Lambertian Reflectance and Linear Subspaces

Ronen Basri, *Member*, *IEEE*, and David W. Jacobs, *Member*, *IEEE*

Abstract—We prove that the set of all Lambertian reflectance functions (the mapping from surface normals to intensities) obtained with arbitrary distant light sources lies close to a 9D linear subspace. This implies that, in general, the set of images of a convex Lambertian object obtained under a wide variety of lighting conditions can be approximated accurately by a low-dimensional linear subspace, explaining prior empirical results. We also provide a simple analytic characterization of this linear space. We obtain these results by representing lighting using spherical harmonics and describing the effects of Lambertian materials as the analog of a convolution. These results allow us to construct algorithms for object recognition based on linear methods as well as algorithms that use convex optimization to enforce nonnegative lighting functions. We also show a simple way to enforce nonnegative lighting when the images of an object lie near a 4D linear space. We apply these algorithms to perform face recognition by finding the 3D model that best matches a 2D query image.

Index Terms—Face recognition, illumination, Lambertian, linear subspaces, object recognition, specular, spherical harmonics.

1 Introduction

Variability in lighting has a large effect on the appearance of objects in images, as is illustrated in Fig. 1. But we show in this paper that the set of images an object produces under different lighting conditions can, in some cases, be simply characterized as a nine dimensional subspace in the space of all possible images. This characterization can be used to construct efficient recognition algorithms that handle lighting variations.

Under normal conditions, light coming from all directions illuminates an object. When the sources of light are distant from an object, we may describe the lighting conditions by specifying the intensity of light as a function of its direction. Light, then, can be thought of as a nonnegative function on the surface of a sphere. This allows us to represent scenes in which light comes from multiple sources, such as a room with a few lamps and, also, to represent light that comes from extended sources, such as light from the sky, or light reflected off a wall.

Our analysis begins by representing these lighting functions using *spherical harmonics*. This is analogous to Fourier analysis, but on the surface of the sphere. With this representation, low-frequency light, for example, means light whose intensity varies slowly as a function of direction. To model the way diffuse surfaces turn light into an image, we look at the amount of light reflected as a function of the surface normal (assuming unit albedo), for each lighting condition. We show that these reflectance functions are produced through the analog of a convolution of the lighting function using a kernel that represents Lambert's reflection. This kernel acts as a low-pass filter with 99.2 percent of its energy in the first nine components, the zero, first, and second order harmonics. (This part of our analysis was derived independently also by Ramamoorthi and Hanrahan [31].) We use this and the nonnegativity of light to prove that under *any* lighting conditions, a nine-dimensional linear subspace, for example, accounts for at least 98 percent of the variability in the reflectance function. This suggests that in general the set of images of a convex, Lambertian object can be approximated accurately by a low-dimensional linear subspace. We further show how to analytically derive this subspace from a model of an object that includes 3D structure and albedo.

To provide some intuition about these results, consider the example shown in Fig. 2. The figure shows a white sphere made of diffuse material, illuminated by three distant lights. The lighting function can be described in this case as the sum of three delta functions. The image of the sphere, however, is smoothly shaded. If we look at a cross-section of the reflectance function, describing how the sphere reflects light, we can see that it is a very smoothed version of three delta functions. The diffuse material acts as a filter, so that the reflected light varies much more slowly than the incoming light.

Our results help to explain recent *experimental* work (e.g., Epstein et al. [10], Hallinan [15], Yuille et al. [40]) that has indicated that the set of images produced by an object under a wide range of lighting conditions lies near a low dimensional linear subspace in the space of all possible images. Our results also allow us to better understand several existing recognition methods. For example, previous work showed that, if we restrict every point on the surface of a diffuse object to face every light source (that is, ignoring attached shadows), then the set of images of the object lies in a 3D linear space (e.g., Shashua [34] and Moses [26]). Our analysis shows that, in fact, this approach uses the linear space spanned by the three first order harmonics, but omits the significant zeroth order (DC) component. Koenderink and van Doorn [21] augmented this space in order to account for an additional, perfect diffuse component. The additional component in their method is the missing DC component.

Our analysis also leads us to new methods of recognizing objects with unknown pose and lighting conditions. In particular, we discuss how the harmonic basis, which is derived analytically from a model of an object, can be used in a linear subspace-based object recognition algorithm, in place

- *R. Basri is with the Department of Computer Science, The Weizmann Institute of Science, Rehovot, 76100 Israel. E-mail: ronen.basri@weizmann.ac.il.*
- *D.W. Jacobs is with NEC Research Institute, 4 Independence Way, Princeton, NJ 08540. E-mail: dwj@research.nj.nec.com.*

Manuscript received 19 June 2001; revised 31 Dec. 2001; accepted 30 Apr. 2002.
Recommended for acceptance by P. Belhumeur.
For information on obtaining reprints of this article, please send e-mail to: tpami@computer.org, and reference IEEECS Log Number 114379.

Fig. 1. The same face, under two different lighting conditions.

of a basis derived by performing SVD on large collections of rendered images. Furthermore, we show how we can enforce the constraint that light is nonnegative everywhere by projecting this constraint to the space spanned by the harmonic basis. With this constraint recognition is expressed as a nonnegative least-squares problem that can be solved using convex optimization. This leads to an algorithm for recognizing objects under varying pose and illumination that resembles Georghiades et al. [12], but works in a low-dimensional space that is derived analytically from a model. The use of the harmonic basis, in this case, allows us to rapidly produce a representation to the images of an object in poses determined at runtime. Finally, we discuss the case in which a first order approximation provides an adequate approximation to the images of an object. The set of images then lies near a 4D linear subspace. In this case, we can express the nonnegative lighting constraint analytically. We use this expression to perform recognition in a particularly efficient way, without complex, iterative optimization techniques.

The paper is divided as follows: Section 2 briefly reviews the relevant studies. Section 3 presents our analysis of Lambertian reflectance. Section 4 uses this analysis to derive new algorithms for object recognition. Finally, Section 5 discusses extensions to specular reflectance.

2 Past Approaches

Our work is related to a number of recent approaches to object recognition that represent the set of images that an object can produce using low-dimensional linear subspaces of the space of all images. Ullman and Basri [38] analytically derive such a representation for sets of 3D points undergoing scaled orthographic projection. Shashua [34] and Moses [26] (see also Nayar and Murase [28] and Zhao and Yang [41]) derive a 3D linear representation of the set of images produced by a Lambertian object as lighting changes, but ignoring shadows. Hayakawa [16] uses factorization to build 3D models using this linear representation. Koenderink and van Doorn [21] extend this to a 4D space by allowing the light to include a diffuse component. Our work differs from these in that our representation accounts for *attached* shadows. These shadows occur when a surface faces away from a light source. We do not account for *cast* shadows, which 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. As is mentioned in Section 1, we show below that the 4D space used by Koenderink and van Doorn is in fact the space obtained by a first order harmonic approximation of the images of the object. The 3D space used by Shashua, Moses, and Hayakawa is the same space, but it lacks the significant DC component.

Researchers have collected large sets of images and performed PCA to build representations that capture within class variations (e.g., Kirby and Sirovich [19], Turk and Pentland [37], and Cootes et al. [7]) and variations due to pose and lighting (Murase and Nayar [27], Hallinan [15], Belhumeur et al. [3], and Yuille et al. [40]; see also Malzbender et al. [24]). This approach and its variations have been extremely popular in the last decade, particularly in applications to face recognition. Hallinan [15], Epstein et al. [10], and Yuille et al. [40] perform experiments that show that large numbers of images of real, Lambertian objects, taken with varied lighting conditions, do lie near a low-dimensional linear space, justifying this representation.

Belhumeur and Kriegman [4] have shown that the set of images of an object under arbitrary illumination forms a convex cone in the space of all possible images. This analysis accounts for attached shadows. In addition, for convex, Lambertian objects, they have shown that this cone (called the *illumination cone*) may have unbounded dimension. They have further shown how to construct the cone from as few as three images. Georghiades et al. [11], [12] use this representation for object recognition. To simplify the representation (an accurate representation of the illumination cone requires all the images that can be obtained with a single directional source), they further projected the images to low-dimensional subspaces obtained by rendering the objects and applying PCA to the rendered images. Our analysis allows us to further simplify this process by using instead the harmonic basis, which is derived analytically from a model of the object. This leads to a significant speed up of the recognition process (see Section 4).

Spherical harmonics have been used in graphics to efficiently represent the bidirectional reflection distribution function (BRDF) of different materials by, e.g., Cabral et al. [6] and Westin et al. [39]. Koenderink and van Doorn [20] proposed replacing the spherical harmonics basis with a basis for functions on the half-sphere that is derived from the Zernike polynomials, since BRDFs are defined over a half sphere. Nimeroff et al. [29], Dobashi et al. [8], and Teo et al.

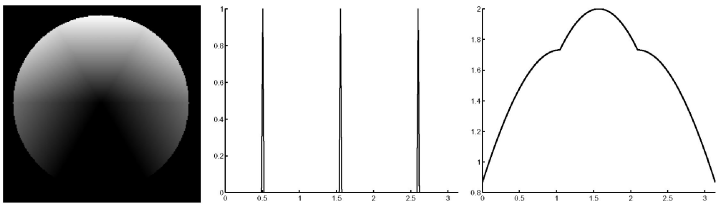

Fig. 2. On the left, a white sphere illuminated by three directional (distant point) sources of light. All the lights are parallel to the image plane, one source illuminates the sphere from above and the two others illuminate the sphere from diagonal directions. In the middle, a cross-section of the lighting function with three peaks corresponding to the three light sources. On the right, a cross-section indicating how the sphere reflects light. We will make precise the intuition that the material acts as a low-pass filtering, smoothing the light as it reflects it.

[35] explore specific lighting configurations (e.g., daylight) that can be represented efficiently as a linear combination of basis lightings. Dobashi et al. [8], in particular, use spherical harmonics to form such a basis.

