, \phi) = N_l^m P_l^m(\cos\theta)e^{im\phi}, \tag{A.1}$$

where

$$N_l^m = \sqrt{\frac{(2l+1)(l-m)!}{4\pi(l+m)!}} \tag{A.2}$$

is a normalization factor and $P_l^m(z)$ are the associated Legendre polynomials. The spherical harmonics form a complete and orthonormal basis for functions on the sphere.

$$\int_{S^2} y_l^m(\theta, \phi)\overline{y_{l'}^{m'}}(\theta, \phi)d\omega = \delta_{ll'}\delta_{mm'} \tag{A.3}$$

A rotated spherical harmonic basis function can be described as a linear combination of the spherical harmonic basis functions of the same azimuthal order m. When the arguments of the function are rotated using the ZYZ Euler angle parameterization in (2.7) then

$$y_l^m(R_{\alpha,\beta,\gamma}(\theta,\phi)) = \sum_{n=-l}^{l} d_l^{mn}(\alpha, \beta, \gamma)y_l^n(\theta, \phi), \tag{A.4}$$

where

$$d_l^{mn}(\alpha, \beta, \gamma) = P_l^{mn}(\cos\alpha)e^{im\beta+in\gamma} \tag{A.5}$$

and P_l^{mn} are the generalized associated Legendre polynomials. For their explicit form and more details on rotating spherical harmonics see (Chirikjian and Kyatkin 2001; Vilenkin and Klimyk 1991).

A.2   Real Spherical Harmonics

Real spherical harmonics can be defined as follows.

$$Y_l^m(\theta, \phi) = N_l^m P_l^m \Phi_m(\phi), \tag{A.6}$$

where

$$\Phi_m(\beta) = \begin{cases} \sqrt{2}\cos m\beta & m > 0 \\ 1 & m = 0 \\ \sqrt{2}\sin m\beta & m < 0 \end{cases} \tag{A.7}$$

The functions Φ_m are orthogonal,

$$\int_0^{2\pi} \Phi_m(\beta)\Phi_{m'}(\beta)d\beta = 2\pi\delta_{mm'}. \tag{A.8}$$

From this it is easy to see that the real spherical harmonics also are orthogonal,

$$\int_{S^2} Y_l^m(\theta, \phi)Y_{l'}^{m'}(\theta, \phi)d\omega = \delta_{ll'}\delta_{mm'}. \tag{A.9}$$

In order to convert between complex and real spherical harmonics it is convenient to express the real harmonics as a combination of complex harmonics. The complex conjugate of a complex basis functions is $\overline{y_l^m} = (-1)^m y_l^{-m}$ so

$$Y_l^m(\theta, \phi) = \begin{cases} \frac{1}{\sqrt{2}}\left(y_l^m(\theta, \phi) + (-1)^m y_l^{-m}(\theta, \phi)\right) & m > 0 \\ y_l^m(\theta, \phi) & m = 0 \\ \frac{1}{i\sqrt{2}}\left(y_l^m(\theta, \phi) - (-1)^m y_l^{-m}(\theta, \phi)\right) & m < 0 \end{cases}. \tag{A.10}$$

A.3   Conversion Between Complex and Real
Spherical Harmonics

There is a one-to-one correspondence between representations in real and complex spherical harmonics. A function $f(\theta, \phi)$ can be expressed in either representations

$$f(\theta, \phi) = \sum_{l=0}^{\infty}\sum_{m=-l}^{l} a_l^m y_l^m(\theta, \phi) = \sum_{l=0}^{\infty}\sum_{m=-l}^{l} b_l^m Y_l^m(\theta, \phi). \tag{A.11}$$

Using (A.10) a function represented in real harmonics can be written as a sum of complex basis functions.

$$\sum_{m=-l}^{l} b_l^m Y_l^m(\theta, \phi) = \sum_{m=-l}^{1} b_l^m \frac{1}{i\sqrt{2}}(y_l^m - (-1)^m y_l^{-m})$$

$$+ b_l^0 y_l^0 + \sum_{m=1}^{l} b_l^m \frac{1}{\sqrt{2}}(y_l^m + (-1)^m y_l^{-m}) \tag{A.12}$$

In the above expression the complex coefficients a_l^m can be identified.

$$a_l^m = \begin{cases} \frac{1}{\sqrt{2}}(b_l^m - \frac{1}{i}(-1)^m b_l^{-m}) & m > 0 \\ b_l^0 & m = 0 \\ \frac{1}{\sqrt{2}}(\frac{1}{i} b_l^m + (-1)^m b_l^{-m}) & m < 0 \end{cases} \tag{A.13}$$

and from this the inverse conversion from the complex to the real representation can be found.

$$b_l^m = \begin{cases} \frac{1}{\sqrt{2}}(a_l^m + (-1)^m a_l^{-m}) & m > 0 \\ a_l^0 & m = 0 \\ \frac{i}{\sqrt{2}}(a_l^m - (-1)^m a_l^{-m}) & m < 0 \end{cases} \tag{A.14}$$

A.4   Rotating Real Spherical Harmonics

Similarly to their complex counterparts the real spherical harmonics rotated can be written as a linear combinations of the non-rotated real functions.

$$Y_l^m(R_{\alpha,\beta,\gamma}(\theta, \phi)) = \sum_{n=-l}^{l} D_l^{mn}(\alpha, \beta, \gamma) Y_l^n(\theta, \phi), \tag{A.15}$$

The rotation reparameterization functions, D_l^{mn} can be found by writing the real functions as a sum of the complex functions using (A.10). These are then rotated using (A.4). The rotated complex functions are then converted back to real functions according to (A.14) and the reparameterization functions can be identified as

$$D_l^{mn}(\alpha, \beta, \gamma) = \begin{cases} \cos(m\beta + n\gamma)P_l^{mn}(\cos\alpha) \\ \quad +(-1)^n \cos(m\beta - n\gamma)P_l^{m,-n}(\cos\alpha) & m > 0 \text{ and } n > 0 \\ \sqrt{2}\cos m\beta P_l^{m0}(\cos\alpha) & m > 0 \text{ and } n = 0 \\ -\sin(m\beta + n\gamma)P_l^{mn}(\cos\alpha) \\ \quad +(-1)^n \sin(m\beta - n\gamma)P_l^{m,-n}(\cos\alpha) & m > 0 \text{ and } n < 0 \\ \\ \sqrt{2}\cos n\gamma P_l^{0n}(\cos\alpha) & m = 0 \text{ and } n > 0 \\ P_l^{00}(\cos\alpha) & m = 0 \text{ and } n = 0 \\ -\sqrt{2}\sin n\gamma P_l^{0n}(\cos\alpha) & m = 0 \text{ and } n < 0 \\ \\ \sin(m\beta + n\gamma)P_l^{mn}(\cos\alpha) \\ \quad +(-1)^n \sin(m\beta - n\gamma)P_l^{m,-n}(\cos\alpha) & m < 0 \text{ and } n > 0 \\ \sqrt{2}\sin m\beta P_l^{m0}(\cos\alpha) & m < 0 \text{ and } n = 0 \\ \cos(m\beta + n\gamma)P_l^{mn}(\cos\alpha) \\ \quad -(-1)^n \cos(m\beta - n\gamma)P_l^{m,-n}(\cos\alpha) & m < 0 \text{ and } n < 0 \end{cases}$$

$$(A.16)$$

or in a more compact form

$$D_l^{mn}(\alpha, \beta, \gamma) = \frac{\text{sgn}(n)}{2} \left[(\Phi_m(\beta)\Phi_n(\gamma) - \Phi_{-m}(\beta)\Phi_{-n}(\gamma)) P_l^{mn}(\cos\alpha) \right. \\ \left. + (\Phi_m(\beta)\Phi_n(\gamma) + \Phi_{-m}(\beta)\Phi_{-n}(\gamma)) (-1)^n P_l^{m,-n}(\cos\alpha) \right] \qquad (A.17)$$

where $\Phi_m(\beta)$ is the same as (A.7) and

$$\text{sgn}(n) = \begin{cases} 1 & n \geq 0 \\ -1 & n < 0 \end{cases}. \qquad (A.18)$$

