Appearance Sampling for Obtaining A Set of Basis Images for Variable Illumination

Imari Sato   Takahiro Okabe   Yoichi Sato   Katsushi Ikeuchi
Institute of Industrial Science, The University of Tokyo, Japan
{imarik, takahiro, ysato, ki}@iis.u-tokyo.ac.jp

Abstract

Previous studies have demonstrated that the appearance of an object under varying illumination conditions can be represented by a low-dimensional linear subspace. A set of basis images spanning such a linear subspace can be obtained by applying the principal component analysis (PCA) for a large number of images taken under different lighting conditions. While the approaches based on PCA have been used successfully for object recognition under varying illumination conditions, little is known about how many images would be required in order to obtain the basis images correctly. In this study, we present a novel method for analytically obtaining a set of basis images of an object for arbitrary illumination from input images of the object taken under a point light source. The main contribution of our work is that we show that a set of lighting directions can be determined for sampling images of an object depending on the spectrum of the object's BRDF in the angular frequency domain such that a set of harmonic images can be obtained analytically based on the sampling theorem on spherical harmonics. In addition, unlike the previously proposed techniques based on spherical harmonics, our method does not require the 3D shape and reflectance properties of an object used for rendering harmonics images of the object synthetically.

1. Introduction

It is well known that the appearance of an object changes significantly under different illumination conditions. For instance, the appearance change of someone's face often becomes much larger than the difference of two different faces under the same lighting. Thus, for the task of object recognition, it is very important to be able to predict the variation of objects' appearance under varying illumination conditions.

Previous studies have demonstrated that appearance changes of an object under varying illumination can be represented with a low-dimensional linear subspace. In the most simplistic case of a convex Lambertian object under distant illumination without attached and cast shadows, the appearance of the object can be completely described with a 3-D linear subspace defined with three input images taken under linearly independent lighting conditions [9, 15, 18]. However, the assumption of this model would be too restrictive to be used for object recognition in more realistic settings.

Other researchers have reported empirical studies for representing image variation due to varying illumination for human faces and other objects [7, 4, 17]. Interestingly enough, most of the image variation caused by varying illumination can be explained with a low-dimensional linear subspace slightly higher than 3D even when images contain a significant amount of shadows. For instance, Hallinan reported a) that a 5-D subspace would suffice to represent most of the image variation due to illumination change including extreme cases where a face is lit from its sides, and b) that a 3-D subspace would be sufficient when a face is lit mainly from its front [7]. Georghiades et al. used a similar observation more specifically for object recognition under varying lighting conditions [5, 6].

A set of basis images spanning such a linear subspace can be obtained by applying the principal component analysis to a large number of images of an object taken under different lighting, e.g., by moving a point light source along a sphere surrounding the object. One might ask whether a certain set of input images of an object would be sufficient to fully span the subspace of the object for arbitrary illumination conditions. Previous empirical studies do not necessarily provide enough insight to this important question, and thus it has been a common practice to use as many input images as possible to ensure that the set of input images span a subspace entirely. Another interesting observation of Hallinan's early work is that a subspace obtained by PCA does not vary widely between sparsely and densely sampled sets of lighting directions in the case of faces [7]. However, it was not known how many images would be sufficient in order to obtain the basis images correctly.

Recently, it was shown through the frequency-space analysis of reflection that appearance of a convex Lambertian object can be well represented with a 9-D linear subspace spanned by basis images of the object, called *harmonic images*, each of which corresponds to an image of the object

1

illuminated under *harmonic lights* whose distributions are specified in terms of spherical harmonics[10, 11, 1]. Basri and Jacobs successfully used this 9-D subspace defined with harmonic images for face recognition under varying illumination [1]. Other researchers have also used a set of harmonic images for the purpose of efficient rendering of an object under complex illumination [12, 16]. More recently, Ramamoorthi has shown theoretically that, under certain conditions of lighting and object shape and reflectance, a set of basis images obtained from PCA on input images taken under varying lighting coincides with a set of harmonic images [13].

While harmonic images are specified analytically, it is difficult to obtain harmonic images for various kinds of real objects because harmonic lights consist of both negative and positive values distributed in a complex way and thus do not exist as real lighting. Therefore, the previously proposed techniques for object recognition and image synthesis based on harmonic images require the model of an object about its 3-D shape and reflectance properties such as albedo so that harmonic images of the object under harmonic light can be rendered synthetically. This is not a problem for the case of synthetic objects because the model of a synthetic object is given a priori by definition. On the other hand, acquiring shape and reflectance properties of real objects is not an easy problem and therefore, most of the previous studies treated real objects as Lambertian objects and estimated only albedos by using vision techniques such as photometric stereo.