Miller and Hoffman [25] were first to describe the process of turning incoming light into reflection as a convolution. D'Zmura [9] describes this process in terms of spherical harmonics. With this representation, after truncating high order components, the reflection process can be written as a linear transformation and, so, the low-order components of the lighting can be recovered by inverting the transformation. He used this analysis to explore ambiguities in lighting. We extend this work by deriving subspace results for the reflectance function, providing analytic descriptions of the basis images, and constructing new recognition algorithms that use this analysis while enforcing nonnegative lighting.

Independent of and contemporaneous with our work, Ramamoorthi and Hanrahan [31], [32], [33] have described the effect of Lambertian reflectance as a convolution and analyzed it in terms of spherical harmonics. Like D'Zmura, they use this analysis to explore the problem of recovering lighting from reflectances. Both the work of Ramamoorthi and Hanrahan and ours (first described in [1]) show that Lambertian reflectance acts as a low-pass filter with most of the energy in the first nine components. In addition to this, we show that the space spanned by the first nine harmonics accurately approximates the reflectance function under *any* light configuration, even when the light is dominated by high frequencies. Furthermore, we show how to use this space for object recognition.

Since the first introduction of our work, a number of related papers have further used and extended these ideas in a number of directions. Specifically, Ramamoorthi [30] analyzed the relationship between the principal components of the images produced by an object and the first nine harmonics. Lee et al. [23] constructed approximations to this space using physically realizable lighting. Basri and Jacobs [2] used the harmonic formulation to construct algorithms for photometric stereo under unknown, arbitrary lighting. Finally, Thornber and Jacobs [36] and Ramamoorthi and Hanrahan [32] further examined the effect of specularity and cast shadows.

3 Modeling Image Formation

In this section, we construct an analytically derived representation of the images produced by a convex, Lambertian object illuminated by distant light sources. We restrict ourselves to convex objects, so we can ignore the effect of shadows cast by one part of the object on another part of it. We assume that the surface of the object reflects light according to Lambert's law [22], which states that materials absorb light and reflect it uniformly in all directions. The only parameter of this model is the *albedo* at each point on the object, which describes the fraction of the light reflected at that point. This relatively simple model applies to diffuse (nonshiny) materials. It has been analyzed and used effectively in a number of vision applications.

By a "distant" light source we mean that it is valid to make the approximation that a light shines on each point in the scene from the same angle, and with the same intensity (this also rules out, for example, slide projectors). Lighting, however, may come from multiple sources, including diffuse sources such as the sky. We can therefore describe the intensity of the light as a single function of its direction that does not depend on the position in the scene. It is important to note that our analysis accounts for *attached shadows*, which occur when a point in the scene faces away from a light source.

While we are interested in understanding the images created by an object, we simplify this problem by breaking it into two parts. We use an intermediate representation, the *reflectance function* (also called the *reflectance map*, see Horn [17, chapters 10, 11]). Given our assumptions, the amount of light reflected by a white surface patch (a patch with albedo of one) depends on the surface normal at that point, but not on its spatial position. For a specific lighting condition, the reflectance function describes how much light is reflected by each surface normal. In the first part of our analysis, we consider the set of possible reflectance functions produced under different illumination conditions. This analysis is independent of the structure of the particular object we are looking at; it depends only on lighting conditions and the properties of Lambertian reflectance. Then, we discuss the relationship between the reflectance function and the image. This depends on object structure and albedo, but not on lighting, except as it determines the reflectance function. We begin by discussing the relation of lighting and reflectance.

Before we proceed, we would like to clarify the relation between the reflectance function and the bidirectional reflection distribution function (BRDF). The BRDF of a surface material is a function that describes the ratio of radiance, the amount of light reflected by the surface in every direction (measured in power per unit area per solid angle), to irradiance, the amount of light falling on the surface in every direction (measured in power per unit area). BRDF is commonly specified in a *local* coordinate frame, in which the surface normal is fixed at the north pole. The BRDF of a Lambertian surface is constant, since such a surface reflects light equally in all direction, and it is equal to $1/\pi$. In contrast, the reflectance function describes the radiance of a unit surface area given the *entire* distribution of light in the scene. The reflectance function is obtained by integrating the BRDF over all directions of incident light, weighting the intensity of the light by the foreshortening of the surface as seen from each light source. In addition, the reflectance function is specified in a *global*, *viewer centered* coordinate frame in which the viewing direction is fixed at the north pole. For example, if a scene is illuminated by a single directional source (a distant point source) of unit intensity, the reflectance function for every surface normal will contain the appropriate foreshortening of the surface with respect to the light source direction scaled by $1/\pi$. (For surface normals that face away from the light source the reflectance function will vanish.) For simplicity, we omit below the extra factor of $1/\pi$ that arises from the Lambertian BRDF since it only scales the intensities in the image by a constant factor.

3.1 Image Formation as the Analog of a Convolution

Both lighting and reflectance can be described as functions on the surface of the sphere. We describe the intensity of light as a function of its direction. This formulation allows us to consider multiple light sources that illuminate an object simultaneously from many directions. We describe reflectance as a function of the direction of the surface normal. To begin, we introduce notation for describing such functions.

Let S^2 denote the surface of a unit sphere centered at the origin. We will use u, v to denote unit vectors. We denote their Cartesian coordinates as (x, y, z), with $x^2 + y^2 + z^2 = 1$. When appropriate, we will denote such vectors by a pair of angles, (θ, ϕ), with

$$u = (x, y, z) = (\cos\phi \sin\theta, \sin\phi \sin\theta, \cos\theta), \tag{1}$$

where $0 \leq \theta \leq \pi$ and $0 \leq \phi \leq 2\pi$. In this coordinate frame, the poles are set at $(0, 0, \pm 1)$, θ denotes the angle between u and $(0, 0, 1)$, and it varies with latitude, and ϕ varies with longitude. We will use (θ_l, ϕ_l) to denote a direction of light and (θ_r, ϕ_r) to denote a direction of reflectance, although we will drop this subscript when there is no ambiguity. Similarly, we may express the lighting or reflectance directions using unit vectors such as u_l or v_r. Since we assume that the sphere is illuminated by a distant set of lights all points are illuminated by identical lighting conditions. Consequently, the configuration of lights that illuminate the sphere can be expressed as a nonnegative function $\ell(\theta_l, \phi_l)$, giving the intensity of the light reaching the sphere from each direction (θ_l, ϕ_l). We may also write this as $\ell(u_l)$, describing lighting direction with a unit vector.

According to Lambert's law, if a light ray of intensity l and coming from the 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

$$i = l(u_l)\rho \max(u_l \cdot v_r, 0). \tag{2}$$

If we fix the lighting, and ignore ρ for now, then the reflected light is a function of the surface normal alone. We write this function as $r(\theta_r, \phi_r)$, or $r(v_r)$. If light reaches a point from a multitude of directions, then the light reflected by the point would be the sum of (or in the continuous case the integral over) the contribution for each direction. If we denote $k(u \cdot v) = \max(u \cdot v, 0)$, then we can write:

$$r(v_r) = \int_{S^2} k(u_l \cdot v_r)\ell(u_l)du_l, \tag{3}$$

where $\int_{S^2}$ denotes integration over the surface of the sphere.

Below, we will occasionally abuse notation and write $k(u)$ to denote the max of zero and the cosine of the angle between u and the north pole (that is, omitting v means that v is the north pole). We therefore call k the *half-cosine* function. We can also write $k(\theta)$, where θ is the latitude of u, since k only depends on the θ component of u. For any fixed v, as we vary u (as we do while integrating (3)), then $k(u \cdot v)$ computes the half cosine function centered around v instead of the north pole. That is, since v_r is fixed inside the integral, we can think of k as a function just of u, which gives the max of zero and the cosine of the angle between u and v_r.

Thus, intuitively, (3) is analogous to a convolution, in which we center a kernel (the half-cosine function defined by k), and integrate its product with a signal (ℓ). In fact, we will call this a convolution, and write

$$r(v_r) = k * \ell \stackrel{\text{def}}{=} \int_{S^2} k(u_l \cdot v_r)\ell(u_l)du_l. \tag{4}$$

Note that there is some subtlety here since we cannot, in general, speak of convolving a function on the surface of the sphere with an arbitrary kernel. This is because we have three degrees of freedom in how we position a convolution kernel on the surface of the sphere, but the output of the convolution should be a function on the surface of the sphere, which has only two degrees of freedom. However, since k is rotationally symmetric this ambiguity disappears. In fact, we have been careful to only define convolution for rotationally symmetric k.

3.2 Spherical Harmonics and the Funk-Hecke Theorem

Just as the Fourier basis is convenient for examining the results of convolutions in the plane, similar tools exist for understanding the results of the analog of convolutions on the sphere. We now introduce these tools, and use them to show that in producing reflectance, k acts as a low-pass filter.

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. We denote these functions by Y_{nm}, with $n = 0, 1, 2, \ldots$ and $-n \leq m \leq n$:

$$Y_{nm}(\theta, \phi) = \sqrt{\frac{(2n+1)}{4\pi}\frac{(n-|m|)!}{(n+|m|)!}}P_{n|m|}(\cos\theta)e^{im\phi}, \tag{5}$$

where P_{nm} are the *associated Legendre functions*, defined as

$$P_{nm}(z) = \frac{(1-z^2)^{m/2}}{2^n n!}\frac{d^{n+m}}{dz^{n+m}}\left(z^2 - 1\right)^n. \tag{6}$$

We say that Y_{nm} is an nth *order* harmonic.

In the course of this paper, it will sometimes be convenient to parameterize Y_{nm} as a function of space coordinates (x, y, z) rather than angles. The spherical harmonics, written $Y_{nm}(x, y, z)$, then become polynomials of degree n in (x, y, z). The first nine harmonics then become

$$\begin{aligned}
&Y_{00} = \tfrac{1}{\sqrt{4\pi}} && Y_{10} = \sqrt{\tfrac{3}{4\pi}}z \\
&Y^e_{11} = \sqrt{\tfrac{3}{4\pi}}x && Y^o_{11} = \sqrt{\tfrac{3}{4\pi}}y \\
&Y_{20} = \tfrac{1}{2}\sqrt{\tfrac{5}{4\pi}}(3z^2 - 1) && Y^e_{21} = 3\sqrt{\tfrac{5}{12\pi}}xz \\
&Y^o_{21} = 3\sqrt{\tfrac{5}{12\pi}}yz && Y^e_{22} = \tfrac{3}{2}\sqrt{\tfrac{5}{12\pi}}(x^2 - y^2) \\
&Y^o_{22} = 3\sqrt{\tfrac{5}{12\pi}}xy,
\end{aligned} \tag{7}$$

where the superscripts e and o denote the even and the odd components of the harmonics, respectively, (so $Y_{nm} = Y^e_{n|m|} \pm iY^o_{n|m|}$, according to the sign of m; in fact the even and odd versions of the harmonics are more convenient to use in practice since the reflectance function is real).