A similar derivation of the real rotation reparameterization functions can be found in (Blanco, Florez, and Bermejo 1997).

The functions D_l^{mn} are orthogonal according to

$$\int_{SO(3)} D_l^{mn}(R) D_{l'}^{m'n'}(R) dR = \frac{\delta_{ll'}\delta_{mm'}\delta_{nn'}}{2l+1}. \qquad (A.19)$$

Note that $dR = \frac{1}{8\pi^2}\sin\alpha \, d\alpha d\beta d\gamma$.

Appendix B

Analytic Derivation of C_{lo}^q

This part of the appendix derives the analytic expression of the constants C_{lo}^q used in the image irradiance basis functions in Chapter 3.

B.1 C_{lo}^q

C_{lo}^q is the solution to the integral

$$C_{lo}^q = \int_0^{\frac{\pi}{2}} P_l^q(\cos\theta_i) R_o^q(\sqrt{2}\sin\frac{\theta_i}{2}) \cos\theta_i \sin\theta_i d\theta_i. \tag{B.1}$$

The Zernike polynomials in explicit form:

$$R_o^q(\sqrt{2}\sin\frac{\theta}{2}) = R_o^q(\sqrt{1-\cos\theta})$$

$$= \sum_{s=0}^{\frac{o-q}{2}} \frac{(o-s)!(-1)^s}{s!(\frac{o+q}{2}-s)!(\frac{o-q}{2}-s)!}(1-\cos\theta)^{\frac{o}{2}-s} \tag{B.2}$$

The associated Legendre functions in explicit form, (Mulder 2000):

$$P_l^m(\cos\theta) = \frac{(l+m)!}{2^l l!}\sin^m\theta \sum_{k=0}^{[\frac{l-m}{2}]}(-1)^{m+k}\binom{l}{k}\binom{2l-2k}{l+m}\cos^{l-m-2k}\theta \tag{B.3}$$

Inserting the explicit forms of the polynomials and doing the variable substitution $t = \cos\theta$ in the integral results in

$$C_{lo}^q = \frac{(l+q)!}{2^l l!}\sum_{s=0}^{\frac{o-q}{2}}\sum_{k=0}^{[\frac{l-q}{2}]}\frac{(o-s)!(-1)^{s+q+k}}{s!(\frac{o+q}{2}-s)!(\frac{o-q}{2}-s)!}\binom{l}{k}\binom{2l-2k}{l+q}$$

$$\times I(\frac{o+q}{2}-s, l-q-2k+1, \frac{q}{2}) \tag{B.4}$$

where

$$I(a,b,c) = \int_0^1 (1-t)^a t^b (1+t)^c dt. \tag{B.5}$$

Although (B.5) looks simple it's general solution can only be expressed in terms of the hypergeometric function, $_2F_1$. However, in all our cases both a and b are positive integers, which enables ut to find a closed form solution. First we expand $(1-t)^a$ into a binomial sum,

$$(1-t)^a = \sum_{i=0}^{a} \binom{a}{i} (-1)^i t^i. \tag{B.6}$$

This leaves us with an integral that can be solved using integration by parts.

$$
\begin{aligned}
\int_0^1 t^d (1+t)^c dt &= \left[\frac{(1+t)^{c+1}}{c+1} t^d \right]_0^1 - \int_0^1 \frac{d}{c+1} t^{d-1} (1+t)^{c+1} dt \\
&= \frac{2^{c+1}}{c+1} - \left[\frac{d(1+t)^{c+2} t^{d-1}}{(c+1)(c+2)} \right]_0^1 + \int_0^1 \frac{d(d-1)}{(c+1)(c+2)} t^{d-2} (1+t)^{c+2} dt \\
&= \frac{2^{c+1}}{c+1} - \frac{d\,2^{c+2}}{(c+1)(c+2)} + \frac{d(d-1)2^{c+3}}{(c+1)(c+2)(c+3)} \\
&\quad - \int_0^1 \frac{d(d-1)(d-2)}{(c+1)(c+2)(c+3)} t^{d-2} (1+t)^{c+3} dt \\
&= \cdots = \sum_{j=0}^{d-1} \frac{(-1)^j 2^{c+1+j} (d-j+1)_j}{(c+1)_{j+1}} + (-1)^d \int_0^1 \frac{d!}{(c+1)_d} (1+t)^{c+d} dt \\
&= \sum_{j=0}^{d-1} \frac{(-1)^j 2^{c+1+j} (d-j+1)_j}{(c+1)_{j+1}} + (-1)^d \frac{d!(2^{c+d+1}-1)}{(c+1)_{d+1}} \\
&= \sum_{j=0}^{d} \frac{(-1)^j 2^{c+1+j} (d-j+1)_j}{(c+1)_{j+1}} - \frac{(-1)^d d!}{(c+1)_{d+1}}
\end{aligned}
\tag{B.7}
$$

where

$$(a)_n = \frac{\Gamma(a+n)}{\Gamma(a)} = a(a+1)(a+2)\cdots(a+n-1) \tag{B.8}$$

Putting it all together we obtain

$$
\begin{aligned}
\int_0^1 & (1-t)^a t^b (1+t)^c dt \\
&= \sum_{i=0}^{a} \binom{a}{i} (-1)^i \left(2^{c+1} \sum_{j=0}^{b+i} \frac{(-1)^j 2^j (b+i-j+1)_j}{(c+1)_{j+1}} - \frac{(-1)^{b+i}(b+i)!}{(c+1)_{b+i+1}} \right)
\end{aligned}
\tag{B.9}
$$

Appendix C

Variance of Image Irradiance Modes

In this part of the appendix we derive the analytical expression for the average variance of the image irradiance modes, E_{lop}^{mq}.

First a theorem. I have not found this or an equivalent theorem in any text book, but it is easy to proove and necessary for the calculations in this chapter.

