

ABSTRACT

Title of thesis: ILLUMINATION RECOVERY FROM IMAGES
WITH CAST SHADOWS

Xue Mei
Master of Science, 2007

Thesis directed by: Professor Gang Qu
Department of Electrical and Computer Engineering
and
Professor David Jacobs
Department of Computer Science

The effects produced in an image by cast shadows can be quite complex, especially when light comes from all directions. This makes it difficult to recover the illumination from a scene and recognize objects from the images. In this paper, we show that such images can be well approximated using much simpler lighting represented by a combination of low frequency spherical harmonics, and a small number of directional sources. Therefore, the illumination of the scene can be recovered by summing the spherical harmonic lighting and a small number of directional light sources. To demonstrate the effectiveness of the proposed method, we have successfully tested it by using sets of synthesized images rendered by directional light sources or environment maps with different objects.

ILLUMINATION RECOVERY FROM IMAGES WITH CAST
SHADOWS

by

Xue Mei

Thesis submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Master of Science
2007

Advisory Committee:
Professor Gang Qu, Chair/Advisor
Professor David Jacobs, Co-Chair/Advisor
Professor Adrian Papamarcou

© Copyright by
Xue Mei
2007

Dedication

This thesis is dedicated to my beloved parents, wife and daughter.

Acknowledgements

I would like to thank many people who helped me to bring this work together. First, to my advisors, Prof. Gang Qu and Prof. David Jacobs, whose guidance and support over the last two years have been of great help. To Prof. Adrian Papamarcou, for his valuable assistance and discussions.

I would also like to thank Sameer Shirdhonkar for providing me with the SDP code.

Table of Contents

List of Tables	v
List of Figures	vi
1 Introduction	1
2 Spherical Harmonic Analysis of Lambertian Objects with shadows	5
2.1 What are the Shadows?	5
2.2 Lambertian Model	8
2.3 Spherical Harmonic Analysis	10
3 Analysis of Images with Cast Shadows	16
3.1 An Example	16
3.2 Analysis	18
3.3 Several Algorithms	23
3.3.1 Indirect Algorithms	23
3.3.2 Direct Algorithms	26
3.4 Multiscale Search Algorithm	28
3.5 Comments on Multiscale Search Algorithm	28
4 Experiments	31
4.1 Testing the Validity of Assumptions	31
4.2 Illumination Recovery	37
4.3 Performance Evaluation	37
5 Conclusions and Future Work	43
Bibliography	44

List of Tables

4.1	Best directional source images at different image size. The scenes are from sample images.	35
-----	--	----

List of Figures

2.1	Umbra and penumbra generation. Area light sources, as the line source in (a), generate penumbra where the light is only partially obstructed by the shadow casting object. The umbra and penumbra structure is clearly visible in (b).	6
2.2	Examples of attached and cast shadow. The generation of attached and cast shadow is illustrated in (a). (b) gives an example of attached and cast shadows in real life. The left side of the cat condo, which faces away from the light source, is in the attached shadow. The cat condo casts a hard shadow on the background.	7
2.3	Shadow provide information about the relative positions of objects. We cannot determine the position of the wooden dog from shadowless image (a), whereas on the other three images we understand that it is more and more distant from the ground.	8
2.4	Shadow provide information about the shape and geometry of the occluder.	9
2.5	Lambertian surface and law. In (a), the light falling on the Lambertian surface is reflected equally in all directions and appears the same to the viewer from any viewing direction. According to Lambertian law, the intensity of a point is the dot product of surface normal and lighting direction as in (b).	10
2.6	The first nine spherical harmonics.	12
2.7	A graph representation of the first 9 coefficients and the cumulative energy of the Lambertian kernel.	14
2.8	Images of two boxes. The only difference is the thickness. They can only be differentiated by their shadows cast on the ground.	15
3.1	Images of a flat playground with a thin flag pole. (a) is rendered with only single directional source, while (b) is rendered with 2 directional sources. (b) has two shadows, but each shadow only has half the intensity of the shadow in (a).	17
3.2	Best directional source is selected by comparing it with its 5 nearest neighbors on each level.	29

4.1	Sample images of two scenes, each generated with a single, directional source.	32
4.2	The angle between the two point source images vs. the angle between the two point light sources. The scenes are from sample images. . . .	36
4.3	Each blue star points shows the error that occurs in approximating a different image with linear subspace and a set of directional sources. The red and magenta dashed line shows our prediction and prediction compensated with sampling error about the maximum possible error in this approximation, respectively.	38
4.4	The improvement of the accuracy by adding directional sources. From the plot in the figure, we can clearly see that the accuracy improves gradually as we increase the number of the directional sources. The image is rendered with 8 directional sources.	39
4.5	Comparison of different methods to approximate the image with cast shadows. The scene is rendered with 8 directional sources.	41
4.6	Comparison of different methods to approximate the image with cast shadows. The scene is rendered with an environment map.	42

Chapter 1

Introduction

When a scene is illuminated from all directions, shadows abound. Each light direction produces a different set of cast and attached shadows. These shadows create rich and complex images. In this paper, we consider shadows cast by environment maps in which lighting intensity varies as a function of direction, but does not vary spatially. For Lambertian objects we show that the shadows cast by environment maps are not as complex as they might seem, if we measure their complexity in terms of the number of parameters needed to capture the lighting that produces them. In particular, we show that shadows cast by environment maps can be approximated using a combination of diffuse lighting, described with low order spherical harmonics, and a small number of directional lights.

There has been a series of work aimed at understanding the complexity of the set of images produced by Lambertian objects lit by environment maps. [1, 2] show that when we ignore all shadows, the images of a Lambertian scene lie in a three-dimensional linear subspace in the space of all images. [3] make use of this result in rendering. [4] consider attached shadows, which occur when a surface faces away from a light. They show that with attached shadows, the set of images produced by a Lambertian scene forms a convex cone that has non-zero volume in the space of all images. However, they, and also [5], provide empirical evidence that

this space is well-approximated by a low-dimensional linear subspace. [6, 7] explain this analytically by showing that Lambertian reflectance acts as low-pass filter on the environment map. This implies that the set of images produced by a convex Lambertian object is well approximated by the images produced by lighting that consists of nine, low-frequency spherical harmonics. So the set of images produced by a convex Lambertian object is approximately nine-dimensional. [8] shows that the fact that only some surface normals will face a camera reduces this dimension further.

These results do not consider cast shadows. Recently, [9] analyzes some canonical situations in which shadows are cast. The set of images produced by a scene with cast shadows can be of much higher dimension, although [9] provides some empirical evidence that in many cases this dimension does not grow too rapidly.

Some methods have been proposed to recover illumination distributions from images. [14] proposes a framework to accomplish photo-realistic view-dependent image synthesis from a sparse image set and a geometric model. It uses residual images to estimate a rough approximation of the illumination and introduces a Torrance-Sparrow reflection model to refine the result. [17] presents two methods for recovering the light source position from a single image without the distant illumination assumption. [15] extracts much more accurate multiple illumination information from the shading of a sphere. These works haven't dealt with cast shadows specifically. The complexity of determining lighting grows dramatically when we must account for cast shadows. The method that is most closely related to our work is [10]. In [10], they demonstrate the effectiveness of using occluding

information of incoming light in the estimation of the illumination distribution of a scene. They use adaptive sampling to avoid dense sampling of the illumination. Since they use a shadow image to estimate the radiance distribution, the pixels in the shadow have to be known beforehand and it is hard to do this automatically in many situations. Moreover, it cannot handle the non-convex object which will cast shadows on itself.

To overcome their shortcomings, we take a different approach to measure the complexity of cast shadows in this thesis. We represent images using a basis and images produced by single, directional light sources. We show that the effects of cast shadows can be represented sparsely in this basis, using a small number of directional light sources. Therefore, the illumination of the scene can be recovered by adding a certain number of directional sources to the spherical harmonic lighting. This leads to a representation that has few parameters, although it is not low-dimensional.

This thesis is organized as follows: In Chapter 2, we describe the basic shadow concepts and the spherical harmonic analysis on the Lambertian model we assume for the objects. Shadow concepts such as the umbra and the penumbra, attached and cast shadows are introduced and examples are given to illustrate the difference between them. Examples are also given to show the rich information provided by cast shadows. A spherical harmonic representation of Lambertian objects is briefly reviewed and its shortcomings are shown by a simple example using images with prominent cast shadows.

In Chapter 3, we begin with a simple example to illustrate this. We show that the effects of cast shadows may not be well approximated by any low-dimensional

representation. However, when only a few directional light sources illuminate a scene, they may be compactly represented, while a large number of directional sources do not produce strong shadows. Then we present a simple model of this situation that we can analyze. We derive a bound on the possible effects of cast shadows that are not captured by representing light with low-frequency harmonics and a few directional light source. Using one directional light source reduces this bound by half, and further light sources decrease the bound by the square root of the number of sources. Next, we briefly review several algorithms designed to either find the best directional sources or approximate the images with cast shadows. Finally, we propose a fast and accurate multiscale search algorithm to find the best directional sources, and we give some comments.

In Chapter 4, we conduct numerous experiments on synthesized data to test the validity of our assumptions and apply our proposed algorithm to illumination recovery. We evaluate the performance of several algorithms and our proposed algorithm in terms of processing time and accuracy.

Finally, conclusions and future perspectives are drawn in Chapter 5.