This motivated us to develop a method for determining a set of harmonic images of a real object by using only input images of the object under simple lighting such as a point light source. Lee et al. recently proposed an interesting method to determine a configuration of 9 light source directions such that input images taken under those light source directions approximate a 9-D subspace spanned by harmonic images [8]. They have reported that such a configuration of 9 light source directions does not vary much for similar objects, e.g., different faces. However, Lee et al.'s method still needs a set of harmonic images that are rendered synthetically by using an object's model in order to determine a set of 9 lighting directions. Moreover, Lee et al.'s method chooses 9 light directions from a large number of candidates and therefore, a large number of input images are required for each new object or new class of objects.

In this study, we present a novel method for analytically obtaining a set of basis images of a convex object for arbitrary illumination from input images of the object taken under a point light source. The main contribution of our work is that we show that a set of lighting directions can be determined for sampling images of an object depending on the spectrum of the object's BRDF in the angular frequency domain such that a set of harmonic images can be obtained analytically based on the sampling theorem on spherical har-

monics [3]. Using those sampling directions determined from the sampling theory, we are able to obtain harmonic images by using a significantly smaller number of input images than other techniques which do not take into account a relationship between a spectrum of BRDFs and a sampling density of illumination directions. In addition, unlike other methods based on spherical harmonics, our method does not require the shape and reflectance model of an object used for rendering harmonics images of the object synthetically. Thus, our method can be easily applied for determining a set of basis images for representing the appearance change of a real object under varying illumination conditions.

The rest of the paper is organized as follows. We briefly review the spherical harmonics transformation of a function over the unit sphere and harmonics images based on spherical harmonics in Section 2. We show details of the sampling theorem used in our method for obtaining harmonic images of real objects in Section 3, and consider issues in obtaining harmonic images of real objects based on this sampling theorem in Section 4. In particular, we discuss what kind of artifacts should be expected when input images provide insufficient samplings of an object's appearance. Finally, we show experimental results of the proposed method applied to images of real objects in Section 5, and present concluding remarks in Section 6.

2. A Set of Basis Images for Variable Illumination

2.1 Review of Spherical Harmonics

Spherical Harmonics define an orthonormal basis over the unit sphere. Consider the unit sphere in R^3, a unit vector on the sphere can be described by the polar coordinate system $\theta, (0 \leq \theta \leq \pi)$ in elevation and $\phi, (0 \leq \phi < 2\pi)$ in azimuth. In this coordinate system, spherical harmonics $Y_l^m(\theta, \phi), (l \geq 0, -l \leq m \leq l)$ are defined as

$$Y_l^m(\theta, \phi) = N_l^m P_l^m(cos\theta)e^{Im\phi}, \qquad (1)$$

where N_l^m are the normalized constants, and $P_l^m(\cdot)$ are the associated Legendre functions of degree l and order m.

A function $f(\theta, \phi)$ defined over the unit sphere is expanded as a linear combination of spherical harmonics as

$$f(\theta, \phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} f_l^m Y_l^m(\theta, \phi), \qquad (2)$$

and f_l^m denote coefficients in its spherical harmonic expansion computed as [1]

$$f_l^m = \int_0^{2\pi} \int_0^{\pi} f(\theta, \phi) Y_l^m(\theta, \phi) \, sin\theta d\theta d\phi. \qquad (3)$$

---

[1] In this study, we consider spherical harmonics in a real form.

2

2.2 Harmonic Image Representation

The reflectance property of an object is characterized by a bidirectional reflectance distribution function (BRDF) $\rho(\theta_i', \phi_i', \theta_o', \phi_o')$, where (θ_i', ϕ_i') and (θ_o', ϕ_o') are incident and reflection directions with respect to the surface normal of the object surface whose local coordinate is denoted by using $'$.

Then, the brightness E of the object surface is represented by the global coordinate defined on the unit sphere as

$$E = \int_0^\pi \int_0^{2\pi} L(\theta, \phi) \rho(Rot(\theta, \phi), \theta_o', \phi_o') \cos\theta \sin\theta d\theta d\phi, \tag{4}$$

where $L(\theta, \phi)$ is light source distribution, and $Rot(\cdot)$ represents a rotation operator that rotates (θ, ϕ) into the local coordinate.