Because the spherical harmonics form an orthonormal basis, this means that 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,

$$f(u) = \sum_{n=0}^{\infty}\sum_{m=-n}^{n} f_{nm}Y_{nm}(u), \tag{8}$$

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

$$f_{nm} = \int_{S^2} f(u)Y^*_{nm}(u)du, \tag{9}$$

and $Y^*_{nm}(u)$ denotes the complex conjugate of $Y_{nm}(u)$.

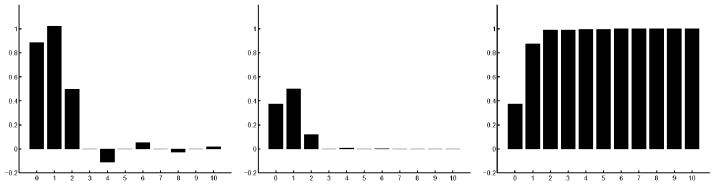

Fig. 3. From left to right: A graph representation of the first 11 coefficients of the Lambertian kernel, the relative energy captured by each of the coefficients, and the cumulative energy.

If we rotate a function f, this acts as a phase shift. Define for every n the nth order amplitude of f as

$$A_n \stackrel{\text{def}}{=} \sqrt{\frac{1}{2n+1}\sum_{m=-n}^{n} f_{nm}^2}. \tag{10}$$

Then, rotating f does not change the amplitude of a particular order. It may shuffle values of the coefficients, f_{nm}, for a particular order, but it does not shift energy between harmonics of different orders. For example, consider a delta function. As in the case of the Fourier transform, the harmonic transform of a delta function has equal amplitude in every order. If the delta function is at the north pole, its transform is nonzero only for the *zonal* harmonics, in which $m = 0$. If the delta function is, in general, position, it has some energy in all harmonics. But in either case, the nth order amplitude is the same for all n.

Both the lighting function, ℓ, and the Lambertian kernel, k, can be written as sums of spherical harmonics. Denote by

$$\ell = \sum_{n=0}^{\infty}\sum_{m=-n}^{n} l_{nm}Y_{nm}, \tag{11}$$

the harmonic expansion of ℓ, and by

$$k(u) = \sum_{n=0}^{\infty} k_n Y_{n0}. \tag{12}$$

Note that, because $k(u)$ is circularly symmetric about the north pole, only the zonal harmonics participate in this expansion, and

$$\int_{S^2} k(u)Y_{nm}^*(u)du = 0, \quad m \neq 0. \tag{13}$$

Spherical harmonics are useful in understanding the effect of convolution by k because of the Funk-Hecke theorem, which is analogous to the convolution theorem. Loosely speaking, the theorem states that we can expand ℓ and k in terms of spherical harmonics and, then, convolving them is equivalent to multiplication of the coefficients of this expansion.

We will state the Funk-Hecke theorem here in a form that is specialized to our specific concerns. Our treatment is based on Groemer [13], but Groemer presents a more general discussion in which, for example, the theorem is stated for spaces of arbitrary dimension.

Theorem 1 (Funk-Hecke). *Let $k(u \cdot v)$ be a bounded, integrable function on [-1,1]. Then:*

$$k * Y_{nm} = \alpha_n Y_{nm}$$

with

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

That is, the theorem states that the convolution of a (circularly symmetric) function k with a spherical harmonic Y_{mn} (as defined in (4)) results in the same harmonic, scaled by a scalar α_n. α_n depends on k and is tied directly to k_n, the nth order coefficient of the harmonic expansion of k.

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

$$r = k * \ell = \sum_{n=0}^{\infty}\sum_{m=-n}^{n}(\alpha_n l_{nm})Y_{nm}. \tag{14}$$

This is the chief implication of the Funk-Hecke theorem for our purposes.

3.3 Properties of the Convolution Kernel

The Funk-Hecke theorem implies that in producing the reflectance function, r, the amplitude of the light, ℓ, at every order n is scaled by a factor α_n that depends only on the convolution kernel, k. We can use this to infer analytically what frequencies will dominate r. To achieve this, we treat ℓ as a signal and k as a filter, and ask how the amplitudes of ℓ change as it passes through the filter.

The harmonic expansion of the Lambertian kernel (12) can be derived (with some tedious manipulation detailed in Appendix A) yielding

$$k_n = \begin{cases} \frac{\sqrt{\pi}}{2} & n = 0 \\ \sqrt{\frac{\pi}{3}} & n = 1 \\ (-1)^{\frac{n}{2}+1}\frac{\sqrt{(2n+1)\pi}}{2^n(n-1)(n+2)}\binom{n}{\frac{n}{2}} & n \geq 2, \text{even} \\ 0 & n \geq 2, \text{odd}. \end{cases} \tag{15}$$

The first few coefficients, for example, are

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

($k_3 = k_5 = k_7 = 0$), $|k_n|$ approaches zero as $O(n^{-2})$. A graph representation of the coefficients is shown in Fig. 3.

The energy captured by every harmonic term is measured commonly by the square of its respective coefficient divided by the total squared energy of the transformed function. The total squared energy in the half cosine function is given by

$$\int_0^{2\pi}\int_0^{\pi} k^2(\theta)\sin\theta d\theta d\phi = 2\pi\int_0^{\frac{\pi}{2}}\cos^2\theta\sin\theta d\theta = \frac{2\pi}{3}. \tag{17}$$

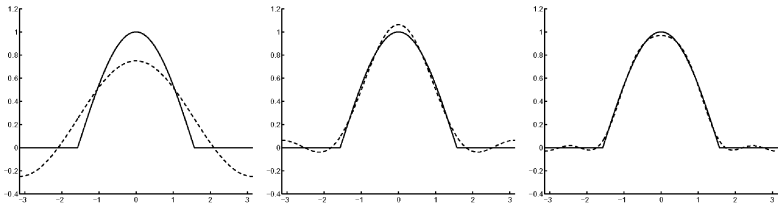

Fig. 4. A slice of the Lambertian kernel (solid) and its approximations (dashed) of first (left), second (middle), and fourth order (right).

(Here, we simplify our computation by integrating over θ and ϕ rather than u. The $\sin\theta$ factor is needed to account for the varying length of the latitude over the sphere.) Table 1 shows the relative energy captured by each of the first several coefficients. It can be seen that the kernel is dominated by the first three coefficients. Thus, a second order approximation already accounts for $(\frac{\pi}{4}+\frac{\pi}{3}+\frac{5\pi}{64})/\frac{2\pi}{3} \approx 99.22\%$ of the energy. With this approximation the half cosine function can be written as:

$$k(\theta) \approx \frac{3}{32} + \frac{1}{2}\cos\theta + \frac{15}{32}\cos^2\theta. \tag{18}$$

The quality of the approximation improves somewhat with the addition of the fourth order term (99.81%) and deteriorates to 87.5% when a first order approximation is used. Fig. 4 shows a 1D slice of the Lambertian kernel and its various approximations.

The maximum error in the second and fourth order approximations to k is, respectively, 0.0938 and 0.0586, while k varies between zero and one. This error occurs exactly at $\theta = \pi/2$, the point where k's first derivative is discontinuous and k becomes most difficult to model using low-frequency functions. The fact that k has a discontinuous first derivative, while its approximation does not, suggests that the two may be perceptually different. This is not a significant issue for the object recognition methods we discuss, in which we compare images using their sum of square differences. However, it may be more relevant to the use of these results in graphics.

3.4 Approximating the Reflectance Function

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. This implies that we can approximate the reflectance function that occurs under any lighting conditions using only low-order spherical harmonics. In this section, we show that this leads to an approximation that is always quite accurate.

We achieve a low-dimensional approximation to the reflectance function by truncating the sum in (14). That is, we have:

$$r = k * \ell = \sum_{n=0}^{\infty}\sum_{m=-n}^{n}(\alpha_n l_{nm})Y_{nm} \approx \sum_{n=0}^{N}\sum_{m=-n}^{n}(\alpha_n l_{nm})Y_{nm} \tag{19}$$

for some choice of order N. This means considering only the effects of the low order components of the lighting on the reflectance function. Intuitively, we know that since k_n and, therefore, α_n, is small for large n, this approximation should be good. However, the accuracy of the approximation also depends on l_{nm}, the harmonic expansion of the lighting.

To evaluate the quality of the approximation consider first, as an example, lighting, $\ell = \delta$, generated by a unit directional (distant point) source at the z direction ($\theta = \phi = 0$). In this case, the lighting is simply a delta function whose peak is at the north pole ($\theta = \phi = 0$). It can be readily shown that

$$r(v) = k * \delta = k(v). \tag{20}$$

If the sphere is illuminated by a single directional source in a direction other than the z direction the reflectance obtained would be identical to the kernel, but shifted in phase. Shifting the phase of a function distributes its energy between the harmonics of the same order n (varying m), but the overall energy in each n is maintained. The quality of the approximation, therefore, remains the same, but now for an Nth order approximation we need to use all the harmonics with $n \leq N$ for all m. Recall that there are $2n+1$ harmonics in every order n. Consequently, a first order approximation requires four harmonics. A second order approximation adds five more harmonics yielding a 9D space. The third order harmonics are eliminated by the kernel and, so, they do not need to be included. Finally, a fourth order approximation adds nine more harmonics yielding an 18D space.

We have seen that the energy captured by the first few coefficients k_i ($1 \leq i \leq N$) directly indicates the accuracy of the approximation of the reflectance function when the light consists of a single point source. Other light configurations may lead to different accuracy. Better approximations are obtained when the light includes enhanced diffuse components of low frequency. Worse approximations are anticipated if the light includes mainly high-frequency patterns.