Chapter 2

Spherical Harmonic Analysis of Lambertian Objects with shadows

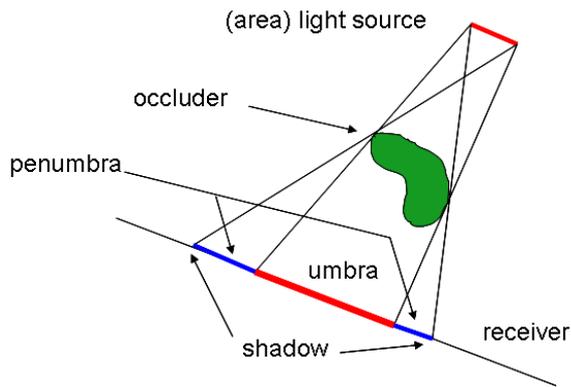
2.1 What are the Shadows?

The first step in the development of efficient tools for recognizing objects with shadows in digital images and image sequences is an understanding of how shadows appear in images and what is peculiar to them. Shadows are crucial for human perception of the 3D world.

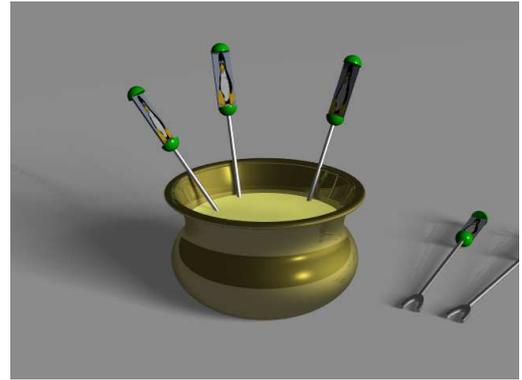
The darkest part of a shadow is the *umbra*. A point P of the scene is considered to be in the *umbra* if it is completely blocked by the object causing the shadow, i.e. it doesn't receive any light from the light source. If the light is partially blocked, hence P can view a part of the light source, it is in the *penumbra*. The union of the umbra and the penumbra is the shadow, the set of points for which at least one point of the light source is occluded. Objects that hide a point from the light source are called occluders. Figure 2.1 (a) shows the generation of umbra and penumbra. Area light sources emit light to the receiver and some lights are blocked by the occluder. The blue part in the penumbra on the receiver can only see part of the light source, while the red part in the umbra cannot see anything from the light source. The umbra and penumbra structure is clearly visible in figure 2.1 (b).

Shadows are categorized as two types:

- attached shadows or self shadows, that occur when a surface faces away from



(a)



(b)

Figure 2.1: Umbra and penumbra generation. Area light sources, as the line source in (a), generate penumbra where the light is only partially obstructed by the shadow casting object. The umbra and penumbra structure is clearly visible in (b).

a light source.

- cast shadows, that occur when an intervening part of an object blocks the light from reaching a different part of the surface.

For convex objects, only attached shadows occur. Attached shadows are easily modeled, since they depend only upon the local geometry of the surface which is the normal direction of the point on the surface. On the other hand, cast shadows are caused when an entirely different region of the surface intersects the path from the light source to the point. Since they are dependent on the global geometry of the surface, cast shadows are more complex to model. Figure 2.2 (a) illustrates the generation of attached and cast shadows. An example of attached and cast shadows in real life is given in figure 2.2 (b). The left side of the cat condo, which faces away from the light source, is in the attached shadow. The cat condo casts a hard shadow

on the background.

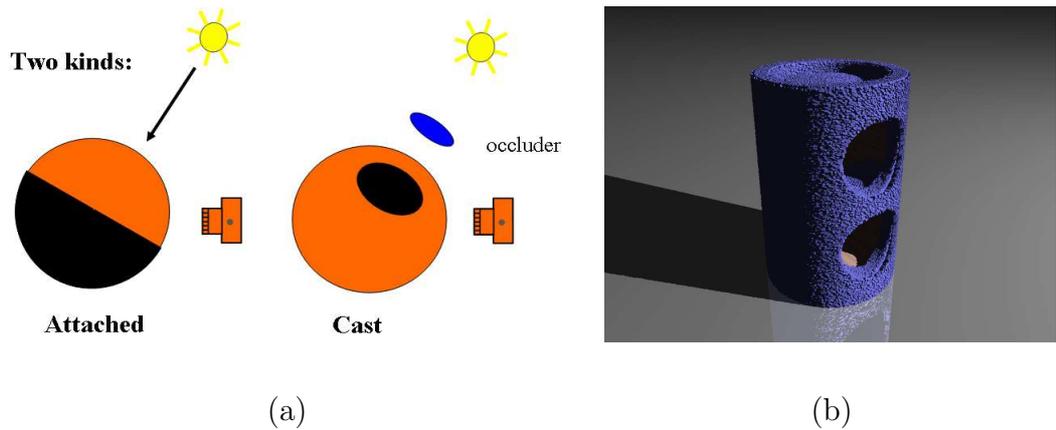


Figure 2.2: Examples of attached and cast shadow. The generation of attached and cast shadow is illustrated in (a). (b) gives an example of attached and cast shadows in real life. The left side of the cat condo, which faces away from the light source, is in the attached shadow. The cat condo casts a hard shadow on the background.

Shadows provide a strong source of information about the shape of surfaces, which we now illustrate with a concrete example.

Shadows help to understand relative object position and size in a scene. In figure 2.3 (a), we are unable to determine the position of an object in space without a cast shadow, whereas on the other three images we understand that it is more and more distant from the ground.

Shadows can also provide information about the shape and geometry of a complex occluder. From the shadow cast on the ground in figure 2.4 (a), we can see the occluder is a man on a racing horse. We can also tell there is a pin sticking in the ground from 2.4 (b).



(a)



(b)



(c)



(d)

Figure 2.3: Shadow provide information about the relative positions of objects. We cannot determine the position of the wooden dog from shadowless image (a), whereas on the other three images we understand that it is more and more distant from the ground.

2.2 Lambertian Model

We assume that a surface exhibits Lambertian reflectance, which states that light falling on it is reflected equally in all directions and the brightness of the surface is the same to an observer regardless of the observer's angle of view. In figure 2.5 (a), the light falling on the Lambertian surface is reflected equally in all directions and appears the same to the viewer from any viewing direction. The bidirectional



(a)



(b)

Figure 2.4: Shadow provide information about the shape and geometry of the occluder.

reflection distribution function(BRDF) for a Lambertian surface is known to be a constant. Under the Lambertian law, the intensity of a point p is calculated by taking the dot product of the surface normal and lighting direction (see figure 2.5 (b))

$$I(p) = n(p) \cdot l \quad (2.1)$$

If the image of a Lambertian object has no shadows, the set of all images under all lighting conditions is a 3-dimensional space. For any lighting direction l , it can be represented by 3 non-coplanar basis lighting directions $\{l_1, l_2, l_3\}$. Given that n is the surface normal of a surface point, the image brightness value of the point is

$$I(p) = n(p) \cdot l = n \cdot (a_1 l_1 + a_2 l_2 + a_3 l_3) = a_1 n(p) \cdot l_1 + a_2 n(p) \cdot l_2 + a_3 n(p) \cdot l_3 \quad (2.2)$$

The brightness value of a surface point is a linear combination of the brightness value of the same point under 3 fixed, independent illuminations.

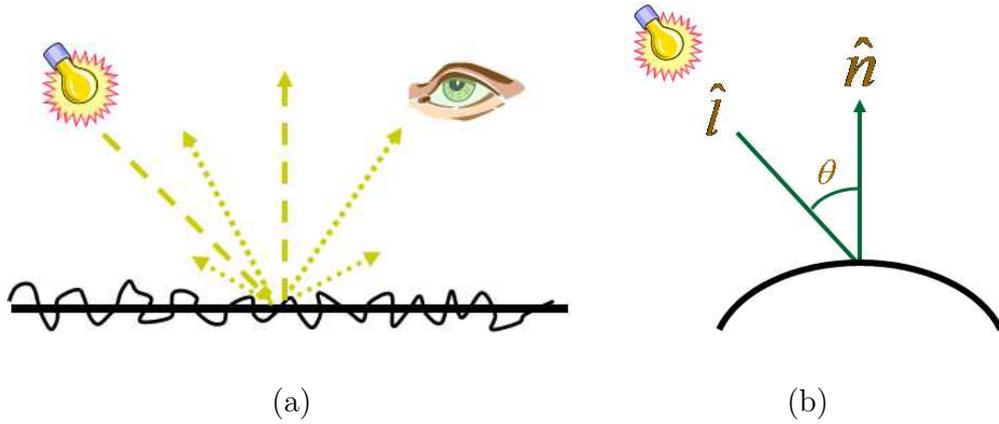


Figure 2.5: Lambertian surface and law. In (a), the light falling on the Lambertian surface is reflected equally in all directions and appears the same to the viewer from any viewing direction. According to Lambertian law, the intensity of a point is the dot product of surface normal and lighting direction as in (b).

2.3 Spherical Harmonic Analysis

Problems arise when the point is in the attached shadow. The angle between the surface normal and the lighting direction is obtuse ($n(p) \cdot l < 0$), so space for the set of images under all lighting conditions is no longer 3-dimensional. Basri and Jacobs [6] show that it is very close to 9D linear subspace using spherical harmonic analysis. It is analogous to Fourier analysis, but on the surface of sphere. With a spherical harmonic representation, low frequency light means light whose intensity varies slowly as a function of direction. The image formation is analogous to the convolution of the lighting function with a cosine function. The reflectance function acts as a low-pass filter with 99.2 percent of its energy in the first nine components.