Theorem 1. The P_l^{mn} functions obey the following relations.

$$\sum_{n=-l}^{l} P_l^{mn}(z)P_l^{m'n}(z) = \delta_{mm'} \tag{C.1}$$

$$\sum_{m=-l}^{l} P_l^{mn}(z)P_l^{mn'}(z) = \delta_{nn'} \tag{C.2}$$

Proof. Two complex spherical harmonic basis function of the same polar are orthogonal according to

$$\int_{S^2} y_l^m(\theta,\phi)\overline{y_l^{m'}}(\theta,\phi) = \delta_{mm'}. \tag{C.3}$$

This orthogonality still holds if the two functions are rotated with the same arbitrary rotation $R = (\alpha,\beta,\gamma) \in SO(3)$. A rotated basis function can be written as a linear combination of the basis functions of the same polar order l according to (A.4). The integral over the rotated functions results in

$$\int_{S^2} \sum_{n=-l}^{l} \sum_{n'=l}^{l} d_l^{mn}(R)y_l^n(\theta,\phi)\overline{d_l^{m'n'}}(R)\overline{y_l^{n'}}(\theta,\phi)\sin\theta d\theta$$

$$= \sum_{n=-l}^{l} \sum_{n'=l}^{l} d_l^{mn}(R)\overline{d_l^{m'n'}}(R)\delta_{nn'} = \sum_{n=-l}^{l} d_l^{mn}(R)\overline{d_l^{m'n}}(R). \tag{C.4}$$

Since (C.3) still holds, we have

$$\sum_{n=-l}^{l} d_l^{mn}(R)\overline{d_l^{m'n}}(R) = \delta_{mm'}. \tag{C.5}$$

Expanding d_l^{mn} from (A.5) results in

$$\sum_{n=-l}^{l} d_l^{mn}(R)\overline{d_l^{m'n}}(R) = \sum_{n=-l}^{l} e^{(m-m')\beta+(n-n)\gamma} P_l^{mn}(\cos\alpha) P_l^{m'n}(\cos\alpha)$$

$$= e^{(m-m')\beta} \sum_{n=-l}^{l} P_l^{mn}(\cos\alpha) P_l^{m'n}(\cos\alpha). \tag{C.6}$$

Since (C.6) holds for any rotation R it also holds when $\beta = 0$. Hence,

$$\sum_{n=-l}^{l} P_l^{mn}(\cos\alpha) P_l^{m'n}(\cos\alpha) = \delta_{mm'} \tag{C.7}$$

(C.2) follows from that $P_l^{mn}(z) = (-1)^{m+n} P_l^{nm}(z)$. ∎

C.1   Variance of a single reflectance mode

In this section we derive the attenuation factors,

$$V_{B,I_{op}^q}(l) = \sum_m \int_{\mathbf{H_{S^2}}} (E_{lop}^{mq}(\alpha,\beta))^2 d\omega \tag{C.8}$$

for a single material mode I_{op}^q. E_{lop}^{mq} can be separated into a product of a polar function $\Theta_{lop}^{mq}(\alpha)$ and an azimuthal function $\Phi_m(\beta)$ as described in Section 3.3.1. The functions $\Phi_m(\beta)$ are orthogonal (A.8) so solving the azimuthal integral in (C.8) results in

$$V_{B,I_{op}^q}(l) = \sum_m \int_{\mathbf{H_{S^2}}} (E_{lop}^{mq}(\alpha,\beta))^2 d\omega = 2\pi \sum_m \int_0^{\pi/2} (\Theta_{lop}^{mq}(\alpha))^2 sin\alpha d\alpha$$

$$= 2\pi \sum_m \int_0^{\pi/2} (N_l^q)^2 \frac{(o+1)(p+1)}{4(1+\delta_{op})(1+\delta_{q0})} \left((-1)^q P_l^{mq}(\cos\alpha) + P_l^{m,-q}(\cos\alpha)\right)^2$$

$$\times \left(C_{lo}^q R_p^q(\sqrt{2}\sin\frac{\alpha}{2}) + C_{lp}^q R_o^q(\sqrt{2}\sin\frac{\alpha}{2})\right)^2 \sin\alpha d\alpha \tag{C.9}$$

If the sum in the above expression is moved into the integral it will eliminate the P_l^{mq} according to Theorem 1.

$$\sum_{m=-l}^{l} ((-1)^q P_l^{mq}(z) + P_l^{m,-q}(z))^2 = 1 + 2(-1)^q \delta_{q0} + 1 \tag{C.10}$$

$$= 2(1 + \delta_{q0}).$$

Furthermore, the Zernike polynomials are orthogonal (Koenderink and van Doorn 1998) according to

$$\int_0^{\pi/2} R_o^q(\sqrt{2}\sin\frac{\alpha}{2}) R_p^q(\sqrt{2}\sin\frac{\alpha}{2}) \sin\alpha \, d\alpha = \frac{\delta_{op}}{o+1} \tag{C.11}$$

so the integral

$$\int_0^{\pi/2} \left(C_{lo}^q R_p^q(\sqrt{2}\sin\frac{\alpha}{2}) + C_{lp}^q R_o^q(\sqrt{2}\sin\frac{\alpha}{2}) \right)^2 \sin\alpha \, d\alpha$$

$$= \left(\frac{(C_{lo}^q)^2}{p+1} + \frac{(C_{lp}^q)^2}{o+1} \right) (1 + \delta_{op}). \tag{C.12}$$

Inserting the results (C.10) and (C.12) in (C.9) results in

$$V_{B,I_{op}^q}(l) = \frac{2\pi(N_l^q)^2(o+1)(p+1)}{4(1+\delta_{op})(1+\delta_{q0})} 2(1+\delta_{q0}) \left(\frac{(C_{lo}^q)^2}{p+1} + \frac{(C_{lp}^q)^2}{o+1} \right)(1+\delta_{op})$$

$$= \pi(N_l^q)^2 \left((o+1)(C_{lo}^q)^2 + (p+1)(C_{lp}^q)^2 \right) \tag{C.13}$$

C.2 Variance of a composite BRDF

For a BRDF represented as a linear combination of the reflectance modes as in (3.4) the attenuation factors $V_B(l)$ are calculated in the same way as for the single mode.

We now derive the expression for the attenuation factors

$$V_B(l) = \sum_{\substack{opq \\ o'p'q'}} b_{op}^q b_{o'p'}^{q'} \sum_m \int_{\mathbf{H_{S^2}}} E_{lop}^{mq}(\alpha,\beta) E_{lo'p'}^{mq'}(\alpha,\beta) d\omega \tag{C.14}$$

Solving the azimuthal integral results in

$$V_B(l) = \sum_{\substack{opq \\ o'p'q'}} b_{op}^q b_{o'p'}^{q'} \sum_m 2\pi \int_0^{\pi/2} \Theta_{lop}^{mq}(\alpha) \Theta_{lo'p'}^{mq'}(\alpha) \sin\alpha \, d\alpha. \tag{C.15}$$