However, even if the light includes mostly high-frequency patterns the accuracy of the approximation is still very high. This is a consequence of the nonnegativity of light. A lower bound on the accuracy of the approximation for *any* light function can be derived as follows.

First, note that Y_{00} is a constant function; this is just a DC component. The following derivation shows that for any nonnegative function the amplitude of the DC component must be at least as high as the amplitude of any of the other components. (The amplitude of a component is defined in (10).) The spherical harmonics satisfy the following useful identity. For every order n and every point u,

$$\sum_{m=-n}^{n}|Y_{nm}(u)|^2 = \frac{2n+1}{4\pi}. \tag{21}$$

(Compare this identity with (10), this identity implies, for example, that the delta function centered at u, for any u, has the same amplitude, $1/\sqrt{4\pi}$, at all orders.) From this identity,

it is straightforward to derive the following inequality. For every order n and for all pairs of points, u and v, on the sphere,

$$\left| \sum_{m=-n}^{n} Y_{nm}(u) Y_{nm}^{*}(v) \right| \leq \frac{2n+1}{4\pi}. \tag{22}$$

(Notice that this inequality, in fact, contains an inner product between the two $(2n+1$-vectors) $(Y_{n,-m}(u) \ldots Y_{nm}(u))$ and $(Y_{n,-m}(v) \ldots Y_{nm}(v))$ whose magnitudes are equal according to (21).)

Now, let $f(u) \geq 0$ be a real, nonnegative function. The squared amplitudes of f are given by

$$A_n^2 = \frac{1}{2n+1} \sum_{m=-n}^{n} f_{nm}^2, \tag{23}$$

where f_{nm} are the harmonic coefficients of f defined in (9). Substituting (9), we obtain

$$A_n^2 = \frac{1}{2n+1} \sum_{m=-n}^{n} \int_{S^2} f(u) Y_{nm}^{*}(u) du \int_{S^2} f(v) Y_{nm}^{*}(v) dv. \tag{24}$$

Rearranging yields

$$A_n^2 = \int_{S^2} \int_{S^2} f(u) f(v) \left(\frac{1}{2n+1} \sum_{m=-n}^{n} Y_{nm}(u) Y_{nm}^{*}(v) \right) du dv. \tag{25}$$

Using (22),

$$A_n^2 \leq \frac{1}{4\pi} \int_{S^2} \int_{S^2} f(u) f(v) du dv. \tag{26}$$

On the other hand, since the zeroth order harmonic is constant, $Y_{00}(u) = 1/\sqrt{4\pi}$, then

$$A_0^2 = \frac{1}{4\pi} \int_{S^2} \int_{S^2} f(u) f(v) du dv. \tag{27}$$

Consequently, for all n,

$$A_n \leq A_0. \tag{28}$$

As a consequence, in an Nth order approximation the worst scenario is obtained when the amplitudes in all frequencies higher than N saturate to the same amplitude as the DC component, while the amplitude of orders $1 \leq n \leq N$ are set to zero. In this case, the relative squared energy becomes

$$\frac{k_0^2}{k_0^2 + \sum_{n=N+1}^{\infty} k_n^2}, \tag{29}$$

or, using (17),

$$\frac{k_0^2}{\frac{2\pi}{3} - \sum_{n=1}^{N} k_n^2}. \tag{30}$$

Table 1 shows the bound obtained for several different approximations. It can be seen that using a second order approximation (involving nine harmonics) the accuracy of the approximation for any light function exceeds 97.96 percent. With a fourth order approximation (involving 18 harmonics) the accuracy exceeds 99.48 percent. Note that the bound computed in (29) is not tight, since the case that all the higher order terms are saturated yields a function with negative values. Consequently, the worst-case accuracy may even be higher than the bound.

TABLE 1

n	0	1	2	4	6	8	10
Energy	37.5	50	11.72	0.59	0.12	0.04	0.02
Cum.	37.5	87.5	99.22	99.80	99.92	99.96	99.98
Bound	37.5	75	97.96	99.48	99.80	99.90	99.94

The top row shows the energy captured by the nth zonal harmonic for the Lambertian kernel, k, $(0 \leq n \leq 10)$. The middle row shows the cumulative energy up to order n. This energy represents the quality of the nth order approximation of $r(\theta, \phi)$ (measured in relative squared error). The bottom row shows a lower bound on the quality of this approximation due to the nonnegativity of the light. The $n = 3$, 5, 7, and 9 are omitted because they contribute no energy. Relative energies are given in percents.

3.5 Generating Harmonic Reflectances

Constructing a basis to the space that approximates the reflectance functions is straightforward—we can simply use the low-order harmonics as a basis (see (19)). However, in many cases, we will want a basis vector for the nm component of the reflectances to indicate the reflectance produced by a corresponding basis vector describing the lighting, Y_{nm}. This makes it easy for us to relate reflectances and lighting, which is important when we want to enforce the constraint that the reflectances arise from nonnegative lighting (see Sections 4.2 and 4.3 below). We call these reflectances *harmonic reflectances* and denote them by r_{nm}. Using the Funk-Hecke theorem, r_{nm} is given by

$$r_{nm} = k * Y_{nm} = \alpha_n Y_{nm}. \tag{31}$$

Then, following (19),

$$r = k * \ell \approx \sum_{n=0}^{N} \sum_{m=-n}^{n} l_{nm} r_{nm}. \tag{32}$$

The first few harmonic reflectances are given by

$$\begin{array}{lll} r_{00} = \pi Y_{00} & r_{1m} = \frac{2\pi}{3} Y_{1m} & r_{2m} = \frac{\pi}{4} Y_{2m} \\ r_{4m} = \frac{\pi}{24} Y_{4m} & r_{6m} = \frac{\pi}{64} Y_{6m} & r_{8m} = \frac{\pi}{128} Y_{8m} \end{array} \tag{33}$$

for $-n \leq m \leq n$ (and $r_{3m} = r_{5m} = r_{7m} = 0$).

3.6 From Reflectances to Images

Up to this point, we have analyzed the reflectance functions obtained by illuminating a unit albedo sphere by arbitrary light. Our objective is to use this analysis to efficiently represent the set of images of objects seen under varying illumination. An image of an object under certain illumination conditions can be constructed from the respective reflectance function in a simple way: each point of the object inherits its intensity from the point on the sphere whose normal is the same. This intensity is further scaled by its albedo.

We can write this explicitly, as follows: Let p_i denote the ith object point. Let n_i denote the surface normal at p_i, and let ρ_i denote the albedo of p_i. Let the illumination be expanded with the coefficients l_{nm} (11)). Then the image, I_i of p_i is:

$$I_i = \rho_i r(n_i), \tag{34}$$

Fig. 5. We show the first nine harmonic images for a model of a face. The top row contains the zeroth harmonic (left) and the three first order harmonic images (right). The second row shows the images derived from the second harmonics. Negative values are shown in black, positive values in white.

where

$$r(n_i) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} l_{nm} r_{nm}(n_i). \tag{35}$$

Then, any image is a linear combination of *harmonic images*, b_{nm}, of the form:

$$b_{nm}(p_i) = \rho_i r_{nm}(n_i) \tag{36}$$

with

$$I_i = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} l_{nm} b_{nm}(p_i). \tag{37}$$

Fig. 5 shows the first nine harmonic images derived from a 3D model of a face.

We would like to now discuss how the accuracy of our low-dimensional linear approximation to a model's images can be affected by the mapping from the reflectance function to images. We will make two points. In the worst case, this can make our approximation arbitrarily bad. But, in typical cases, it will not make our approximation less accurate.

There are two components to turning a reflectance function into an image. One is that there is a rearrangement in the x, y position of points. That is, a particular surface normal appears in one location on the unit sphere, and may appear in a completely different location in the image. This rearrangement has no effect on our approximation. We represent images in a linear subspace in which each coordinate represents the intensity of a pixel. The decision as to which pixel to represent with which coordinate is arbitrary and changing this decision by rearranging the mapping from (x, y) to a surface normal merely reorders the coordinates of the space.

The second and more significant difference between images and reflectance functions is that the distribution of surface normals in the image may be quite different from the distribution of surface normals on the sphere. The distribution of surface normals in the image are determined by the shape of the object and the viewing direction. Thus, for example, occlusion ensures that half the surface normals on the sphere will be facing away from the camera and will not produce any visible intensities. A polygonal surface will contain some surface normals, but not others. These normals may stretch over extended regions in the image. For smooth surfaces, the curvature at a point will determine the extent to which its surface normal contributes to the image. The distribution of surface normals in the image can have quite a significant effect on the quality of our approximation.

Albedo, in fact, has a similar effect. It may scale the intensity of each point in the image independently of the others. This scaling effectively associates a weight to every pixel that affects the comparison of an image with its approximation. Thus, a black point (zero albedo) has no effect on the accuracy of the approximation, and in general darker pixels affect the accuracy of the approximation less than brighter pixels do. Overall, these effects are captured by noticing that the extent to which the reflectance of each point on the unit sphere influences the image can range from zero to the entire image.

Note that, in this discussion, we are still ignoring other real world effects such as shadows cast by nonconvex objects and materials that deviate somewhat from Lambertian. Likewise, this discussion ignores the effect of noise. We address these issues experimentally later, at least for images of faces.

We will next give an example to show that in the worst case these effects can make our approximation arbitrarily bad. Consider the case of an object that is a sphere of constant albedo (this example is adapted from Belhumeur and Kriegman [4]). If the light is coming from a direction opposite the viewing direction, it will not illuminate any visible pixels. We can then shift the light slightly, so that it illuminates just one pixel on the boundary of the object; by varying the intensity of the light we can give this pixel any desired intensity. A series of lights can do this for every pixel on the rim of the sphere. If there are n such pixels, the set of images we get fully occupies the positive orthant of an n-dimensional space. Obviously, points in this space can be arbitrarily far from any 9D space. What is happening is that all the energy in the image is concentrated in those surface normals for which our approximation happens to be poor.