In the following, we briefly review the spherical harmonic analysis for attached shadows. We denote lighting direction and surface normal using unit vector u_l and

v_r , respectively. According to Lambert's law, if a light ray of intensity l coming from direction u_l reaches a surface point with albedo ρ and normal direction v_r , then the intensity, i , reflected by the point due to this light is given by

$$i = \rho l(u_l) \max(u_l \cdot v_r, 0) \quad (2.3)$$

If we fix the lighting and ignore ρ for now, we can write

$$r(v_r) = \int_{S^2} k(u_l \cdot v_r) l(u_l) du_l \quad (2.4)$$

where $k(u \cdot v) = \max(u \cdot v, 0)$. The reflected light on a point is a function of surface normal alone. Intuitively, it can be regarded as a convolution of k and l .

The surface spherical harmonics are a set of functions that form an orthonormal basis for the set of all functions on the surface of the sphere. Any piecewise continuous function f on the surface of the sphere can be written as a linear combination of an infinite series of harmonics. Specifically, for any f ,

$$f(u) = \sum_{n=0}^{\infty} \sum_{m=-n}^n f_{nm} Y_{nm}(u) \quad (2.5)$$

where f_{nm} is a scalar value, computed as:

$$f_{nm} = \int_{S^2} f(u) Y_{nm}^*(u) du \quad (2.6)$$

The first nine spherical harmonics which are a function of space coordinates

(x, y, z) are:

$$\begin{aligned}
 Y_{00} &= \frac{1}{\sqrt{4\pi}} & Y_{10} &= \sqrt{\frac{3}{4\pi}}z \\
 Y_{11}^e &= \sqrt{\frac{3}{4\pi}}x & Y_{11}^o &= \sqrt{\frac{3}{4\pi}}y \\
 Y_{20} &= \frac{1}{2}\sqrt{\frac{5}{4\pi}}(3z^2 - 1) & Y_{21}^e &= 3\sqrt{\frac{5}{12\pi}}xz \\
 Y_{21}^o &= 3\sqrt{\frac{5}{12\pi}}yz & Y_{22}^e &= \frac{3}{2}\sqrt{\frac{5}{12\pi}}(x^2 - y^2) \\
 Y_{22}^o &= 3\sqrt{\frac{5}{12\pi}}xy
 \end{aligned} \tag{2.7}$$

where the superscripts e and o denote the even and odd components of the harmonics, respectively. $Y_{nm} = Y_{n|m}^e \pm iY_{n|m}^o$. Figure 2.6 shows the first nine spherical harmonics.

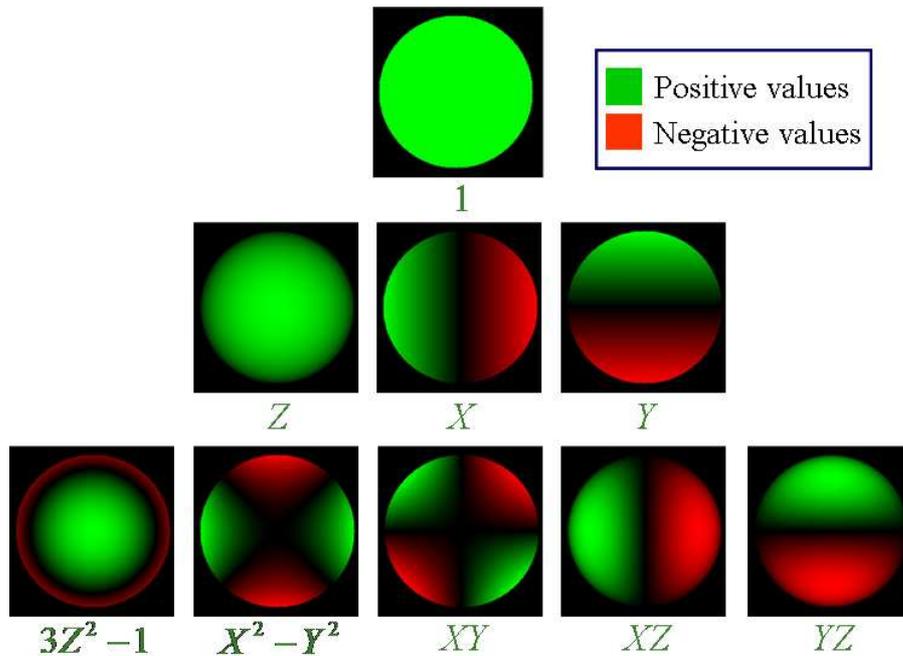


Figure 2.6: The first nine spherical harmonics.

Both the lighting function l , and Lambertian kernel k , can be written as sums

of spherical harmonics. Denoted by

$$l = \sum_{n=0}^{\infty} \sum_{m=-n}^n l_{nm} Y_{nm} \quad (2.8)$$

the harmonic expansion of l , and by

$$k(u) = \sum_{n=0}^{\infty} k_n Y_{n0} \quad (2.9)$$

Note that since $k(u)$ is circular symmetric about the north pole, so

$$\int_{S^2} k(u) Y_{nm}^*(u) du = 0, m \neq 0 \quad (2.10)$$

According to the Funk-Hecke theorem, the harmonic expansion of the reflectance function r can be written as:

$$r = k * l = \sum_{n=0}^{\infty} \sum_{m=-n}^n (\alpha_n l_{nm}) Y_{nm} \quad (2.11)$$

where $\alpha = \sqrt{\frac{4\pi}{2n+1}} k_n$.

The first few coefficients of Lambertian kernel are

$$\begin{aligned} k_0 &= \frac{\sqrt{\pi}}{2} \approx 0.8862 & k_1 &= \sqrt{\frac{\pi}{3}} \approx 1.0233 \\ k_2 &= \frac{\sqrt{5\pi}}{8} \approx 0.4954 & k_4 &= -\frac{\sqrt{\pi}}{16} \approx -0.1108 \\ k_6 &= \frac{\sqrt{13\pi}}{128} \approx 0.0499 & k_8 &= -\frac{\sqrt{17\pi}}{256} \approx -0.0285 \end{aligned} \quad (2.12)$$

Figure 2.7 is a graph representation of the first 9 coefficients and the cumulative energy of the Lambertian kernel. Because the Lambertian kernel k acts as a low-pass filter, the high-frequency components of the lighting have little effect on the reflectance function. We achieve a low-dimensional approximation to the reflectance function by truncating the sum in (2.11).

$$r = k * l = \sum_{n=0}^{\infty} \sum_{m=-n}^n (\alpha_n l_{nm}) Y_{nm} \approx \sum_{n=0}^N \sum_{m=-n}^n (\alpha_n l_{nm}) Y_{nm} = \sum_{n=0}^N \sum_{m=-n}^n l_{nm} r_{nm} \quad (2.13)$$

where r_{nm} is given by

$$r_{nm} = k * Y_{nm} = \alpha_n Y_{nm} \quad (2.14)$$

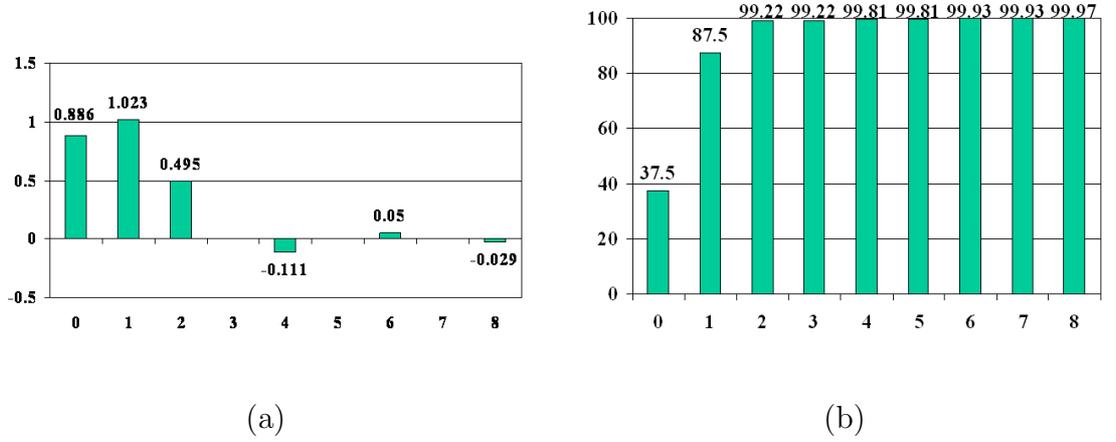


Figure 2.7: A graph representation of the first 9 coefficients and the cumulative energy of the Lambertian kernel.

Using this analysis, we can efficiently represent the set of images of objects seen under varying illumination. Let p_i denote the i th point on the surface of object. Let n_i and ρ_i denote the surface normal and albedo of p_i , respectively. Then, the image I_i of p_i is:

$$I_i = \rho_i r(n_i) = \rho_i \sum_{n=0}^{\infty} \sum_{m=-n}^n l_{nm} r_{nm} = \sum_{n=0}^{\infty} \sum_{m=-n}^n l_{nm} b_{nm}(p_i) \quad (2.15)$$

where $b_{nm}(p_i) = \rho_i r_{nm}(n_i)$ is the harmonic image. We can see any image is a linear combination of harmonic images.