The sum over m is moved into the integral. This sum expands to

$$\sum_{m=-l}^{l} \Theta_{lop}^{mq}(\alpha)\Theta_{lo'p'}^{mq'}(\alpha)$$

$$= \sum_{m=-l}^{l} N_l^q N_l^{q'} \sqrt{\frac{(o+1)(p+1)(o'+1)(p'+1)}{16(1+\delta_{op})(1+\delta_{q0})(1+\delta_{o'p'})(1+\delta_{q'0})}}$$

$$\times \left((-1)^q P_l^{mq}(\cos\alpha) + P_l^{m,-q}(\cos\alpha)\right)\left((-1)^{q'} P_l^{mq'}(\cos\alpha) + P_l^{m,-q'}(\cos\alpha)\right)$$

$$\times \left(C_{lo}^q R_p^q(\sqrt{2}\sin\frac{\alpha}{2}) + C_{lp}^q R_o^q(\sqrt{2}\sin\frac{\alpha}{2})\right)$$

$$\times \left(C_{lo'}^{q'} R_{p'}^{q'}(\sqrt{2}\sin\frac{\alpha}{2}) + C_{lp'}^{q'} R_{o'}^{q'}(\sqrt{2}\sin\frac{\alpha}{2})\right) \tag{C.16}$$

Due to (C.2) this sum is greatly simplifyed, since

$$\sum_{m=-l}^{l} ((-1)^q P_l^{mq}(z) + P_l^{m,-q}(z))((-1)^{q'} P_l^{mq'}(z) + P_l^{m,-q'}(z))$$

$$= (-1)^{q+q'}\delta_{qq'} + (-1)^q \delta_{qq'}\delta_{q0} + (-1)^{q'}\delta_{qq'}\delta_{q0} + \delta_{qq'}$$

$$= \delta_{qq'}2(1+\delta_{q0}). \tag{C.17}$$

Inserting (C.17) in (C.16) results in

$$\sum_{m=-l}^{l} \Theta_{lop}^{mq}(\alpha)\Theta_{lo'p'}^{mq'}(\alpha)$$

$$= (N_l^q)^2 \sqrt{\frac{(o+1)(p+1)(o'+1)(p'+1)}{4(1+\delta_{op})(1+\delta_{o'p'})}} \tag{C.18}$$

$$\times \left(C_{lo}^q R_p^q(\sqrt{2}\sin\frac{\alpha}{2}) + C_{lp}^q R_o^q(\sqrt{2}\sin\frac{\alpha}{2})\right)$$

$$\times \left(C_{lo'}^q R_{p'}^q(\sqrt{2}\sin\frac{\alpha}{2}) + C_{lp'}^q R_{o'}^q(\sqrt{2}\sin\frac{\alpha}{2})\right)\delta_{qq'}$$

Due to the orthogonality of the Zernike polynomials

$$\int_0^{\pi/2} \left(C_{lo}^q R_p^q(\sqrt{2}\sin\frac{\alpha}{2}) + C_{lp}^q R_o^q(\sqrt{2}\sin\frac{\alpha}{2})\right)$$

$$\times \left(C_{lo'}^q R_{p'}^q(\sqrt{2}\sin\frac{\alpha}{2}) + C_{lp'}^q R_{o'}^q(\sqrt{2}\sin\frac{\alpha}{2})\right)\sin\alpha\, d\alpha \tag{C.19}$$

$$= \frac{C_{lo}^q C_{lo'}^q}{p+1}\delta_{pp'} + \frac{C_{lo}^q C_{lp'}^q}{p+1}\delta_{po'} + \frac{C_{lp}^q C_{lo'}^q}{o+1}\delta_{op'} + \frac{C_{lp}^q C_{lp'}^q}{o+1}\delta_{oo'}.$$

Finally, combining (C.15) (C.18) and (C.19) results in

$$
V_B(l) = \sum_{\substack{opq \\ o'p'q'}} b_{op}^q b_{o'p'}^{q'} \pi (N_l^q)^2 \sqrt{\frac{(o+1)(p+1)(o'+1)(p'+1)}{(1+\delta_{op})(1+\delta_{o'p'})}}
$$
$$
\times \left(\frac{C_{lo}^q C_{lo'}^q}{p+1}\delta_{pp'} + \frac{C_{lo}^q C_{lp'}^q}{p+1}\delta_{po'} + \frac{C_{lp}^q C_{lo'}^q}{o+1}\delta_{op'} + \frac{C_{lp}^q C_{lp'}^q}{o+1}\delta_{oo'} \right) \delta_{qq'}
$$