However, generally, things will not be so bad. In general, occlusion will render an arbitrary half of the normals on the unit sphere invisible. Albedo variations and curvature will emphasize some normals, and deemphasize others. But, in general, the normals whose reflectances are poorly approximated will not be emphasized more than any other reflectances, and we can expect our approximation of reflectances on the entire unit sphere to be about as good over those pixels that produce the intensities visible in the image.

4 Recognition

We have developed an analytic description of the linear subspace that lies near the set of images that an object can produce. We now show how to use this description to recognize objects. Although our method is suitable for general objects, we will give examples related to the problem of face recognition. We will assume that a database of models of 3D objects that includes their surface normals and albedos is available. Our approach will allow us to recognize the objects under complicated lightings that include combinations of point and extended sources. Below, we assume that the pose of the object is already known, but that its identity and lighting conditions are not. For example, we may wish to identify a face that is known to be facing the camera. Or, we may assume that either a human or an automatic system have identified features, such as the eyes and the tip of the nose, that allow us to determine pose for each face in the database, but that the database is too big to allow a human to select the best match.

Recognition proceeds by comparing a new query image to each model in turn. To compare to a model, we compute the distance between the query image and the nearest image that the model can produce. We present two classes of algorithms that vary in their representation of a model's images. The linear subspace can be used directly for recognition, or we can restrict ourselves to a subset of the linear subspace that corresponds to physically realizable lighting conditions.

We will stress the advantages we gain by having an *analytic* description of the subspace available, in contrast to previous methods in which PCA could be used to derive a subspace from a sample of an object's images. One advantage of an analytic description is that we know this provides an accurate representation of an object's possible images, not subject to the vagaries of a particular sample of images. A second advantage is efficiency; we can produce a description of this subspace much more rapidly than PCA would allow. The importance of this advantage will depend on the type of recognition problem that we tackle. In particular, we are interested in recognition problems in which the position of an object is not known in advance, but can be computed at runtime using feature correspondences. In this case, the linear subspace must also be computed at runtime, and the cost of doing this is important. Finally, we will show that when we use a 4D linear subspace, an analytic description of this subspace allows us to incorporate the constraint that the lighting be physically realizable in an especially simple and efficient way.

4.1 Linear Methods

The most straightforward way to use our prior results for recognition is to compare a novel image to the linear subspace of images that correspond to a model, as derived by our harmonic representation (D'Zmura [9] also makes this suggestion). To do this, we produce the harmonic basis images of each model, as described in Section 3.6. 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 $\|Ba - I\|$. B is $p \times r$, p is the number of points in the image, and r is the number of basis images used. As discussed above, nine is a natural value to use for r, but $r = 4$ provides greater efficiency, while $r = 18$ offers even better potential accuracy. Every column of B contains one harmonic image b_{nm}. These images form a basis for the linear subspace, though not an orthonormal one. So we apply a QR decomposition to B to obtain such a basis. We compute Q, a $p \times r$ matrix with orthonormal columns, and R, an $r \times r$ matrix so that $QR = B$ and Q^TQ is an $r \times r$ identity matrix. Then Q is an orthonormal basis for B, and Q^TQI is the projection of I into the space spanned by B. We can then compute the distance from the image, I, and the space spanned by B as $\|QQ^TI - I\|$. The cost of the QR decomposition is $O(pr^2)$, assuming $p >> r$.

A number of previous methods have performed recognition by measuring the distance from an image to a linear subspace that represents a model's possible images. Some prior methods have performed PCA on a sample of images to find a linear subspace representing an object (see also Belhumeur et al. [3] for discussion of a different linear projection using the Fisher linear discriminant). Hallinan [15] performed experiments indicating that PCA can produce a five- or six-dimensional subspace that accurately models a face. Epstein et al. [10] and Yuille et al. [40] describe experiments on a wider range of objects that indicate that images of Lambertian objects can be approximated by a linear subspace of between three and seven dimensions. Specifically, the set of images of a basketball were approximated to 94.4 percent by a 3D space and to 99.1 percent by a 7D space, while the images of a face were approximated to 90.2 percent by a 3D space and to 95.3 percent by a 7D space. Georghiades et al. [12] render the images of an object and find an 11D subspace that approximates these images.

These numbers are roughly comparable to the 9D space that, according to our analysis, approximates the images of a Lambertian object. Additionally, we note that the basis images of an object will not generally be orthogonal and can in some cases be quite similar. For example, if the z components of the surface normals of an object do not vary much, then some of the harmonic images will be quite similar, such as $b_{00} = \rho$ versus $b_{10} = \rho z$. This may cause some components to be less significant, so that a lower-dimensional approximation can be fairly accurate (see a recent analysis by Ramamoorthi [30]).

When s sampled images are used (typically $s >> r$), with $s << p$ PCA requires $O(ps^2)$. Also, in MATLAB, PCA of a thin, rectangular matrix seems to take exactly twice as long as its QR decomposition. Therefore, in practice, PCA on the matrix constructed by Georghiades et al. would take about 150 times as long as using our method to build a 9D linear approximation to a model's images (this is for $s = 100$ and $r = 9$. One might expect p to be about 10,000, but this does not affect the relative costs of the methods). This may not be too significant if pose is known ahead of time and this computation takes place offline. But when pose is computed at runtime, the advantages of our method can become very great.

It is also interesting to compare our method to another linear method, due to Shashua [34] and Moses [26]. Shashua points out that in the absence of attached shadows, every possible image of an object is a linear combination of the x, y, and z components of the surface normals, scaled by the albedo. He therefore proposes using these three components to produce a 3D linear subspace to represent a model's images. Notice that these three vectors are identical, up to a scale factor, to the basis images produced by the first order harmonics in our method.

While this equivalence is clear algebraically, we can also explain it as follows: The first order harmonic images are images of any object subjected to a lighting condition

described by a single harmonic. The Funk-Hecke theorem ensures that all components of the kernel describing the reflectance function will be irrelevant to this image except for the first order components. In Shashua's work, the basis images are generated by using a point source as the lighting function which contains all harmonics. But the kernel used is the full cosine function of the angle between the light and the surface normal. This kernel has components only in the first harmonic. So all other components of the lighting are irrelevant to the image. In either case, the basis images are due only to the first set of harmonics.

We can therefore interpret Shashua's method as also making an analytic approximation to a model's images, using low order harmonics. However, our previous analysis tells us that the images of the first harmonic account for only 50 percent of the energy passed by the half-cosine kernel. Furthermore, in the worst case it is possible for the lighting to contain *no* component in the first harmonic. Most notably, Shashua's method does not make use of the DC component of the images, i.e., of the zeroth harmonic. These are the images produced by a perfectly diffuse light source. Nonnegative lighting must always have a significant DC component. Koenderink and van Doorn [21] have suggested augmenting Shashua's method with this diffuse component. This results in a linear method that uses the four most significant harmonic basis images, although Koenderink and van Doorn propose this as apparently an heuristic suggestion, without analysis or reference to a harmonic representation of lighting.

Note finally that, as with our method, the basis proposed by Shashua to represent the set of images of the object, the three components of the surface normals scaled by the albedo, do not represent physically realizable images, since they may contain negative values. Shashua suggested replacing this basis with a basis composed of three images obtained under directional sources. This physically realizable basis is accurate only if pixels that face away from the light source are discarded.

4.2 Enforcing Nonnegative Light

When we take arbitrary linear combinations of the harmonic basis images, we may obtain images that are not physically realizable. This is because the corresponding linear combination of the harmonics representing lighting may contain negative values. That is, rendering these images may require negative "light," which of course is physically impossible. In this section, we show how to use the basis images while enforcing the constraint of nonnegative light. Belhumeur and Kriegman [4] have shown that the set of images of an object produced by nonnegative lighting is a convex cone in the space of all possible images. They call this the *illumination cone*. We show how to compute approximations to this cone in the space spanned by the harmonic basis images.

Specifically, given an image I, we attempt to minimize $\|Ba - I\|$ subject to the constraint that the light is nonnegative everywhere along the sphere. A straightforward method to enforce positive light is to infer the light from the images by inverting the convolution. This would yield linear constraints in the components of a, $Ha \geq 0$, where the columns of H contain the spherical harmonics Y_{nm}. Unfortunately, this naive method is problematic since the light may contain higher order terms that cannot be recovered from a low order approximation of the images of the object. In addition, the harmonic approximation of nonnegative light may at times include negative values. Forcing these values to be nonnegative will lead to an incorrect recovery of the light. Below we describe a different method in which we project the illumination cone onto the low-dimensional space and use this projection to enforce nonnegative lighting.

We first present a method that can use any number of harmonic basis images. A nonnegative lighting function can be written as a nonnegative combination of delta functions, each representing a directional source. Denote by $\delta_{\theta_0\phi_0}$ the function returning a nonzero value at (θ_0, ϕ_0), 0 elsewhere, and integrating to 1. This lighting function represents a distant point source at direction (θ_0, ϕ_0). To project the delta function onto the first few harmonics, we need to look at the harmonic transform of the delta function. Since the inner product of $\delta_{\theta_0\phi_0}$ with a function f returns simply $f(\theta_0, \phi_0)$, we can conclude that the harmonic transform of the delta function is given by

$$\delta_{\theta_0\phi_0} = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} Y_{nm}(\theta_0, \phi_0) Y_{nm}. \tag{38}$$

The projection of the delta function onto the first few harmonics is obtained by taking the sum only over the first few terms.

Suppose now that a nonnegative lighting function $\ell(\theta, \phi)$ is expressed as a nonnegative combination of delta functions

$$\ell = \sum_{j=1}^{J} a_j \delta_{\theta_j\phi_j}, \tag{39}$$

for some J. Obviously, due to the linearity of the harmonic transform, the transform of ℓ is a nonnegative combination of the transforms of the delta functions with the same coefficients. That is,

$$\ell = \sum_{j=1}^{J} a_j \sum_{n=0}^{\infty} \sum_{m=-n}^{n} Y_{nm}(\theta_j, \phi_j) Y_{nm}. \tag{40}$$

Likewise, the image of an object illuminated by ℓ can be expressed as a nonnegative combination as follows:

$$I = \sum_{j=1}^{J} a_j \sum_{n=0}^{\infty} \sum_{m=-n}^{n} Y_{nm}(\theta_j, \phi_j) b_{nm}, \tag{41}$$

where b_{nm} are the harmonic images defined in the previous section.