We recognize objects by comparing a new query image to the linear subspace of images that correspond to each model in turn. Given an image I , we seek the distance from I to the space spanned by the basis images. Let B denote the basis images. Then we seek a vector a that minimizes $\|Ba - I\|$. Every column of B contains one harmonic image b_{mn} . This shows very good results for face recognition.

If there are cast shadows in the image, the dimension of the linear subspace spanned by the set of images under varying illuminations can be very large. To see the impact of cast shadows on object recognition, we consider a simple example. In figure 3.1, two boxes have the same width and height, but different thickness. They appear the same to the viewer from the front. We can't differentiate them without considering cast shadows on the ground. By applying recognition with nine harmonics, we obtain wrong recognition results. This encourages us to develop another algorithm which can handle images with significant cast shadows.

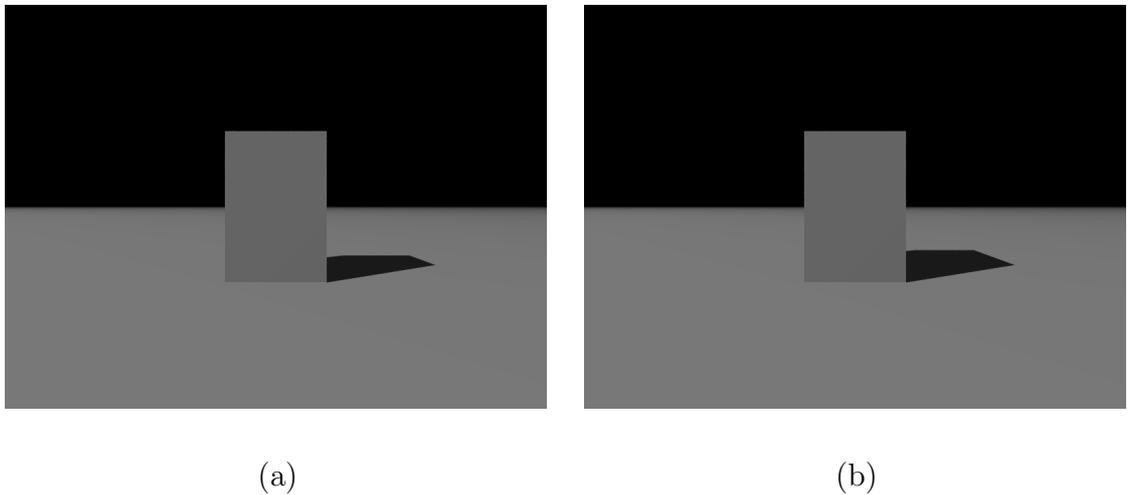


Figure 2.8: Images of two boxes. The only difference is the thickness. They can only be differentiated by their shadows cast on the ground.

Chapter 3

Analysis of Images with Cast Shadows

3.1 An Example

To strengthen our intuitions, we consider a very simple example of a scene consisting of a flat playground with an infinitely thin flag pole. We view the scene from directly above, so that the playground is visible, but the flag pole appears only as a negligible point. Suppose the scene is illuminated by an arbitrary set of directional lights of equal intensity that each have an elevation of 45 degrees. In this case, the intensity of the lighting can be described as a one-dimensional function of azimuth. A single directional light illuminates the playground to constant intensity except for a thin, black shadow on it. The entire set of lights can cause shadows in multiple directions. None of these shadows overlap, because the pole is infinitely thin.

Now consider the linear subspace spanned by the images that this scene can produce. We first consider the set of images that are each produced by a single directional source. All images are nonnegative, linear combinations of these. We represent each image as a vector. The principal component of these images will be the constant image produced in the absence of cast shadows; since each shadow is infinitely thin, the non-shadowed part dominates each image. Suppose we then project each image into the space orthogonal to this principal component. We get

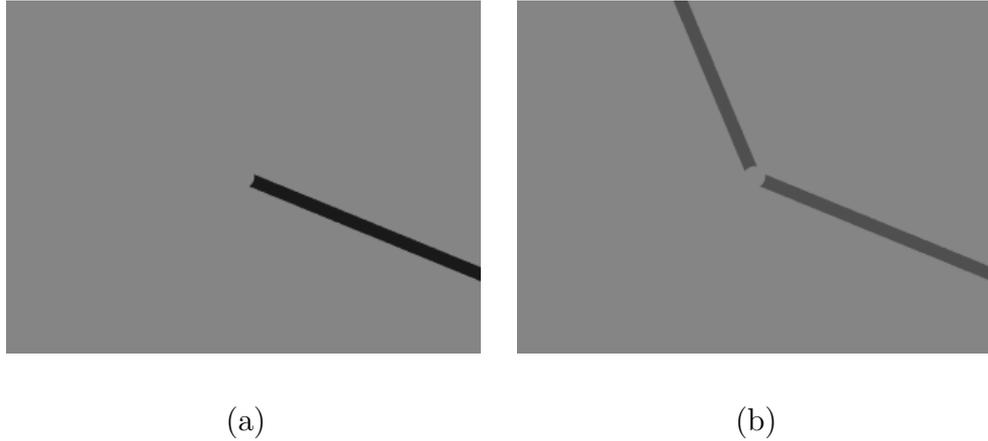


Figure 3.1: Images of a flat playground with a thin flag pole. (a) is rendered with only single directional source, while (b) is rendered with 2 directional sources. (b) has two shadows, but each shadow only has half the intensity of the shadow in (a). an image that is near zero, except for a large negative component at the shadow. All these images have equal magnitude, and are orthogonal to each other. Therefore, they span an infinite-dimensional space, and PCA will produce an infinite number of equally significant components. A finite-dimensional linear subspace cannot capture any significant fraction of the effects of cast shadows.

But, let's look at the images of this scene differently. A single directional source produces a single, black shadow. Two sources produce two shadows, but each shadow has half the intensity of the rest of the playground, because each shadow is lit by one of the lights. The more lights we have the more shadows we have, but the lighter these shadows are. Intuitively, we expect many very light shadow to be less important perceptually than a few dark ones. In this paper we will measure the magnitude of an image, or the magnitude of the error in an approximation by the root mean square (RMS) value of each pixel, that is, the norm of the image.

With this measure, we can quantify our intuition by pointing out that the norm of the difference between a shadowed image generated with N sources and a constant image will be proportional to $\sqrt{N(\frac{1}{N})^2}$, and so it will decrease in proportion to \sqrt{N} .

Now, suppose we approximate any possible image using one image of constant intensity, and a small number of images that are each produced by a directional source. If the actual image is produced by a small number of directional sources, we can represent its shadows exactly. If the image is produced by a large number of directional sources, we cannot represent the shadows well with a few sources, but we do not need to, because they have only a small effect.

3.2 Analysis

We now analyze an idealized model of cast shadows. This model will not hold for every possible scene, but we argue that it captures significant properties of common situations. We make several key assumptions. First, we assume that all images produced by a single, directional source will have a significant component in a common, low-dimensional linear subspace. This is motivated by the fact that images without shadows are known to be well-approximated by a low-dimensional subspace, and shadowed portions of the scene will typically be relatively small. Second, we assume that the norm of the image will be proportional to the number of directional sources and their intensity. This assumption can be false. For example, if each source illuminates a single, separate pixel with an intensity of 1, then an image with N sources will have a norm of \sqrt{N} . However, we claim that both assumptions

are reasonable for many situations in which cast shadows appear, and are worth investigating. We note that these assumptions are true for the example of the flag pole discussed in the previous section.

We will analyze the case of a scene illuminated by a finite number of directional light sources, N . We let I_k denote the image that would result when the scene is illuminated by just light source k , normalized so that $\|I_k\| = 1$. The actual image, then, is a nonnegative combination of these single-source images, that is:

$$I = \sum_{k=1}^N a_k I_k \quad (3.1)$$

for $a_k \geq 0$. We will use the notation: $A \equiv \sum_{k=1}^N a_k$. We also assume, without loss of generality, that the single-source images are ordered so that $a_1 \geq a_2 \geq \dots \geq a_N$.

Each single-source image can be described as having some component in a common, d -dimensional subspace, S , and some component orthogonal to this space. We write this as:

$$I_k = \alpha \vec{c}_k + \beta \vec{r}_k \quad (3.2)$$

where \vec{c}_k is a unit vector in the subspace S , and \vec{r}_k is a unit vector orthogonal to this subspace. We make the simplifying assumption that α is constant for all images produced by a single, directional source. We will also assume that $\langle \vec{r}_k, \vec{r}_j \rangle = 0$ for $k \neq j$. We relax these assumptions later. Because I_k is normalized, we have $\alpha^2 + \beta^2 = 1$.

Now we approximate I using the linear subspace and m images generated with single directional sources, I_1, I_2, \dots, I_m . Note that these are the m images that play the biggest role in creating I . Let \bar{I}_m be the best approximation to I that we can

generate from a linear combination of these images and the subspace. We have:

$$\bar{I}_m = \sum_{k=1}^N a_k \alpha \vec{c}_k + \sum_{k=1}^m a_k \beta \vec{r}_k \quad (3.3)$$

We can measure the quality of this approximation by looking at the angle, θ , between I and \bar{I}_m . We have:

$$\|I\| \sin \theta = \|I - \bar{I}_m\| \quad (3.4)$$

At this point we make use of the assumption that $\|I\| \propto A$. This is our second key assumption, that the magnitude of the image grows proportionally with the sum of the intensity of light falling on the scene. Then we have:

$$\sin \theta \propto \frac{\|\sum_{k=m+1}^N a_k \beta \vec{r}_k\|}{A} = \frac{\beta (\sum_{k=m+1}^N a_k^2)^{\frac{1}{2}}}{A} \quad (3.5)$$

The value of θ will vary depend on the distribution of the a_k values. Of course, θ can be as small as zero when $a_k = 0$ for $k > m$, in which case our sparse approximation perfectly captures the lighting. We now want to understand how badly a sparse approximation can do. This is described by the maximum possible value of θ , and we want to see how this depends on the choice of m .