$(C.20)$

$V_B(l)$ for all the materials in the CUReT database are plotted in Figures C.1-C.6.
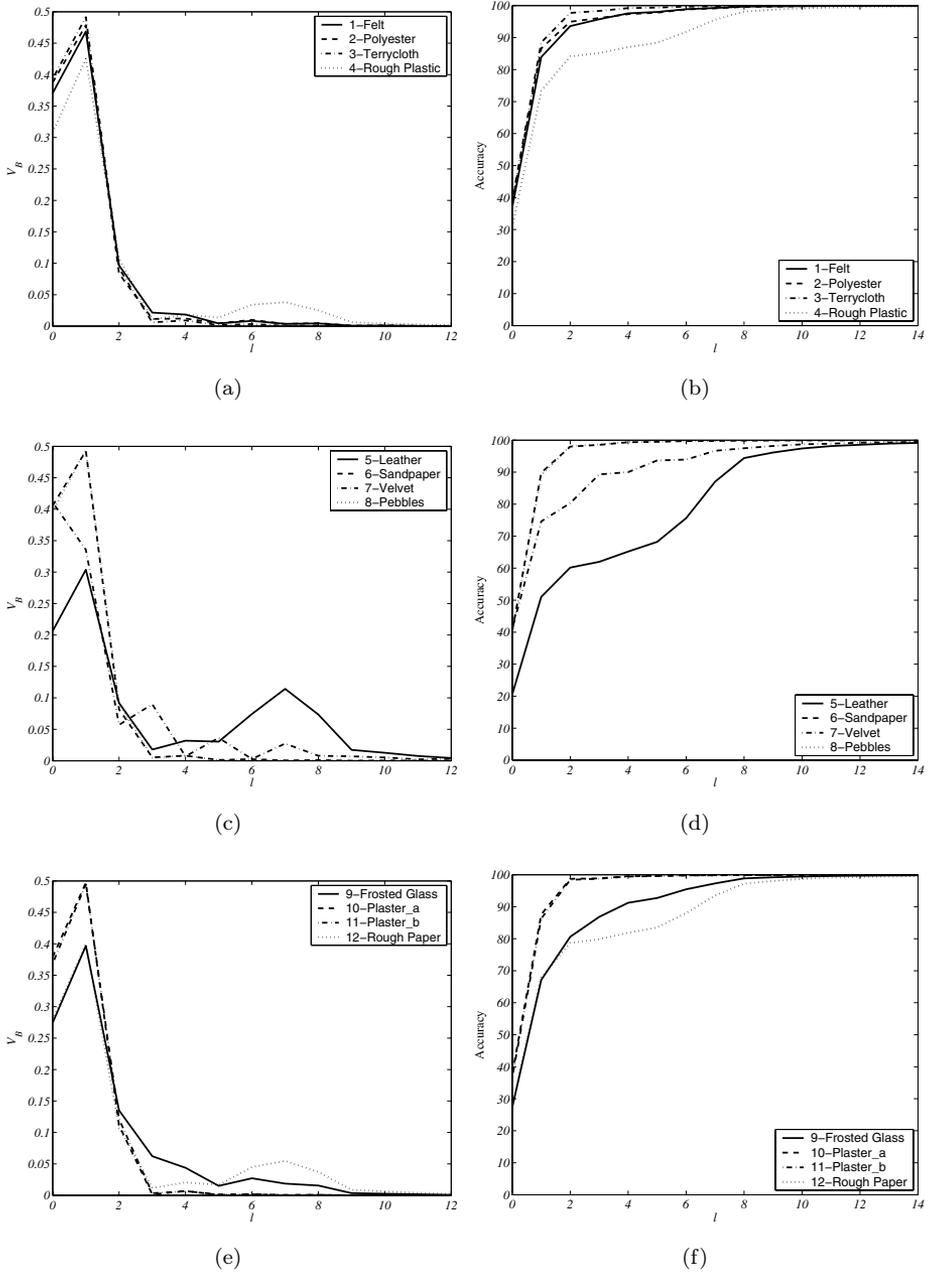
Figure C.1: Left column: Attenuation factors on each iilumination order l for the materials in the CUReT database. Right column: Accuracy of image irradiance representation when using illumination orders up to l.
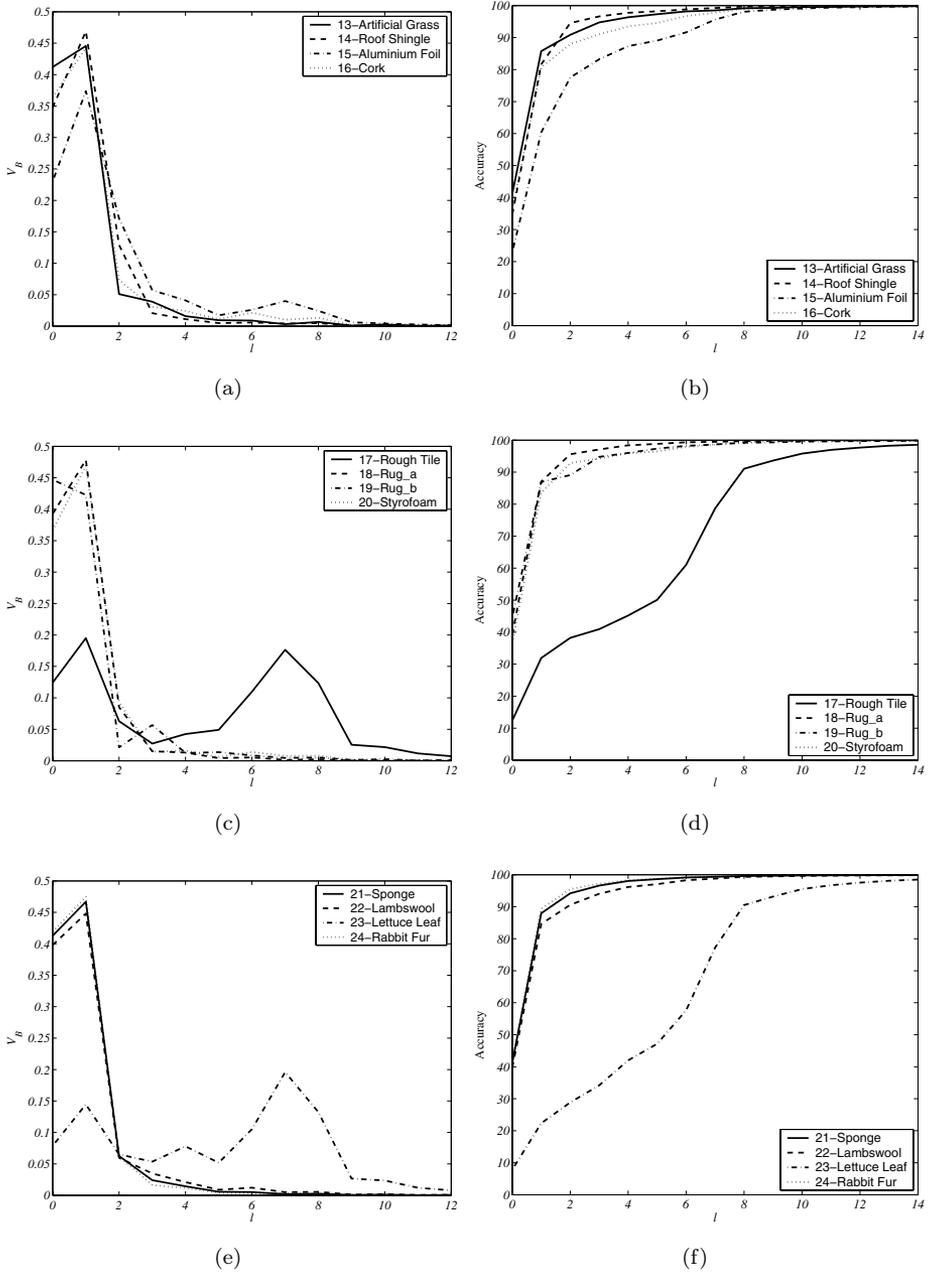
Figure C.2: Left column: Attenuation factors on each iilumination order l for the materials in the CUReT database. Right column: Accuracy of image irradiance representation when using illumination orders up to l.
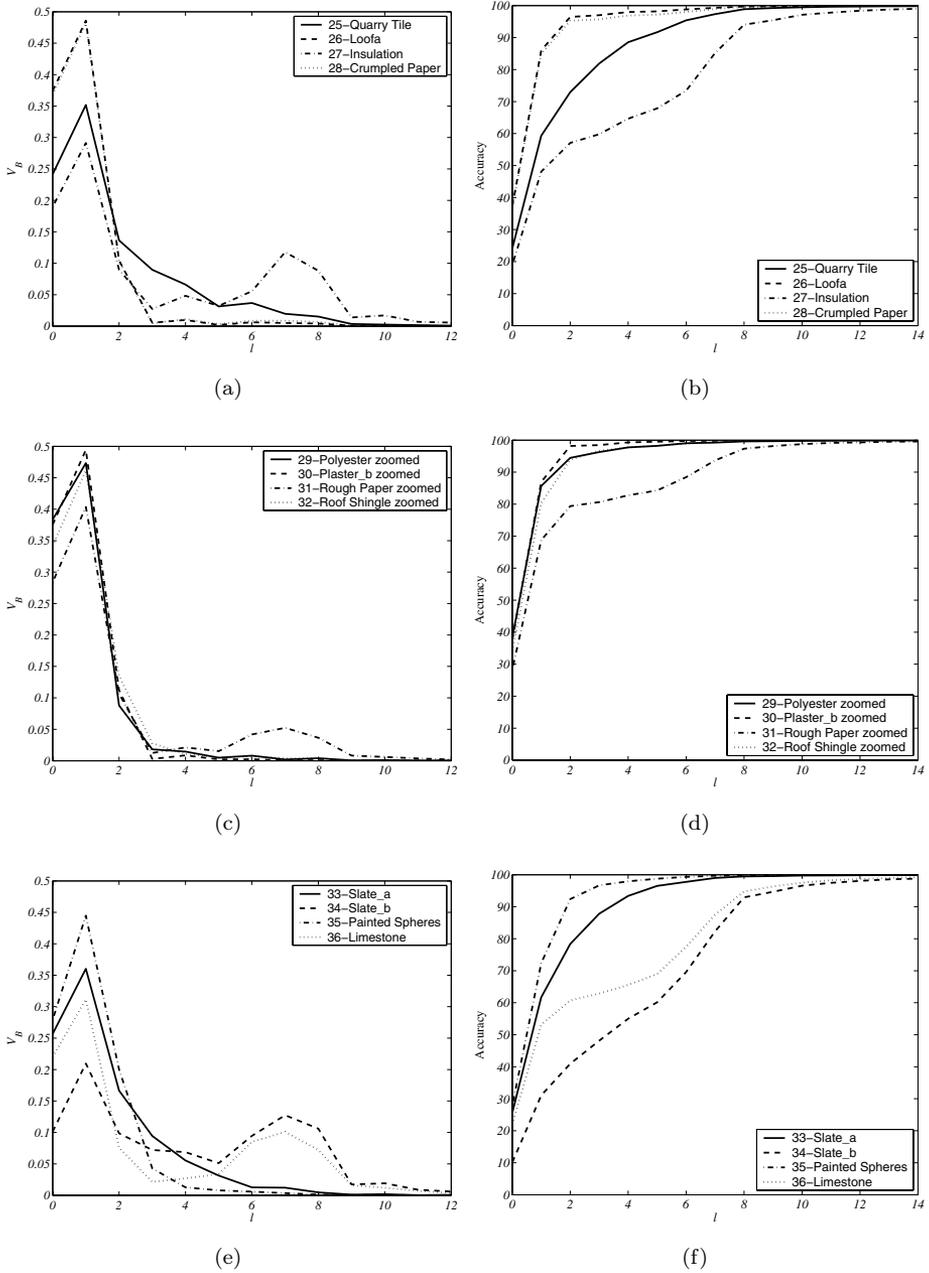
Figure C.3: Left column: Attenuation factors on each iilumination order l for the materials in the CUReT database. Right column: Accuracy of image irradiance representation when using illumination orders up to l.