In this paper, we consider the appearance of an object under variable illumination seen from a fixed viewpoint and therefore, we represent $\rho(Rot(\theta, \phi), \theta_o', \phi_o') \cos\theta$ by using the global coordinate as $R(\theta, \phi)$, and refer to it as the *reflection kernel*.

Since both the light source distribution and the reflection kernel are functions defined on the unit sphere, we can represent them as

$$L(\theta, \phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} L_l^m Y_l^m(\theta, \phi), \tag{5}$$

$$R(\theta, \phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} R_l^m Y_l^m(\theta, \phi). \tag{6}$$

where L_l^m and R_l^m are coefficients in their spherical harmonic expansion from (3). Assuming spherical light sources with $Y_l^m(\theta, \phi)$ radiance in its (θ, ϕ) direction, R_l^m represents the brightness seen under these spherical light sources called harmonic lights in (3).

From (5), (6), and the orthonormality of spherical harmonics, the surface brightness E in (4) is represented as

$$E = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} L_l^m R_l^m. \tag{7}$$

Here, if we prepare R_l^m for all pixels, that is, for all corresponding points on the object surface, images under variable illumination are represented from (7). Images containing R_l^m are called harmonic images [1].

3 Methods for Obtaining Harmonic Image

There are several approaches for obtaining harmonic images of an object. One approach is to provide a reflection kernel $R(\theta, \phi)$ of the object from the knowledge of its 3D shape and reflectance properties. Since $Y_l^m(\theta, \phi)$ are predefined functions, R_l^m are computed from (3). Most of the previously proposed methods employed this approach to compute harmonic images synthetically.

One might think of observing an object under physically constructed harmonic lights. In this case, each pixel value directly corresponds to coefficients R_l^m, and thus the 3D shapes and reflectance properties of the object need not be given a priori. However, as pointed out by other researchers [1, 8], harmonic lights are complex diffuse lighting consisting of both negative and positive radiance. Thus, it is difficult to physically construct such lighting in a real setting.

In this work, we take an alternative approach of observing a reflection kernel $R(\theta, \phi)$ for each surface point on an object directly by taking an image of the object under a point light source located at the direction (θ, ϕ). This approach is based on the assumption that a point light source used in this method is sufficiently distant from the objects and thus, this light source projects parallel rays onto the object surface.

The reflectance kernel $R(\theta, \phi)$ represents radiance of reflected light due to incoming light with unit radiance from the direction (θ, ϕ). Thus if we suppose that this point light source has unit radiance, the value of $R(\theta, \phi)$ can be obtained simply as an observed brightness of the surface point when the object is illuminated under a light source from the direction (θ, ϕ).[2] Once we determine the values of $R(\theta, \phi)$ for all directions, we can compute pixel values of harmonic images as R_l^m from (3). In this way, we do not need to synthetically compute the reflection kernel $R(\theta, \phi)$ of the object, nor are the 3D shapes and reflectance properties of the object required.

Since the function $R(\theta, \phi)$ are continuously distributed on the unit sphere in (3), we first need to approximate its distribution by a discrete set of the function $R(\theta, \phi)$ so that we can sample $R(\theta, \phi)$ using a point light source physically located at (θ, ϕ) direction. Then, the question we have to ask is how densely $R(\theta, \phi)$ need be sampled in order to correctly compute coefficients R_l^m from them. In other words, we want to know how densely a point light source needs to be positioned around an object to obtain harmonics images of the object correctly. In the next section, we will consider this issue in terms of the characteristics of spherical harmonics $Y_l^m(\theta, \phi)$.

3.1 Sampling Theorem on Spherical Harmonics Transformation

There have been several methods proposed in the field of applied mathematics to efficiently compute coefficients of a

---

[2]Note that we can determine $R(\theta, \phi)$ only up to some unknown scaling factor, so it is reasonable to treat a point light source to have unit radiance.

3

function $R(\theta, \phi)$ in its spherical harmonic expansion using fewer samplings of the function $R(\theta, \phi)$. We have adapted one of their theories to enable us to compute harmonic images using fewer input images of objects taken by moving a point light source to particular locations.

It is common knowledge that the sampling theorem on the 1D line tells us that a band-limited function can be reconstructed accurately from properly sampled discrete data. Namely, Fourier coefficients of the function can be determined by weighted sums of the function sampled at even intervals. In the case of a function defined on the sphere, the similar theorem, the sampling theorem on the sphere, has been proved [3]. In this section, we outline the theorem.