To get at this issue, we consider the maximum possible value that can be obtained by the right-hand side of Equation 3.5. First, we note that for any fixed value of a_m , which is the least influential light source used in our approximation, this expression will be maximized when $a_1 = a_2 = \dots = a_m$. This minimizes the effect of the lights used in the approximation, and maximizes the effect of the rest of the lights. So this expression is maximized when $\sum_{k=m+1}^N a_k^2$ is maximized subject to $\sum_{k=m+1}^N a_k = A - ma_m$ and $0 \leq a_k \leq a_m$ for $k > m$. This maximum occurs

when the leading coefficients, a_{m+1}, \dots have the maximum possible value, and all other coefficients are 0. That is, the right-hand side of (3.5) is maximized when $a_{m+1}, \dots, a_{m+p} = a_m$ for some choice of p . That is, we must find p to maximize:

$$\frac{\sqrt{pa_m^2}}{(m+p)a_m} = \frac{\sqrt{p}}{m+p} \quad (3.6)$$

Taking the derivative with respect to p and setting it to zero shows that this is maximized when $p = m$. This tells us that in the worst case, for an image generated by any set of directional light sources, we have: $a_1 = a_2 = \dots = a_{2m}, a_{2m+1} = \dots = a_N = 0$. Substituting this into Equation 3.5 we have:

$$\sin\theta \propto \frac{\beta\sqrt{m}}{2m} = \frac{\beta}{2\sqrt{m}} \quad (3.7)$$

To summarize, we have considered the angle between an image and an approximation to it based on a low-dimensional subspace and a sparse set of m directional light sources. For a simplified model of this situation, we have shown that maximum possible value of the sine of this angle will decrease at the rate of $\frac{1}{\sqrt{m}}$. For example, if we use one directional source in our approximation ($m = 1$), the worst case for us will be when there are two, equal intensity directional sources. In this case, we capture half of the difference between the image and our low-dimensional subspace using one source. Four directional sources guarantee that we will capture $\frac{3}{4}$ of the effect, and so on.

We have made a few, critical assumptions to arrive at this conclusion. First of all, our derivation assumes that the vectors describing the effects of cast shadows for each source, \vec{r}_k , were all orthogonal. We can derive the same result without this assumption. First, we define \bar{I}_k to be the component of I that is orthogonal to the

low-dimensional linear subspace and the images produced by the first k sources. We reorder the light sources so that $\langle a_1 \vec{r}_1, \bar{I}_0 \rangle \geq \langle a_k \vec{r}_k, \bar{I}_0 \rangle$ for any $k > 1$. (Note that when the \vec{r}_k are orthogonal this amounts to the ordering we use above, in which $a_1 \geq a_k$). Then we rewrite the remaining images expressing only their component orthogonal to \bar{I}_1 , so that each \vec{r}_k is orthogonal to this space, and each a_k represents the magnitude of each image orthogonal to \bar{I}_1 . Next, we reorder the remaining sources so that $\langle a_2 \vec{r}_2, \bar{I}_1 \rangle \geq \langle a_k \vec{r}_k, \bar{I}_1 \rangle$ for any $k > 2$. We continue this process until all the sources are accounted for. We can then apply our initial derivation to this representation of the images.

A second simplifying assumption appeared in our use of a single β value for all images. In reality, we expect β to vary somewhat from image to image, as the effect of cast shadows varies. For example, we know that there will be no cast shadows when the light direction is the same as the camera direction, while a light near the horizon may produce long prominent shadows. We expect that this variation will not be too great, and that it will suffice to use an average β value in our predictions.

A more critical assumption in our derivation was that $\|I\|$ will be approximately A . Intuitively, we expect that the image produced by each directional source will lie mostly in a low-dimensional linear subspace, and that as we sum together the components of these images in that subspace, the magnitude of the resulting image will grow with the number of images. This occurs when images from different directional sources are not orthogonal, but are highly correlated. It can be seen that this is exactly true for the example of the flag pole in the playground. In general, one can construct images with cast shadows and different directional sources that

are orthogonal; our approach rests on the conjecture that in most situations, when a Lambertian scene is lit by a large number of directional sources, the images will not tend to be orthogonal.

On the other hand, for example, this assumption would not be true for the non-Lambertian effects produced by a mirrored ball. In that case, each directional source lights a different point on the ball, and these images are orthogonal. Consequently, our sparse representation of an image should be appropriate only for situations in which images due to different directional sources are highly correlated.

3.3 Several Algorithms

From the previous derivation, we arrive at the conclusion that an image with cast shadows can be well approximated using a combination of low frequency spherical harmonics, and a small number of directional sources. So the problem is reduced to finding the directional sources which can best approximate the image. There are several algorithms proposed to either find the best directional sources (indirect) or approximate the images with cast shadows (direct). In the following, we categorize them to be direct and indirect algorithms and briefly review and analyze each of them.

3.3.1 Indirect Algorithms

1. Brute-force search: First, we define \bar{I}_k to be the projection error of I projected on the low-dimensional linear subspace and the images produced by the first k

sources. We reorder the light sources so that $\langle a_1 \vec{r}_1, \bar{I}_0 \rangle \geq \langle a_k \vec{r}_k, \bar{I}_0 \rangle$ for any $k > 1$. (Note that when the \vec{r}_k are orthogonal this amounts to the ordering we use above, in which $a_1 \geq a_k$). Then we rewrite the remaining images expressing only their component orthogonal to \bar{I}_1 , so that each \vec{r}_k is orthogonal to this space, and each a_k represents the magnitude of each image orthogonal to \bar{I}_1 . Next, we reorder the remaining sources so that $\langle a_2 \vec{r}_2, \bar{I}_1 \rangle \geq \langle a_k \vec{r}_k, \bar{I}_1 \rangle$ for any $k > 2$. We continue this process until all the sources are accounted for.

Brute-force search is the most time consuming of all algorithms. At each step, it has to check all the candidate directional sources to see if they satisfy the minimum error approximation. The complexity depends on the number of points uniformly sampled on the surface of the upper hemisphere n and the number of directional sources m we want to find. The number of search steps to find the best m directional sources is $O(mn)$. Brute-force search guarantees that it will find the best directional source at each step, but the number of search steps is proportional to the number of candidate solutions and grows very quickly as the number of candidate directional sources increases.

2. Simulated Annealing: The optimization technique simulated annealing is applied to avoid the cost of brute-force search by directing the search successfully. It reduces the search time by orders of magnitude. The original version of simulated annealing has been applied to segmentation and noise reduction of degraded images [11]. Simulated annealing was chosen because of its ability to avoid terminating at local minima and to keep searching for the global

minimum. Simulated annealing is analogous to the annealing process for metals, wherein the material is raised to a high temperature and then gradually cooled, allowing the atoms to settle into their most desirable states.

In our case the desirable end state is the best directional sources, and atomic motion due to thermal energy is simulated by a randomization routine which allows 'uphill' acceptances of states with probability controlled by the temperature parameter. Parameters of the simulated annealing process are: starting temperature T , rate of cooling L (measured in iterations at each temperature), and cooling schedule α (change in temperature, $T' = \alpha T$, after each set of iterations, L). These parameters directly affect the run time of simulated annealing.

Simulated annealing is outlined as follows:

- (a) Initialize temperature $T = T_0$.
- (b) Find the best m directional sources p_1, p_2, \dots, p_m out of 65 uniformly sampled sources using brute-force search .
- (c) On the grid uniformly sampled 3871 points on the surface of a hemisphere, we obtain the energy function E_i and E_{ij} for each point p_i and its 5 nearest neighbors p_{ij} , where $i = 1, 2, \dots, m$ and $j = 1, 2, \dots, 5$. If $e^{-\frac{(E_i - E_{ij})}{T}} > \text{random}[0, 1)$, then keep p_i , otherwise, replace p_i with its corresponding nearest neighbor p_{ij} .
- (d) Keep going until reaching the maximum number of iteration.

- (e) Keep reducing the temperature T by $T' = \alpha T$ and go to step 3 until reaching the lowest temperature.

Simulated annealing can reduce computational cost by a certain amount, but it depends on the rate of cooling and might take a long time to converge. Ideally, it will lead to a globally optimal solution as long as the cooling rate is small enough.

3.3.2 Direct Algorithms

- 3. Semidefinite programming(SDP):[12] applies semidefinite programming to perform a constrained optimization to quickly and accurately solve the non-negative linear combination of spherical harmonics. It has been successfully applied for the specular object recognition on both synthetic and real data by better separating the correct and incorrect models. Their SDP algorithm is summarized as follows:

- (a) First, spherical harmonic images are obtained by rendering the 3D objects with each individual harmonic lighting.
- (b) Second, these harmonic images are vectorized and stacked as columns of a matrix M . The resulting image is described as Ma which is the product of spherical harmonic images M and coefficient vector a .
- (c) Third, given the query image $r = I + noise$, a is found by minimizing

$\|Ma - r\|$ subject to $T_L(a) \geq 0$. where

$$\mathbf{T}_L(\mathbf{a}) = \begin{bmatrix} a_0 & a_1 & \cdots & a_n \\ a_{-1} & a_0 & \ddots & \\ \vdots & \ddots & \ddots & a_1 \\ a_{-n} & & a_{-1} & a_0 \end{bmatrix}$$

(d) Fourth, compute the QR decomposition of the matrix M . The problem is reduced to: solve the size $(L+1)^2$ problem: $\min_a \|Ra - Q^T r\|^2$ subject to $T_L(a) \geq 0$. This kind of problems are called semidefinite programming (SDP) problems.