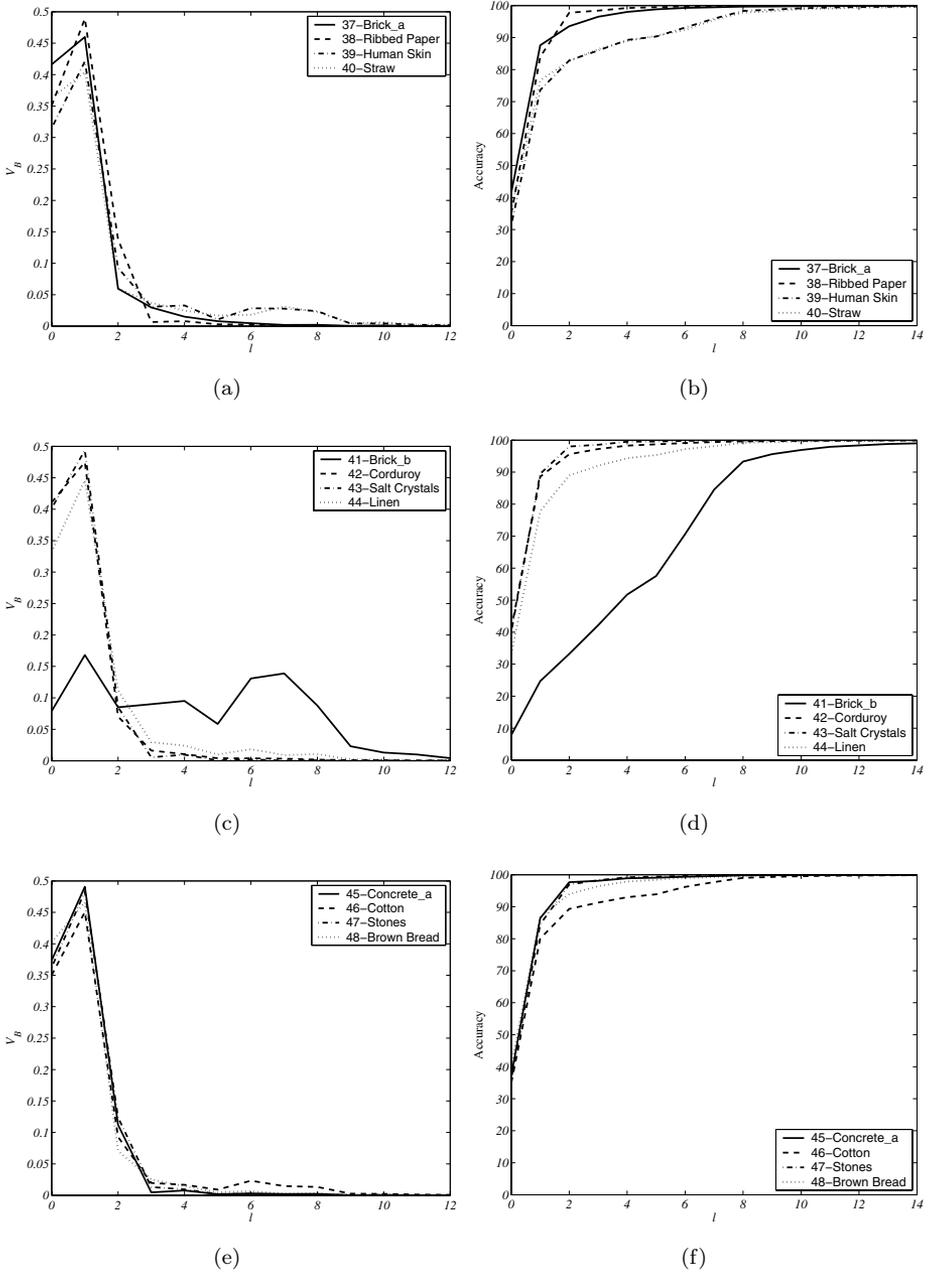
Figure C.4: Left column: Attenuation factors on each iilumination order l for the materials in the CUReT database. Right column: Accuracy of image irradiance representation when using illumination orders up to l.
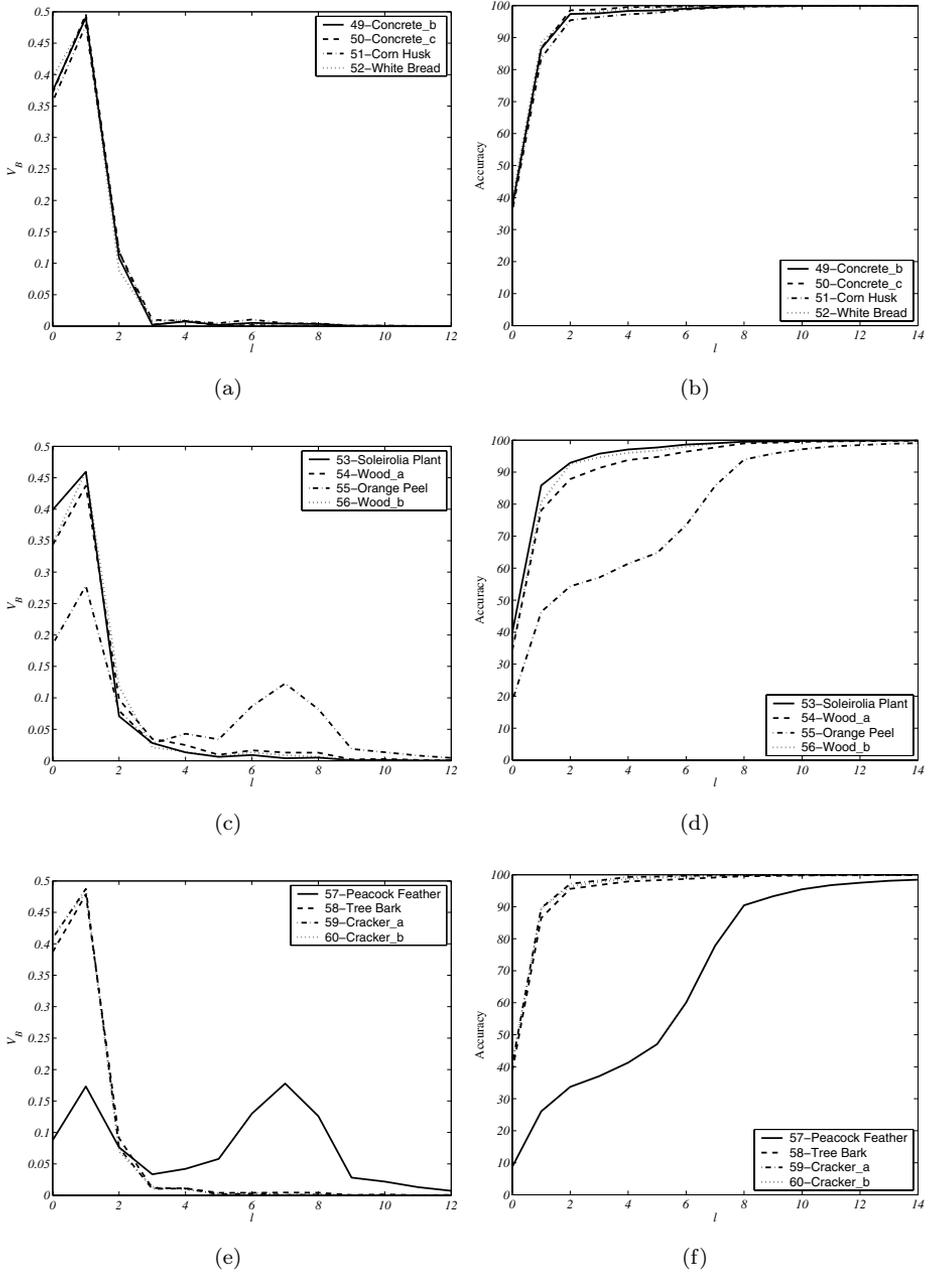
Figure C.5: Left column: Attenuation factors on each iilumination order l for the materials in the CUReT database. Right column: Accuracy of image irradiance representation when using illumination orders up to l.
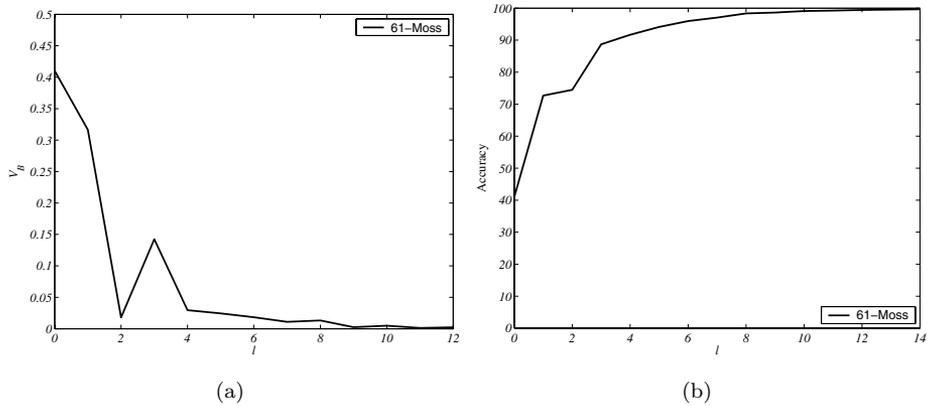
Figure C.6: Left column: Attenuation factors on each iilumination order l for the materials in the CUReT database. Right column: Accuracy of image irradiance representation when using illumination orders up to l.