Given an image our objective is to recover the nonnegative coefficients a_j. Assume we consider an approximation of order N, and denote the number of harmonics required for spanning the space by $r = r(N)$ (e.g., if $N = 2$, then $r = 9$). In matrix notation, denote the harmonic functions by H, H is $s \times r$, where s is the number of sample points on the sphere. In practice, we use a uniform tesselation of the sphere with $s = 122$. The columns of H contain a sampling of the harmonic functions, while its rows contain the transform of the delta functions. Further, denote by B the basis images, B is $p \times r$, where p is the number of points in the image. Every column of B contains one harmonic image b_{nm}. Finally, denote $a = (a_1, \ldots, a_s)^T$. Then, our objective is to solve the nonnegative least-squares problem:

$$\min_a \|BH^T a - I\| \quad \text{s.t.} \quad a \geq 0. \tag{42}$$

Here, $H^T a$ is a $(1 \times r)$ vector that represents the r low-frequency harmonic components of the lighting. $BH^T a$, then gives an image of our object under this lighting. We can further project the image to the r-dimensional space spanned by the harmonic images and solve the optimization problem in this smaller space. To do so, we apply a QR decomposition to B, as described previously. We obtain:

$$\min_a \|RH^T a - Q^T I\| \quad \text{s.t.} \quad a \geq 0. \tag{43}$$

Now, R is $r \times r$ and $Q^T I$ is an r-vector. Once we have solved for a, we measure error as the distance between $BH^T a$ and I.

Note that this method is similar to that presented in Georghiades et al. [11]. The primary difference is that we work in a low-dimensional space constructed for each model using its harmonic basis images. Georghiades et al. perform a similar computation after projecting all images into a 100-dimensional space constructed using PCA on images rendered from models in a 10-model database. Also, we do not need to explicitly render images using a point source, and project them into a low-dimensional space. In our representation, the projection of these images is achieved simply through H^T.

4.3 Recognition with Four Harmonics

A further simplification can be obtained if the set of images of an object is approximated only up to first order. Four harmonics are required in this case. One is the DC component, representing the appearance of the object under uniform ambient light and three are the basis images also used by Shashua. Again, we attempt to minimize $\|Ba - I\|$ (now B is $p \times 4$) subject to the constraint that the light is nonnegative everywhere along the sphere.

As before, we determine the constraints by projecting the delta functions onto the space spanned by the first four harmonics. However, now this projection takes a particularly simple form. Consider a delta function $\delta_{\theta_0\phi_0}$. Its first order approximation is given by

$$\delta_{\theta_0\phi_0} \approx \sum_{n=0}^{1} \sum_{m=-n}^{n} Y_{nm}(\theta_0, \phi_0) Y_{nm}. \tag{44}$$

Using space coordinates this approximation becomes

$$\begin{aligned} &\delta_{\theta_0\phi_0}(x, y, z) \\ &\approx \frac{1}{4\pi} + \frac{3}{4\pi}(x \sin\theta_0 \cos\phi_0 + y \sin\theta_0 \sin\phi_0 + z \cos\theta_0). \end{aligned} \tag{45}$$

Let

$$\ell \approx a_0 + a_1 x + a_2 y + a_3 z \tag{46}$$

be the first order approximation of a nonnegative lighting function ℓ. ℓ is a nonnegative combination of delta functions. It can be readily verified that such a combination cannot decrease the zero order coefficient relative to the first order ones. Consequently, any nonnegative combination of delta functions must satisfy

$$9a_0^2 \geq a_1^2 + a_2^2 + a_3^2. \tag{47}$$

(Equality is obtained when the light is a single delta function, see (45).) Therefore, we can express the problem of recognizing an object with a 4D harmonic space as minimizing $\|Ba - I\|$ subject to (47).

In the four harmonic cases, the harmonic images are just the albedos and the components of the surface normals scaled by the albedos, each scaled by some factor. It is therefore natural to use those directly and hide the scaling coefficients within the constraints. Let I be an image of the object illuminated by ℓ, then, using (33) and (36),

$$I \approx \pi a_0 \rho + \frac{2\pi}{3}(a_1 \rho n_x + a_2 \rho n_y + a_3 \rho n_z), \tag{48}$$

where ρ and (n_x, n_y, n_z) are, respectively, the albedo and the surface normal of an object point. Using the unscaled basis images, ρ, ρn_x, ρn_y, and ρn_z, this equation can be written as:

$$I \approx b_0 \rho + b_1 \rho n_x + b_2 \rho n_y + b_3 \rho n_z, \tag{49}$$

with $b_0 = \pi a_0$ and $b_i = \frac{2\pi}{3} a_i$ $(1 \leq i \leq 3)$. Substituting for the a_is in (47), we obtain

$$\frac{9 b_0^2}{\pi^2} \geq \frac{9}{4\pi^2}(b_1^2 + b_2^2 + b_3^2), \tag{50}$$

which simplifies to

$$4 b_0^2 \geq b_1^2 + b_2^2 + b_3^2. \tag{51}$$

Consequently, to find the nearest image in the space spanned by the first four harmonic images with nonnegative light we may minimize the difference between the two sides of (49) subject to (51). We solve the problem of minimizing $\|Ba - I\|$ subject to this constraint by first performing the minimization ignoring the constraint. If the solution we obtain obeys the constraint, we are done. If not, we know that the minimum occurs when the constraint is satisfied at equality. This minimization problem has the general form:

$$\min_x \|Ax - b\| \quad \text{s.t.} \quad x^T B x = 0. \tag{52}$$

We show in Appendix B that by diagonalizing A and B simultaneously and introducing a Lagrange multiplier the problem can be solved by finding the roots of a sixth degree polynomial with a single variable, the Lagrange multiplier. With this manipulation, solving the minimization problem is straightforward.

4.4 Experiments

We have experimented with these recognition methods using a database of faces collected at NEC, Japan. The database contains models of 42 faces, each includes the 3D shape of the face (acquired using a laser scanner) and estimates of the albedos in the red, green and blue color channels. As query images, we use 42 images each of ten individuals, taken across seven different poses and six different lighting conditions (shown in Fig. 6). In our experiment, each of the query images is compared to each of the 42 models, and then the best matching model is selected.

In all methods, we first obtain a 3D alignment between the model and the image, using the algorithm of Blicher and Roy [5]. In brief, features on the faces were identified by hand and, then, a 3D rigid transformation was found to align the 3D features with the corresponding 2D image features.

In all methods, we only pay attention to image pixels that have been matched to some point in the 3D model of the face. We also ignore image pixels that are of maximum intensity,

Fig. 6. Test images used in the experiments.

since these may be saturated, and provide misleading values. Finally, we subsample both the model and the image, replacing each $m \times m$ square with its average values. Preliminary experiments indicate that we can subsample quite a bit without significantly reducing accuracy. In the experiments below, we ran all algorithms subsampling with 16×16 squares, while the original images were 640×480.

Our methods produce coefficients that tell us how to linearly combine the harmonic images to produce the rendered image. These coefficients were computed on the sampled image, but then applied to harmonic images of the full, unsampled image. This process was repeated separately for each color channel. Then, a model was compared to the image by taking the root mean squared error, derived from the distance between the rendered face model and all corresponding pixels in the image.

Fig. 7 shows the Receiver Operating Characteristic (ROC) curves for three recognition methods: the 9D linear method, and the methods that enforce positive lighting in 9D and 4D. The curves show the fraction of query images for which the correct model is classified among the top k, as k varies from 1 to 40. The 4D positive lighting method performs significantly less well than the others, getting the correct answer about 60 percent of the time. However, it is much faster, and seems to be quite effective under the simpler pose and lighting conditions. The 9D linear method and 9D positive lighting method each pick the correct model first 86 percent of the time. With this data set, the difference between these two algorithms is quite small compared to other sources of error. These may include limitations in our model for handling cast shadows and specularities, but also includes errors in the model building and pose determination processes. In fact, on examining our results we found that one pose (for one person) was grossly wrong because a human operator selected feature points in the wrong order. We eliminated the six images (under six lighting conditions) that used this pose from our results.

5 Future Work: Specularity

In general, it is a subject of future work to consider how this sort of analysis may be applied to more complex imaging situations that include specularities and cast shadows. However, in this section, we will make one basic remark about these situations. These issues are further analyzed in Thornber and Jacobs [36] and Ramamoorthi and Hanrahan [32].

We note that a low-dimensional set of images can also result when the lighting itself is low dimensional. This can occur when the lights are all diffuse, as when the sun is behind clouds or lighting is due to interreflections. In this case, the lighting itself may be well approximated by only low-order harmonics. If the lighting is a linear combination of a small number of harmonics, then images will be a linear combination of those produced when the scene is rendered separately by each of these harmonics. This low-dimensionality is due simply to the linearity of lighting, the fact that the sum of two images produced by any two lighting conditions will be the image produced by the sum of these lighting conditions. Therefore, this will be true under the most general imaging assumptions, including cast shadows and specularities.

We also note that with specular objects, the bidirectional reflection distribution function (BRDF) is generally much more sharply peaked than it is with the cosine function. This provides the intuition that specular objects will be more affected by high-order harmonic components of the lighting. In the extreme case of a mirror, the entire lighting function passes into the reflectance function, preserving all components of the lighting. Therefore, we expect that for specular objects, a low-order approximation to the image set will be less accurate. A representation in terms of harmonic images

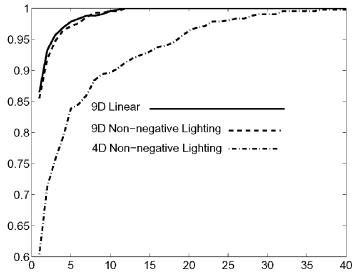

Fig. 7. ROC curves for our recognition methods.

may still provide a useful approximation, however. This is consistent with the experiments of Epstein et al. [10].