(e) Finally, the problem is solved as: $\min_a q$ subject to

$$1 + q \geq \begin{bmatrix} 1 - q \\ Ra - Q^T r \end{bmatrix}$$

and $T_L(a) \geq 0$. It's solved in MATLAB using SDPT3 and YALMIP packages.

SDP is designed to approximate high frequency signals which cannot be captured by the 9D spherical harmonics. It works well on specular objects such as a shiny rubber ball and a ceramic shaker using harmonics up to 10th order. Since images with cast shadows generally have a lot of high frequency signals, it still misses a certain amount of information which is contained in higher order harmonics.

4. Non-Negative Linear: [4] has shown that the set of images of an object produced by nonnegative lighting is a convex cone in the space of all possible

images. Given an image I , we attempt to minimize $\|Ha - I\|$ subject to $a \geq 0$ where H is the matrix whose columns are directional source images. If we densely sample the illumination distribution, it makes the solution exceedingly expensive in terms of processing time and storage requirements because of the high dimensionality of the matrix H formed by point source images.

3.4 Multiscale Search Algorithm

We propose a multiscale search algorithm to find the best directional sources. It employs an s scale nearest neighbor search from the coarsest level to the finest level. It achieves a good balance between processing time and accuracy.

The multiscale search algorithm is summarized as follows (see Figure 3.2).

1. First, we find the best m out of 65 sources at the coarsest level using a brute-force search algorithm.
2. For each m directional sources, we check the k nearest neighbors in a finer level and pick the one which has the smallest angle between the approximated image and the query image.
3. We go to a finer level and repeat step 2, until we reach the finest level.

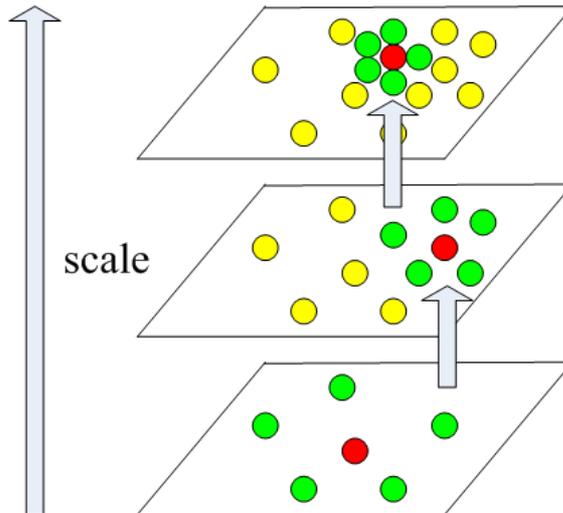


Figure 3.2: Best directional source is selected by comparing it with its 5 nearest neighbors on each level.

3.5 Comments on Multiscale Search Algorithm

First, the search for nearest neighbor is conducted on the surface of the hemisphere. Therefore, practically speaking, there are no 'boundary' points, that is, we can always find a certain number of nearest neighbors for every point no matter where the point is located.

Second, the number of search steps only depends on the number of scales s and point light sources m when the number of nearest neighbors k is given. Thus, the total number of search steps is $s \cdot k \cdot m + O(mn_0)$, where $O(mn_0)$ is the number of exhaustive search steps at the coarsest level with n_0 images. This is much less than the number of exhaustive search steps $O(mn)$ where n is the number of images at the finest level which is 3871 in our experiment.

Third, in the brute-force search algorithm, we find the best directional source

of all the candidate sources first. Then, we find the best one out of the rest of the sources, and so on. Therefore, the best directional sources are ranked according to their importance. Whereas for the multiscale search algorithm, this is only true at the coarsest level and it no longer holds for the finer levels. That is, the order of the importance of the best directional sources at the finest level is changed. That's why we see a smooth curve in the brute-force search rather than some abrupt drops in the multiscale search.

Chapter 4

Experiments

We have conducted numerous experiments on synthesized data to test the validity of our assumptions and applied our proposed algorithm to illumination recovery. We evaluate the performance of several algorithms and our proposed algorithm in terms of processing time and accuracy.

4.1 Testing the Validity of Assumptions

We now present experiments to test the validity of our assumptions. We will check three things. First, for a given scene we can assess the empirical values that occur for α and β . Second, we can measure the extent to which the magnitude of an image is proportional to the magnitude of the sum of the light sources. And third, we can evaluate our prediction that as we use more directional light sources, the maximum possible error in our approximation will be bounded by $\frac{\beta}{2\sqrt{m}}$.

To implement our approach we must choose a low-dimensional linear subspace to approximate the set of images that a scene will produce. We use a nine-dimensional subspace generated by rendering images of the scene, including their cast shadows, using lighting that consists of zero, first, and second order spherical harmonics. [6, 7] have shown that this subspace will approximate a scene's image well in the absence of cast shadows, so it seems a reasonable choice. A reasonable



Figure 4.1: Sample images of two scenes, each generated with a single, directional source.

alternative would be to render the scene ignoring cast shadows. We would expect that ignoring cast shadows would lead to a less accurate subspace, but it can be computed analytically from the scene's surface normals, without ray tracing.

We have experimented using numerous scenes. We show two in Figure 4.1. Results were very similar for all scenes, so for the sake of brevity we only report results for the scenes in Figure 4.1. Using the POV-Ray ray tracer we generate directional images, each using a single directional light source. We obtain directions by uniformly sampling the upper hemisphere. Using these images, we numerically integrate to compute nine images of the scene, each with lighting consisting of a single spherical harmonic.

We first consider the variation that occurs in β for each image. Our analysis used the simplifying assumption of a constant β , but we know that in reality β will vary. For the scene in Figure 4.1-left we find that β has a mean value of 0.1970 with a standard deviation of 0.0524, and minimum and maximum values of 0.0751 and

0.4136. For the scene in Figure 4.1-right we find that β has a mean value of 0.1507 with a standard deviation of 0.04, and minimum and maximum values of 0.0830 and 0.2721. So there is real, but not extremely large variation in β . To predict the effect of using more directional sources to approximate the image we use a value of β that is an average of the β values for each light source, weighted by the magnitude of the image that light source produces.

Second, we test our prediction that $\|I\| \propto A$. We do this for sets of k images in which k varies from two to sixteen. For each k we render k images with randomly chosen light sources. We normalize each image to unit length, and then measure the magnitude of the image we get by summing these images. We repeat this experiment one hundred times for each k . On average, we find that $\frac{\|I\|}{A}$ has a mean value of .96 for $k = 2$, .95 for $k = 3$, .94 for $k = 4, 5$, .93 for $k = 6$ to 15, and .92 for higher values of k up to 32. These numbers all have a standard deviation of about .05. This supports our contention that we can expect the magnitude of an image of a Lambertian scene to grow proportionately to the sum of the magnitude of the individual images produced by each light source; in fact we find that $\|I\|$ is quite close to A .

Finally, our key prediction is that we can represent an image using a harmonic subspace of images and m images rendered with directional light sources so that the error in this representation will be no more than $\frac{\beta}{2\sqrt{m}}$ (or β for $m = 0$). Of course, the error can be much less than this, and as little as zero. We test this first using images generated with directional light sources. We expect well distributed light sources to be the most difficult to approximate. For each value of these images, we

approximate the image with a linear subspace and m directional sources, allowing m to vary from zero to ten. We select the light sources with a greedy algorithm, adding a light source by selecting the directional sources that most improves our approximation.

Before we look at the approximation results, we examine the effect of image size and sampling error. There are 4 scales used in our experiments. They are 65, 249, 977, and 3871 uniform samples on the surface of a hemisphere from coarsest to finest.

We did the experiment to find the best directional sources at different image sizes. We want to find an image size that is as small as possible while maintaining accuracy. We conducted numerous experiments on different scenes and two results are reported in Table 4.1. From the tables, we can see 60×80 is the best image size in terms of processing time and accuracy. Therefore, we will use images with size 60×80 to do all the rest of the experiments.

To examine how sampling affects the final result, we conducted experiments on many scenes. First, we randomly pick 100 directions from the surface of a hemisphere and render an image with each single point source. Second, for each point, we pick points which have angles to this point linearly spaced from 1° to 3° and generate images rendered with these point sources. The reason we choose the angle from 1° to 3° is that the average angle between the nearest points from the 3871 uniform samples on the surface of a hemisphere is 2.2° . The figures illustrating the angle between the two point source images vs. the angle between the two point light sources are shown in Figure 4.2. We can see the error in degrees can be as much as

Image Size	Index of Directional Source Image									
40x30	16	71	84	160	226	21	97	138	130	172
80x60	160	21	226	5	16	71	97	84	67	172
160x120	160	71	16	226	21	5	97	172	67	78
320x240	160	71	226	16	21	97	5	172	140	84

(a)

Image Size	Index of Directional Source Image									
40x30	5	3	209	164	138	152	226	128	221	136
80x60	5	3	209	164	152	138	57	128	221	136
160x120	5	3	209	164	152	138	57	128	221	136
320x240	5	3	209	164	152	138	57	128	221	136

(b)

Table 4.1: Best directional source images at different image size. The scenes are from sample images.