References

Ballard, D. and C. Brown (1982). *C*omputer Vision. Prentice-Hall.

Basri, R. and D. Jacobs (2001a). Lambertian reflectance and linear subspaces. In *I*nternational Conference on Computer Vision, pp. II: 383–390.

Basri, R. and D. Jacobs (2001b). Photometric stereo with general, unknown lighting. In *I*EEE Computer Vision and Pattern Recognition, pp. II:374–381.

Basri, R. and D. Jacobs (2003, February). Lambertian reflectance and linear subspaces. *I*EEE Transactions on Pattern Analysis and Machine Intelligence *2*5(2), 218–233.

Belhumeur, P., D. Kriegman, and A. Yuille (1997). The bas-relief ambiguity. In *I*EEE Computer Vision and Pattern Recognition, pp. 1060–1066.

Belhumeur, P. N. and D. J. Kriegman (1998). What is the set of images of an object under all possible illumination conditions? *I*nternational Journal of Computer Vision *2*8(3), 245–260.

Blanco, M. A., M. Florez, and M. Bermejo (1997). Evaluation of the rotation matrices in the basis of real spherical harmonics. *J*ournal of Molecular Structure: Theochem *4*19, 19–27.

Cabral, B., N. Max, and R. Springmeyer (1987). Bidirectional reflection functions from surface bump maps. *A*CM Computer Graphics *2*1(4), 273–281.

Chirikjian, G. S. and A. B. Kyatkin (2001). *E*ngineering Applications of Noncommutative Harmonic Analysis. CRC Press.

Cook, R. L. and K. E. Torrance (1981). A reflectance model for computer graphics. In *S*IGGraph, pp. 307–316. ACM Press.

Cowell, R., A. Dawid, S. Lauritzen, and D. Spiegelhalter (1999). *P*robabilistic Networks and Expert Systems. Springer.

Craig, J. J. (1989). *I*ntroduction to Robotics. Addison-Wesley.

Dana, K., B. van Ginneken, S. Nayar, and J. Koenderink (1996, December). Reflectance and texture of real-world surfaces. Technical report, Columbia University.

Dana, K., B. van Ginneken, S. Nayar, and J. Koenderink (1999, January). Reflectance and texture of real-world surfaces. *A*CM Transactions on Graphics *1*8(1), 1–34.

Debevec, P. (1998). Rendering synthetic objects into real scenes: Bridging traditional and image-based graphics with global illumination and high dynamic range photography. In *S*IGGraph, pp. 189–198.

Dror, R., E. Adelson, and A. Willsky (2001). Recognition of surface reflectance properties from a single image under unknown real-world illumination. In *W*orkshop of recognizing objects under varying illumination.

D'Zmura, M. (1991). Shading ambiguity: Reflectance and illumination. In M. S. Landy and J. A. Movshon (Eds.), *C*omputational Models of Visual Processing. The MIT Press.