6 Conclusions

Lighting can be arbitrarily complex. But, in many cases, its effect is not. When objects are Lambertian, we show that a simple, nine-dimensional linear subspace can capture the set of images they produce. This explains prior empirical results. It also gives us a new and effective way of understanding the effects of Lambertian reflectance as that of a low-pass filter on lighting.

Moreover, we show that this 9D space can be directly computed from a model, as low-degree polynomial functions of its scaled surface normals. This description allows us to produce efficient recognition algorithms in which we know we are using an accurate approximation to the model's images. We can compare models to images in a 9D space that captures at least 98 percent of the energy of all the model's images. We can enforce the constraint that lighting be positive by performing a nonnegative least-squares optimization in this 9D space. Or, if we are willing to settle for a less accurate approximation, we can compute the positive lighting that best matches a model to an image by just solving a six-degree polynomial in one variable. We evaluate the effectiveness of all these algorithms using a database of models and images of real faces.

Appendix A

The Harmonic Transform of the Lambertian Kernel

In this appendix, we derive an analytic expression of the harmonic expansion of the Lambertian kernel, $k(\theta) = \max(\cos\theta, 0)$. The content of this appendix was derived independently also by Ramamoorthi and Hanrahan [31]. According to (12), this expansion is given by

$$k = \sum_{n=0}^{\infty} k_n Y_{n0},$$

with

$$k_n = 2\pi \int_0^{\pi} k(\theta) Y_{n0}(\theta) \sin\theta d\theta.$$

(Note that the zonal harmonics Y_{n0} depend only on θ.)

We next determine an explicit form for the coefficients k_n. First, we can limit the integration to the positive portion of the cosine function by integrating over θ only to $\pi/2$, that is,

$$k_n = 2\pi \int_0^{\frac{\pi}{2}} \cos\theta Y_{n0}(\theta) \sin\theta d\theta.$$

Now,

$$Y_{n0} = \sqrt{\frac{2n+1}{4\pi}} P_n(\cos\theta),$$

where $P_n(z)$ is the associated Legendre function of order n defined by

$$P_n(z) = \frac{1}{2^n n!} \frac{d^n}{dz^n} (z^2 - 1)^n.$$

Substituting $z = \cos\theta$, we obtain

$$k_n = \sqrt{(2n+1)\pi} \int_0^1 z P_n(z) dz.$$

We now turn to computing the integral

$$\int_0^1 z P_n(z) dz.$$

Integrating by parts yields,

$$\frac{1}{2^n n!} \left[z \frac{d^{n-1}}{dz^{n-1}} (z^2 - 1)^n \Big|_0^1 - \int_0^1 \frac{d^{n-1}}{dz^{n-1}} (z^2 - 1)^n dz \right].$$

The first term vanishes and we are left with

$$-\frac{1}{2^n n!} \int_0^1 \frac{d^{n-1}}{dz^{n-1}} (z^2 - 1)^n dz,$$

which, when integrated, yields

$$-\frac{1}{2^n n!} \frac{d^{n-2}}{dz^{n-2}} (z^2 - 1)^n \Big|_1^0.$$

This formula vanishes for $z = 1$ and, so, we obtain

$$\frac{1}{2^n n!} \frac{d^{n-2}}{dz^{n-2}} (z^2 - 1)^n \Big|_{z=0}.$$

Now,

$$(z^2 - 1)^n = \sum_{k=0}^{n} \binom{n}{k} (-1)^{n-k} z^{2k}.$$

When we take the $n-2$ derivative all terms whose exponent is less than $n-2$ disappear. Moreover, since we are evaluating the derivative at $z = 0$ all the terms whose exponent is larger than $n-2$ vanish. Thus, only the term whose exponent is $2k = n-2$ survives. Denote the respective coefficient by b_{n-2}, then, when n is odd $b_{n-2} = 0$, and when n is even

$$b_{n-2} = \binom{n}{\frac{n}{2} - 1} (-1)^{\frac{n}{2}+1}.$$

In this case,

$$\frac{d^{n-2}}{dz^{n-2}} (z^2 - 1)^n \Big|_{z=0} = (n-2)! b_{n-2},$$

and we obtain

$$\int_0^1 zP_n(z)dz = \frac{(-1)^{\frac{n}{2}+1}}{2^n(n-1)(n+2)}\binom{n}{\frac{n}{2}}.$$

The above derivation holds for $n \geq 2$. The special cases that $n = 0$ and $n = 1$ should be handled separately. In the first case, $P_0(z) = 1$ and, in the second case, $P_1(z) = z$. For $n = 0$, the integral becomes $\int_0^1 zdz = 1/2$, and for $n = 1$ it becomes $\int_0^1 z^2dz = 1/3$. Consequently,

$$k_n = \sqrt{(2n+1)\pi}\int_0^1 zP_n(z)dz = \begin{cases} \frac{\sqrt{\pi}}{2} & n = 0 \\ \sqrt{\frac{\pi}{3}} & n = 1 \\ (-1)^{\frac{n}{2}+1}\frac{\sqrt{(2n+1)\pi}}{2^n(n-1)(n+2)}\binom{n}{\frac{n}{2}} & n \geq 2, \text{ even} \\ 0 & n \geq 2, \text{ odd.} \end{cases}$$

Finally, we analyze the asymptotic behavior of k_n. For even $n \geq 2$,

$$|k_n| = \frac{\sqrt{(2n+1)\pi}}{2^n(n-1)(n+2)}\binom{n}{\frac{n}{2}}.$$

Using the Stirling equation, as $n \to \infty$, $n! \approx n^n e^{-n}\sqrt{2\pi n}$. This implies that $(n/2)! \approx n^{\frac{n}{2}}(2e)^{\frac{-n}{2}}\sqrt{\pi n}$, and we obtain,

$$\binom{n}{\frac{n}{2}} \approx \sqrt{\frac{2}{\pi n}}2^n.$$

Thus,

$$|k_n| = \frac{\sqrt{(2n+1)\pi}}{2^n(n-1)(n+2)}\binom{n}{\frac{n}{2}} \approx \frac{2}{(n-1)(n+2)} \approx O(n^{-2}).$$

APPENDIX B

RECOGNITION WITH FOUR HARMONICS

Finding the nearest image in the 4D harmonic space subject to the constraint that the light is nonnegative has the general form

$$\min_x \|Ax - b\| \quad \text{s.t.} \quad x^T Bx = 0,$$

with $A\,(n \times 4)$, $b\,(n \times 1)$, and $B\,(4 \times 4)$. In this representation, the columns of A contain the unscaled harmonic images, b is the image to be recognized and $B = \text{diag}(4, -1, -1, -1)$. (The method we present below, however, can be used with an arbitrary nonsingular matrix B.)

First, we can solve the linear system

$$\min_x \|Ax - b\|$$

and check if this solution satisfies the constraint. If it does, we are done. If not, we must seek a minimum that occurs when the constraint is satisfied at equality. We will divide the solution into two parts. In the first part we will convert the problem to the form:

$$\min_z \|z - c\| \quad \text{s.t.} \quad z^T Dz \geq 0,$$

Later, we will show how to turn the new problem into a sixth degree polynomial in a single variable.

Step 1. First, we can assume without loss of generality that b resides in the column space of A, since the component of b orthogonal to this space does not affect the solution to the problem. Furthermore, since b lies in the column space of A we can assume that A is 4×4 full rank and b is 4×1. This can be achieved, for example, using a QR decomposition. Now, define b' such that $Ab' = b$ (this is possible because A is full rank). Then, $Ax - b = A(x - b')$, implying that our problem is equivalent to:

$$\min_x \|A(x - b')\| \quad \text{s.t.} \quad x^T Bx = 0.$$

Using the method presented in Golub and van Loan [14] (see the second edition, pp. 466-471, especially Algorithm 8.7.1), we simultaneously diagonalize $A^T A$ and B. This will produce a nonsingular matrix X such that $X^T A^T AX = I$ and $X^T BX = D$, I denotes the identity matrix, and D is a 4×4 diagonal matrix. Thus, we obtain

$$\min_x \|X^{-1}(x - b')\| \quad \text{s.t.} \quad x^T X^{-T} DX^{-1} x = 0,$$

where X^{-1} denotes the inverse of X, and X^{-T} denotes its transpose. Denote $z = X^{-1}x$ and $c = X^{-1}b'$, then we obtain

$$\min_z \|z - c\| \quad \text{s.t.} \quad z^T Dz = 0.$$

This has the desired form.

Step 2. At this point, we attempt to solve a problem of the form

$$\min_z \|z - c\| \quad \text{s.t.} \quad z^T Dz = 0.$$

We solve this minimization problem using Lagrange multipliers. That is,

$$\min_{\{z,\lambda\}} \|z - c\|^2 + \lambda z^T Dz.$$

Taking the derivatives with respect to z and λ, we get

$$z - c + \lambda Dz = 0,$$

and

$$z^T Dz = 0.$$

From the first equation, we get

$$z = (I + \lambda D)^{-1}c.$$

Since D is diagonal, the components of z are given by

$$z_i = \frac{c_i}{1 + \lambda d_i},$$

where $z = (z_1, \ldots, z_4)$, $c = (c_1, \ldots, c_4)$, and $D = \text{diag}(d_1, \ldots, d_4)$. The constraint $z^T Dz = 0$, thus becomes

$$\sum_{i=1}^4 \frac{c_i^2 d_i}{(1 + \lambda d_i)^2} = 0,$$

which, after multiplying out the denominator, becomes a sixth degree polynomial in λ. This polynomial can be efficiently and accurately solved using standard techniques

(we use the MATLAB function *roots*). We plug in all solutions to determine x, as indicated above, and choose the real solution that minimizes our optimization criteria.

ACKNOWLEDGMENTS

The authors would like to thank Bill Bialek, Peter Belhumeur, Peter Blicher, Leonid Gurvits, Mike Langer, and Warren Smith. Their insightful comments and suggestions have been of great assistance to us. Some of our preliminary ideas on this topic were developed with Peter Belhumeur and are described in Jacobs et al. [18]. We are also grateful to Rui Ishiyama, Shizuo Sakamoto, and Johji Tajima for their helpful comments and for providing us with data for our experiments. An early version of this paper has appeared in [1]. Major portions of this research were conducted while Ronen Basri was at NEC Research Institute, Princeton, New Jersey. At the Weizmann Institute Ronen Basri is supported in part by the European Community grant number IST-2000-26001 and by the Israel Science Foundation grant number 266/02. The vision group at the Weizmann Institute is supported in part by the Moross Foundation.