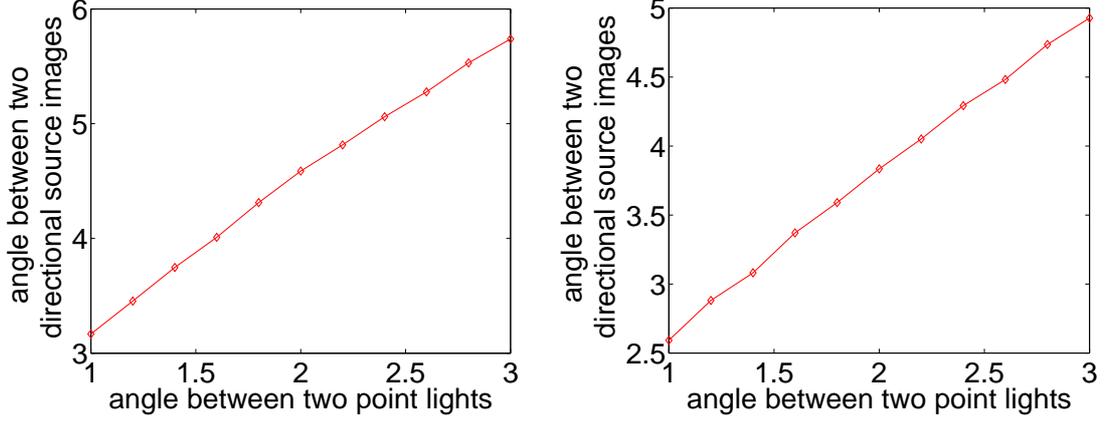


Figure 4.2: The angle between the two point source images vs. the angle between the two point light sources. The scenes are the same as in Figure 4.1.

2.8° considering that the angle between the real point source and the approximated point source can be as much as 1.1° at the worst case. So when we compare the experimental error with the prediction error, we have to take the sampling error which is around 2.8° into account.

We plot the approximation results in Figure 4.3. The horizontal axis shows, m , the number of directional sources used to approximate an image. The vertical axis shows the error in this approximation, in degrees. The red and magenta dashed lines show our prediction and the prediction compensated with sampling error about the maximum possible error in this approximation, respectively. The line with blue stars shows the actual error that we encounter for a different image. In general, the fewer directional sources that are used to approximate an image, the more error there will be. The actual error is slightly bigger than the prediction error, but less than the prediction error taking the sampling error into account. Figure 4.3-top and down show the image with buildings and teacup rendered with 3 and 8

directional sources, respectively. In the figure, we can see the predicted error in this approximation provides a good upper bound on the actual error we see. While we only show results for two scenes, the predicted maximum possible error bounded the true approximation error for all scenes and lighting conditions with which we experimented.

4.2 Illumination Recovery

We estimate the illumination by adding 9D spherical harmonic lighting with a small number of directional sources. Figure 4.4 shows the improvement of the accuracy obtained by increasing the number of directional sources. The vertical axis represents the error in degrees compared with the image rendered with the approximated lighting. The horizontal axis represents the number of directional sources used for estimation. The small picture right next to the plot shows error distributions in the synthesized image. Darker and whiter color represents larger error in a pixel value. The image on the right corner shows the illumination distribution from which it is approximated. We see that the error keeps decreasing as we add more directional sources. In the figure, we can see that the error is cut by 26% when we add 20 directional sources.

4.3 Performance Evaluation

We evaluate the algorithms we mentioned in the previous section in terms of their accuracy and computational cost. All the algorithms were run in MATLAB

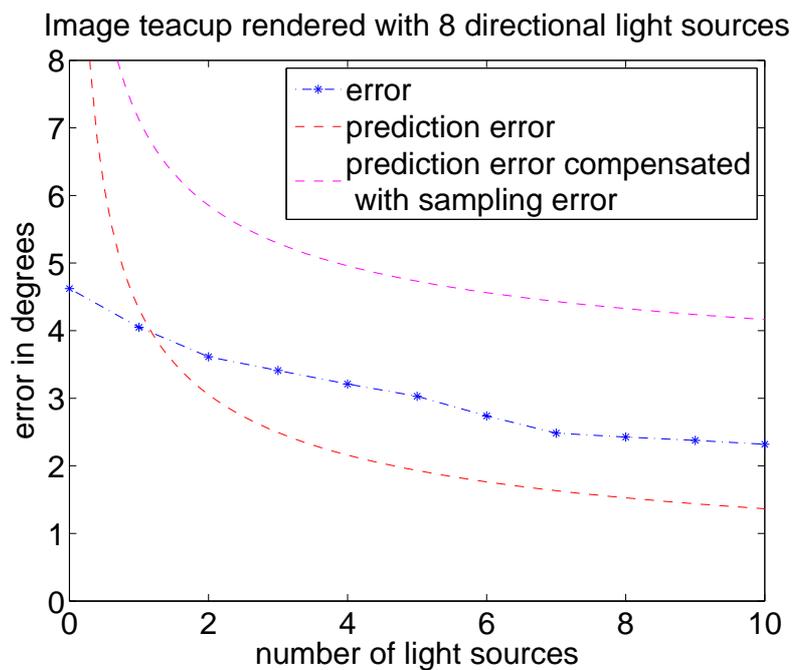
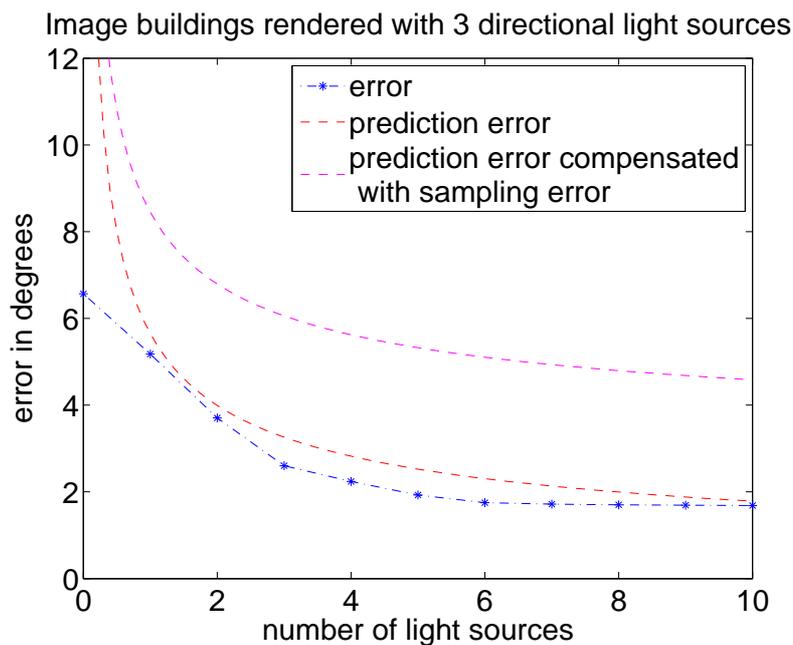


Figure 4.3: Each blue star shows the error that occurs in approximating a different image with linear subspace and a set of directional sources. The red and magenta dashed lines show our prediction and the prediction compensated with sampling error about the maximum possible error in this approximation, respectively.

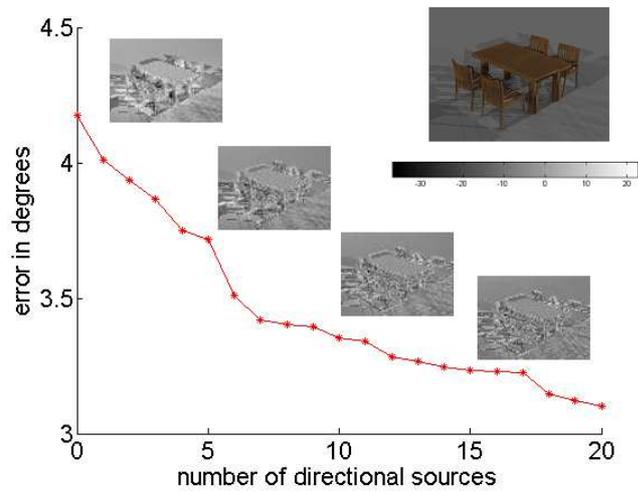
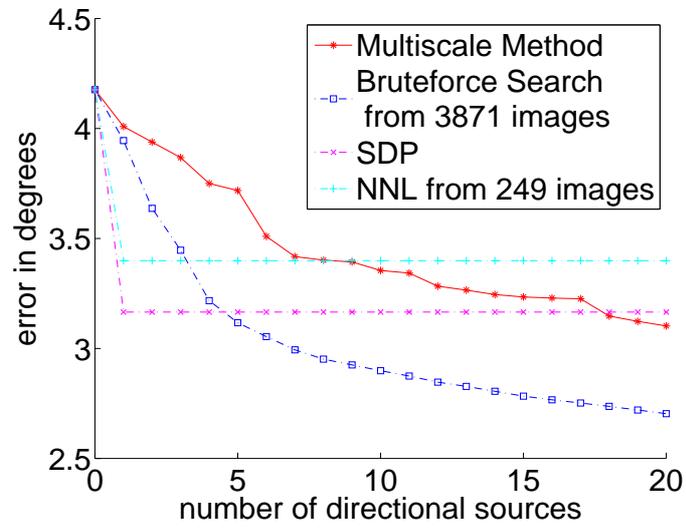


Figure 4.4: The improvement of the accuracy by adding directional sources. From the plot in the figure, we can clearly see that the accuracy improves gradually as we increase the number of the directional sources. The image is rendered with 8 directional sources.