Epstein, R., P. Hallinan, and A. Yuille (1995). 5+/-2 eigenimages suffice: An empirical investigation of low-dimensional lighting models. In *I*EEE Workshop Physics-Based Modeling in Computer Vision, pp. 108–116.

Hallinan, P. (1994). A low-dimensional representation of human faces for arbitrary lighting conditions. In *I*EEE Computer Vision and Pattern Recognition, pp. 995–999.

Horn, B. (1970). *S*hape from Shading: A Method for Obtaining the Shape of a Smooth Opaque Object from One View. Ph. D. thesis, MIT.

Horn, B. (1975). Obtaining shape from shading information. In B. Horn and M. Brooks (Eds.), *S*hape from Shading, pp. 123–173. MIT Press.

Horn, B. (1977, April). Image intensity understanding. *A*rtificial Intelligence *8*(2), 201–231.

Horn, B. (1986). *R*obot Vision. McGraw-Hill.

Horn, B. and M. Brooks (1989). *S*hape from Shading. MIT Press.

Horn, B. K. P. and R. W. Sjoberg (1979, June). Calculating the reflectance map. *A*pplied Optics *1*8(11), 1770–1779.

Jensen, F. (1996). *A*n Introduction to Bayesian Networks. UCL Limited.

Johnson, R. A. and D. W. Wichern (1998). *A*pplied Multivariate Statistical Analysis, 4:th ed. Prentice-Hall.

Klinker, G., S. Shafer, and T. Kanade (1990, January). A physical approach to color image understanding. *I*nternational Journal of Computer Vision *4*(1), 7–38.

Koenderink, J. and A. van Doorn (1998, November). Phenomenological description of bidirectional surface reflection. *J*ournal of the Optical Society of America *1*5(11), 2903–2912.

Lu, R., J. Koenderink, and A. Kappers (1998, September). Optical properties (bidirectional reflection distribution functions) of velvet. *Applied Optics 3*7(25), 5974–5984.

Minnaert, M. (1941). The reciprocity principle in lunar photometry. *T*he Astrophysical Journal *9*3, 403–410.

Mitsunaga, T. and S. Nayar (1999). Radiometric self calibration. In *I*EEE Computer Vision and Pattern Recognition, pp. I: 374–380.

Mulder, J. J. C. (2000, February). Closed-form spherical harmonics: Explicit polynomial expression for the associated legendre functions. *J*ournal of Chemical Education *7*7(2), 244.

Murase, H. and S. Nayar (1995, January). Visual learning and recognition of 3-d objects from appearance. *I*nternational Journal of Computer Vision *1*4(1), 5–24.

Nicodemus, F., J. Richmond, J. Hsia, I. Ginsberg, and T. Limperis (1977, October). Geometrical considerations and nomenclature for reflectance. NBS Monograph 160. U.S. Departmentent of Commerce, National Bureau of Standards.

Nillius, P. and J. Eklundh (2001). Automatic estimation of the projected light source direction. In *I*EEE Computer Vision and Pattern Recognition, pp. I:1076–1083.

Nillius, P. and J. Eklundh (2003a). Low-dimensional representations of shaded surfaces under varying illumination. In *I*EEE Computer Vision and Pattern Recognition, pp. II:185–192.

Nillius, P. and J. Eklundh (2003b). Phenomenological eigenfunctions for image irradiance. In *I*nternational Conference on Computer Vision, pp. 568–575.

Nillius, P. and J. Eklundh (2004). Classifying materials from their reflectance properties. In *E*uropean Conference on Computer Vision.

Noy-Meir, I., D. Walker, and W. T. Williams (1975). Data transformation in ecological ordination. ii. on the meaning of data standardization. *J*ournal of Ecology (63), 779–800.

Oren, M. and S. Nayar (1995, April). Generalization of the lambertian model and implications for machine vision. *I*nternational Journal of Computer Vision *1*4(3), 227–251.

Osadchy, M., D. Jacobs, and R. Ramamoorthi (2003). Using specularities for recognition. In *I*nternational Conference on Computer Vision, pp. 1512–1519.

Pentland, A. (1982). Finding the illuminant direction. *J*ournal of the Optical Society of America *7*2, 448–455.

Ramamoorthi, R. (2002, October). Analytic pca construction for theoretical analysis of lighting variability in images of a lambertian object. *I*EEE Transactions on Pattern Analysis and Machine Intelligence *2*4(10), 1322–1333.

Ramamoorthi, R. and P. Hanrahan (2001a, October). On the relationship between radiance and irradiance: determining the illumination from images of a convex lambertian object. *J*ournal of the Optical Society of America *1*8(10), 2448–2458.

Ramamoorthi, R. and P. Hanrahan (2001b). A signal-processing framework for inverse rendering. In *S*IGGraph.

Ramamoorthi, R. and P. Hanrahan (2002). Frequency space environment map rendering. In *S*IGGraph.

Sato, I., Y. Sato, and K. Ikeuchi (1999a). Illumination distribution from brightness in shadows: Adaptive estimation of illumination distribution with unknown reflectance properties in shadow regions. In *I*nternational Conference on Computer Vision, pp. 875–82.

Sato, I., Y. Sato, and K. Ikeuchi (1999b). Illumination distribution from shadows. In *I*EEE Computer Vision and Pattern Recognition, pp. 306–12.

Simakov, D., D. Frolova, and R. Basri (2003). Dense shape reconstruction of a moving object under arbitrary, unknown lighting. In *I*nternational Conference on Computer Vision, pp. 1202–1207.

Torrance, K. and E. Sparrow (1967). Theory for off-specular reflection from roughened surfaces. *J*ournal of the Optical Society of America *5*7, 1105–1114.

Vega, E. V. and Y.-H. Yang (1994). Default shape theory: With the application to the computation of the direction of the light source. *J*ournal of the Optical Society of America *6*0, 285–299.

Vilenkin, N. J. and A. U. Klimyk (1991). *R*epresentation of Lie Groups and Special Functions, Volume 1-3. Kluwer Academic Publishers.

Westin, S., J. Arvo, and K. Torrance (1992, July). Predicting reflectance functions from complex surfaces. *C*omputer Graphics *2*6(2), 255–264.

Wolff, L. (1994, November). Diffuse-reflectance model for smooth dielectric surfaces. *J*ournal of the Optical Society of America *1*1(11), 2956–2968.

Wolff, L. B., S. A. Shafer, and G. E. Healey (Eds.) (1992). *P*hysics-Based Vision: Principles and Practice: Shape Recovery. Jones and Bartlett Publishers.

Yang, Y. and A. Yuille (1991). Sources from shading. In *I*EEE Computer Vision and Pattern Recognition, pp. 534–539.

Zhang, L. and D. Samaras (2003). Face recognition under variable lighting using harmonic image exemplars. In *I*EEE Computer Vision and Pattern Recognition, pp. I: 19–25.

Zhang, R., P.-S. Tsai, J. Cryer, and M. Shah (1999, aug). Shape from shading: A survey. *I*EEE Transactions on Pattern Analysis and Machine Intelligence *2*1(8), 690–706.

Zhang, Y. and Y.-H. Yang (2000). Illumination direction determination for multiple light sources. In *I*EEE Computer Vision and Pattern Recognition, pp. 269–76.