REFERENCES

[1] R. Basri and D.W. Jacobs, "Lambertian Reflectances and Linear Subspaces," *IEEE Int'l Conf. Computer Vision,* 2001.
[2] R. Basri and D.W. Jacobs, "Photometric Stereo with General, Unknown Lighting," *IEEE Conf. Computer Vision and Pattern Recognition,* vol. II, pp. 374-381, 2001.
[3] P. Belhumeur, J. Hespanha, and D. Kriegman, "Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 19, no. 7, pp. 711-720, July 1997.
[4] P. Belhumeur and D. Kriegman, "What Is the Set of Images of an Object Under All Possible Lighting Conditions?" *IEEE Conf. Computer Vision and Pattern Recognition,* pp. 270-277, 1996.
[5] A.P. Blicher and S. Roy, personal communication.
[6] B. Cabral, N. Max, and R. Springmeyer, "Bidirectional Reflection Functions from Surface Bump Maps," *Computer Graphics,* vol. 21, no. 4, pp. 273-281, 1987.
[7] T.F. Cootes, C.J. Taylor, D.H. Cooper, and J. Graham, "Training Models of Shape from Sets of Examples," *Proc. British Machine Vision Conf.,* pp. 9-18, 1992.
[8] Y. Dobashi, K. Kaneda, H. Nakatani, and H. Yamashita, "A Quick Rendering Method Using Basis Functions for Interactive Lighting Design," *Eurographics,* pp. 229-240, 1995.
[9] M. D'Zmura, "Shading Ambiguity: Reflectance and Illumination," *Computational Models of Visual Processing,* M. Landy and J. Movshon, eds., 1991.
[10] R. Epstein, P. Hallinan, and A. Yuille, "5 ± 2 Eigenimages Suffice: An Empirical Investigation of Low-Dimensional Lighting Models," *Proc. IEEE Workshop Physics-Based Vision,* pp. 108-116, 1995.
[11] A. Georghiades, D. Kriegman, and P. Belhumeur, "Illumination Cones for Recognition Under Variable Lighting: Faces," *Proc. IEEE Conf. Computer Vision and Pattern Recognition,* pp. 52-59, 1998.
[12] A. Georghiades, P. Belhumeur, and D. Kriegman, "From Few to Many: Generative Models for Recognition Under Variable Pose and Illumination," *Proc. Int'l Conf. Automatic Face and Gesture Recognition,* 2000.
[13] H. Groemer, *Geometric Applications of Fourier Series and Spherical Harmonics.* Cambridge Univ. Press, 1996.
[14] G. Golub and C. van Loan, *Matrix Computations.* Baltimore: Johns Hopkins Univ. Press, 1989.
[15] P. Hallinan, "A Low-Dimensional Representation of Human Faces for Arbitrary Lighting Conditions," *Proc. IEEE Conf. Computer Vision and Pattern Recognition,* pp. 995-999, 1994.
[16] H. Hayakawa, "Photometric Stereo under a Light Source with Arbitrary Motion," *J. Optical Soc. of Am.,* vol. 11, no. 11, pp. 3079-3089, 1994.
[17] B.K.P. Horn, *Robot Vision.* Cambridge, Mass.: MIT Press, 1986.
[18] D.W. Jacobs, P.N. Belhumeur, and R. Basri, "Comparing Images under Variable Illumination," Technical Report, NEC Research Inst., pp. 97-183, 1997.
[19] M. Kirby and L. Sirovich, "The Application of the Karhunen-Loeve Procedure for the Characterization of Human Faces," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 12, no. 1, pp. 103-108, Jan. 1990.
[20] J. Koenderink and A. van Doorn, "Bidirectional Reflection Distribution Function Expressed in Terms of Surface Scattering Modes," *Proc. European Conf. Computer Vision,* vol. 2, pp. 28-39, 1996.
[21] J. Koenderink and A. van Doorn, "The Generic Bilinear Calibration-Estimation Problem," *Int'l J. Computer Vision,* vol. 23, no. 3, pp. 217-234, 1997.
[22] J. Lambert, "Photometria Sive de Mensura et Gradibus Luminus, Colorum et Umbrae," Eberhard Klett, 1760.
[23] K.C. Lee, J. Ho, and D. Kriegman, "Nine Points of Light: Acquiring Subspaces for Face Recognition under Variable Lighting," *Proc. IEEE Conf. Computer Vision and Pattern Recognition,* pp. 519-526, 2001.
[24] T. Malzbender, D. Gelb, and H. Wolters, "Polynomial Texture Maps," *Proc. ACM SIGGRAPH,* pp. 519-528, 2001.
[25] G. Miller and R. Hoffman, "Illumination and Reflection Maps: Simulated Objects in Simulated and Real Environments," *Proc ACM SIGGRAPH, Course Notes for Advanced Computer Graphics Animation,* July 1984.
[26] Y. Moses, "Face Recognition: Generalization to Novel Images," PhD thesis, Weizmann Inst. of Science, 1993.
[27] H. Murase and S. Nayar, "Visual Learning and Recognition of 3D Objects from Appearance," *Int'l J. Computer Vision,* vol. 14, no. 1, pp. 5-25, 1995.
[28] S. Nayar and H. Murase, "Dimensionality of Illumination Manifolds in Appearance Matching," *Proc. NSF/ARPA Workshop 3D Object Representation for Computer Vision,* p. 165, 1996.
[29] J. Nimeroff, E. Simoncelli, and J. Dorsey, "Efficient Re-Rendering of Naturally Illuminated Environments," *Proc. Fifth Eurographics Workshop Rendering,* 1994.
[30] R. Ramamoorthi, "Analytic PCA Construction for Theoretical Analysis of Lighting Variability in a Single Image of a Lambertian Object," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 24, no. 10, pp. 1322-1333, Oct. 2002.
[31] R. Ramamoorthi and P. Hanrahan, "On the Relationship between Radiance and Irradiance: Determining the Illumination from Images of Convex Lambertian Object," *J. Optical Soc. Am.,* vol. 18, no. 10, pp. 2448-2459, 2001.
[32] R. Ramamoorthi and P. Hanrahan, "A Signal-Processing Framework for Inverse Rendering," *Proc. ACM SIGGRAPH,* pp. 117-128, 2001.
[33] R. Ramamoorthi and P. Hanrahan, "An Efficient Representation for Irradiance Environment Maps," *Proc. ACM SIGGRAPH,* pp. 497-500, 2001.
[34] A. Shashua, "On Photometric Issues in 3D Visual Recognition from a Single 2D Image, " *Int'l J. Computer Vision,* vol. 21, no. 1-2, pp. 99-122, 1997.
[35] P.C. Teo, E.P. Simoncelli, and D.J. Heeger, "Efficient Linear Re-Rendering for Interactive Lighting Design," Techical Report STAN-CS-TN-97-60, Stanford Univ., 1997.
[36] K. Thornber and D. Jacobs, "Broadened, Specular Reflection and Linear Subspaces," NEC Technical Report 2001-033, 2001.
[37] M. Turk and A. Pentland, "Eigenfaces for Recognition," *J. Cognitive Neuroscience,* vol. 3, no. 1, pp. 71-96, 1991.
[38] S. Ullman and R. Basri, "Recognition by Linear Combinations of Models," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 13, no. 10, pp. 992-1007, Oct. 1991.
[39] S.H. Westin, J.R. Arvo, and K.E. Torrance, "Predicting Reflectance Functions from Complex Surfaces," *Computer Graphics,* vol. 26, no. 2, pp. 255-264, 1992.
[40] A. Yuille, D. Snow, R. Epstein, and P. Belhumeur, "Determining Generative Models of Objects Under Varying Illumination: Shape and Albedo from Multiple Images Using SVD and Integrability," *Int'l J. Computer Vision,* vol. 35, no. 3, pp. 203-222, 1999.
[41] L. Zhao and Y. H. Yang, "Theoretical Analysis of Illumination in PCA-Based Vision Systems," *Pattern Recognition,* vol. 32, pp. 547-564, 1999.

Ronen Basri received the BSc degree in mathematics and computer science from Tel Aviv University in 1985, where he graduated Summa Cum Laude. He received the PhD degree in computer science from the Weizmann Institute of Science in 1990. From 1990 to 1992, he was a postdoctoral fellow at the Massachusetts Institute of Technology in the Department of Brain and Cognitive Science and the Artificial Intelligence Laboratory under the McDonnell-Pew and the Rothchild programs. Since then, he has been affiliated with the Weizmann Institute of Science in the Department of Computer Science and Applied Mathematics, where he currently holds the position of associate professor. Between 1999 and 2000, he spent a sabbatical at NEC Research Institute in Princeton, New Jersey. Dr. Basri's research has focused on computer vision, especially in the areas of object recognition, perceptual organization, and visually guided navigation. He is a member of the IEEE and the IEEE Computer Society.

David W. Jacobs received the BA degree from Yale University in 1982. From 1982 to 1985, he worked for Control Data Corporation on the development of database management systems, and attended graduate school in computer science at New York University. From 1985 to 1992, he attended Massachusetts Institute of Technology, where he received MS and PhD degrees in computer science. Since then, he has been at NEC Research Institute, where he is a senior research Scientist. In 1998, he spent a sabbatical at the Royal Institute of Technology (KTH) in Stockholm. Dr. Jacobs has also been a visiting member of the Courant Institute at New York University, where he has taught classes in computer vision and learning. Dr. Jacobs' research has focused on human and computer vision, especially in the areas of object recognition and perceptual organization. He has also published papers in the areas of motion understanding, memory and learning, and computational geometry. He is an associate editor of *IEEE Transactions on Pattern Analysis and Machine Intelligence*. He and his coauthors received honorable mention for the best paper award at CVPR 2000. He is a member of the IEEE and the IEEE Computer Society.