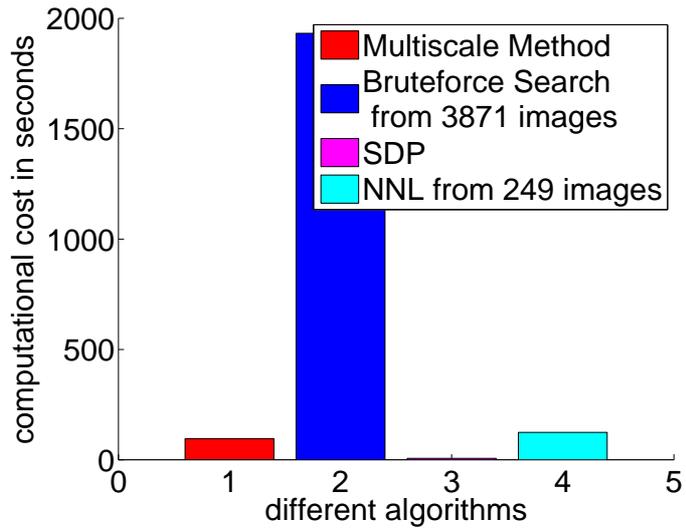
7.1. The computer used was a 2.00GHz Pentium 4 with 1.5GB RAM.

Figure 4.5 evaluates different algorithms to approximate the image with cast shadows in terms of accuracy and computational cost. In Figure 4.5 (a), it shows the approximation error in degrees vs. the number of directional light sources and (b) shows the computational cost. Not to our surprise, brute-force search achieves the best approximation results, but it requires a huge amount of processing time. We only test NNL using 249 images due to the unacceptable processing time if we go to higher level which has 977 images. It doesn't have good results due to the small number of directional source images we test. SDP works well on this case. It requires the least processing time while achieving good results. The multiscale algorithm is the runner-up in both processing time and accuracy. It's difficult to tell which is a better choice between SDP and the multiscale algorithm in this case.

Next we run a similar test using nine environment maps provided by [18], from high dynamic range light probe images. Again, we measure the error and processing time that occur in approximating images rendered with these environment maps using a small number of directional sources. Again, the multiscale algorithm obtains good results in terms of accuracy and processing time in Figure 4.6. While we only show results for one scene, the multiscale algorithm shows good results for all scenes and lighting conditions with which we experimented.

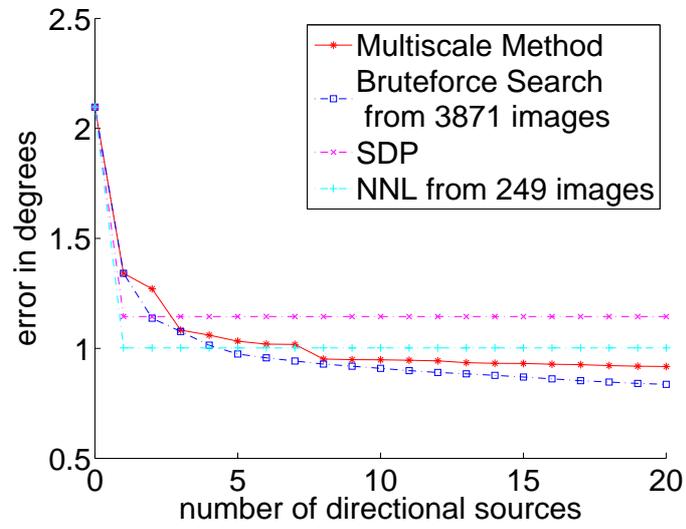


(a)

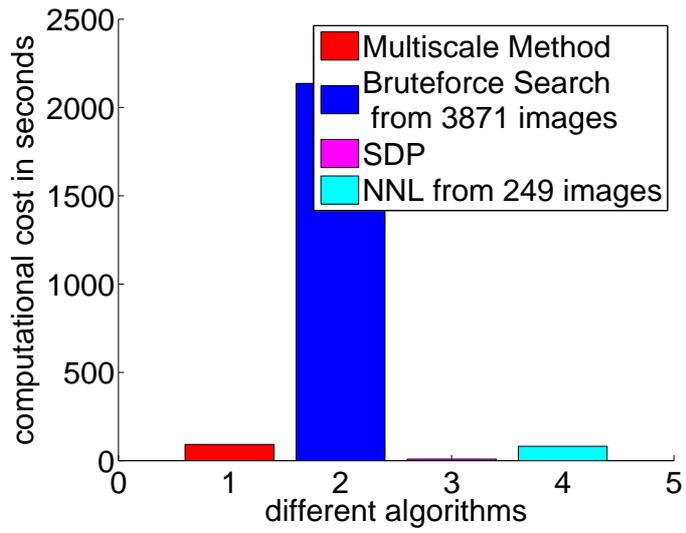


(b)

Figure 4.5: Comparison of different methods to approximate the image with cast shadows. The scene is rendered with 8 directional sources.



(a)



(b)

Figure 4.6: Comparison of different methods to approximate the image with cast shadows. The scene is rendered with an environment map.

Chapter 5

Conclusions and Future Work

In this thesis, we have presented a new method for estimating the illumination distribution from images with cast shadows using directional source images. By using images rendered with single directional source, we can approximate the illumination distribution for the images having significant cast shadows. Moreover, we compared our proposed algorithm with some other algorithms to show superiority of our algorithm.

We conducted experiments on synthesized data and showed the effectiveness and robustness of our proposed algorithm. For future work, more experiments with real data have to be conducted.

Bibliography

- [1] A. Sashua, *On Photometric Issues in 3d Visual Recognition From a Single 2d Image* (International Journal of Computer Vision 21, 1-2, 99-122, 1997).
- [2] Y. Moses, *Face Recognition: Generalization to Novel Images* (PhD thesis, Weizmann Institute of Science).
- [3] J. Nimeroff, E. Simoncelli and J. Dorsey, *Efficient Re-rendering of Naturally Illuminated Environments* (5th Eurographics Workshop on Rendering, 1994).
- [4] P. Belhumeur and D. Kriegman, *What is the Set of Images of an Object Under All Possible Lighting Conditions?* (International Journal of Computer Vision 28, 3, 245-260, 1998).
- [5] P. Hallinan, *A Low-dimensional Representation of Human Faces for Arbitrary Lighting Conditions* (IEEE Conference on Computer Vision and Pattern Recognition, 995-999, 1994).
- [6] R. Basri and D. Jacobs, *Lambertian Reflectances and Linear Subspaces* (IEEE Transaction on Pattern Analysis and Machine Intelligence 25, 2, 218-233, 2003).
- [7] R. Ramamoorthi and P. Hanrahan, *On the relationship between radiance and irradiance: determining the illumination from images of a convex Lambertian object* (Journal of the Optical Society of America, A 18, 10, 2448-2459, 2001).
- [8] R. Ramamoorthi and P. Hanrahan, *Analytic PCA Construction for Theoretical Analysis of Lighting Variability in Images of a Lambertian Object* (IEEE Transaction on Pattern Analysis and Machine Intelligence 24, 10, 1322-1333, 2002).
- [9] R. Ramamoorthi, M. Koudelka and P. Belhumeur, *A Fourier Theory for Cast Shadows* (IEEE Transaction on Pattern Analysis and Machine Intelligence 27, 2, 288-295, 2005).
- [10] I. Sato, Y. Sato and K. Ikeuchi, *Illumination from Shadows* (IEEE Transaction on Pattern Analysis and Machine Intelligence 25, 3, 290-300, 2003).
- [11] S. Geman and D. Geman, *Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images* (IEEE Transaction on Pattern Analysis and Machine Intelligence 6, 6, 721-741, 1984).

- [12] S. Shirdhonkar and D. Jacobs, *Non-Negative Lighting and Specular Object Recognition* (Tenth IEEE International Conference on Computer Vision (ICCV'05)) 2,1323-1330, 2005).
- [13] T. Kim, Y. Seo and K. Hong, *Improving AR using shadows arising from natural illumination distribution in video sequences* , (Eighth IEEE International Conference on Computer Vision (ICCV'01) 2, 329-334, 2001).
- [14] K. Nishino, Z. Zhang and K. Ikeuchi, *Determining reflectance parameters and illumination distribution from a sparse set of images for view-dependent image synthesis* , (Eighth IEEE International Conference on Computer Vision (ICCV'01) 1, 599-606, 2001).
- [15] Y. Zhang and Y. Yang, *Illuminant direction determination for multiple light sources* (IEEE Conference on Computer Vision and Pattern Recognition, 1, 269-276, 2000).
- [16] B.K.P. Horn, *Robot Vision*, (Cambridge, Mass.:MIT Press, 1986).
- [17] K. Hara, K. Nishino and K. Ikeuchi, *Determining reflectance and light position from a single image without distant illumination assumption* , (Ninth IEEE International Conference on Computer Vision (ICCV'03) 1, 560-567, 2003).
- [18] P. Debevec, *Rendering Synthetic Objects into Real Scenes: Bridging Traditional and Image-Based Graphics with Global Illumination and High Dynamic Range Photography* , (Proceedings of SIGGRAPH, 189-198, 1998).