Let us assume that the reflection kernel $R(\theta, \phi)$ is *band-limited* with *bandwidth* N, that is, $R_l^m = 0$ $(l \geq N)$. Then, consider the "comb" function $s(\theta, \phi)$ with equiangular grid $(\Delta\theta, \Delta\phi) = (\pi/2N, 2\pi/2N)$

$$s(\theta, \phi) = \frac{\sqrt{2\pi}}{2N} \sum_{j=0}^{2N-1} \sum_{k=0}^{2N-1} w_j \delta(\theta - \theta_j)\delta(\phi - \phi_k), \quad (8)$$

where $(\theta_j, \phi_k) = (\pi(j + 1/2)/2N, 2\pi k/2N)$ are sampling points on the sphere, and w_j are weight of the points. We can compute the coefficients s_l^m of the comb function from definition (3). For degree $l < 2N$, we obtain

$$s_l^m = \sum_{j=0}^{2N-1} w_j \sqrt{\frac{2l+1}{2}} P_l^0(\cos\theta_j)\delta_{m0}, \quad (9)$$

where the Kronecker delta $\delta_{ij} = 1$ if $i = j$, and $\delta_{ij} = 0$ if $i \neq j$. Here, we can uniquely choose the weight w_j so that $\sum_{j=0}^{2N-1} w_j P_l^0(\cos\theta_j) = \sqrt{2}\delta_{l0}$. Thus, the coefficients are described simply as $s_l^m = \delta_{l0}\delta_{m0}$ for degree $l < 2N$. Equivalently, the comb function is represented by the addition of $Y_0^0(\theta, \phi) = const.$ and higher-degree terms as

$$s(\theta, \phi) = 1 + \sum_{j \geq 2N} \sum_{|k| \leq j} s_j^k Y_j^k(\theta, \phi). \quad (10)$$

Then, from (6) and (10), the product of the reflection kernel and the comb function is written as

$$R(\theta, \phi) \cdot s(\theta, \phi) = R(\theta, \phi)$$
$$+ \sum_{l < N} \sum_{|m| \leq l} \sum_{j \geq 2N} \sum_{|k| \leq j} R_l^m s_j^k Y_l^m(\theta, \phi) Y_j^k(\theta, \phi), \quad (11)$$

where degree $l < N$ because $R(\theta, \phi)$ is band-limited. The second term is known as *aliasing* introduced by discrete sampling. However, it is known that the product of the spherical harmonics $Y_l^m(\theta, \phi) Y_j^k(\theta, \phi)$ is represented as a linear combination of spherical harmonics with a degree greater than or equal to $|l - j|$. Accordingly, aliasing appears to be a degree greater than or equal to $|N - 2N| = N$

in this case. Therefore, for degree $l < N$, $R(\theta, \phi) \cdot s(\theta, \phi)$ is equal to $R(\theta, \phi)$, that is,

$$(R(\theta, \phi) \cdot s(\theta, \phi))_l^m = R_l^m. \quad (12)$$

Hence, the coefficients of the reflection kernel can be computed accurately by the Fourier transform of $R(\theta, \phi) \cdot s(\theta, \phi)$ as

$$R_l^m = \frac{\sqrt{2\pi}}{2N} \sum_{j=0}^{2N-1} \sum_{k=0}^{2N-1} w_j R(\theta_j, \phi_k) Y_l^m(\theta_j, \phi_k), \quad (13)$$

where the weight w_j are analytically given by

$$w_j = \frac{2\sqrt{2}}{2N} \sin\theta_j \sum_{n=0}^{N-1} \frac{1}{2n+1} \sin[(2n+1)\theta_j]. \quad (14)$$

4 Appearance Sampling of Real Objects based on The Sampling Theorem

The sampling theorem described in the previous section tells us the minimum number of sampling $2N \times 2N = 4N^2$ to compute spherical harmonics transformation of a band-limited function with bandwidth N. In this section, we consider issues in defining bandwidth of an object and discuss what kind of artifacts we should expect when the function is not band-limited within N.

4.1 Convex Lambertian Surface

Let us start with the simplest case of convex Lambertian objects. It has been shown in the previous studies that the first nine spherical harmonics with the order $l = 2$ are sufficient to capture more than 99% of the reflection energy of a convex Lambertian surface [1, 10, 11].

Accordingly, we can consider that the function $R(\theta, \phi)$ is band-limited with bandwidth $N = 3$, and this results in $4N^2 = 36$ samplings of $R(\theta, \phi)$ necessary for computing R_l^m correctly. In other words, the coefficients R_l^m are given as a finite weighted sums of the function $R(\theta, \phi)$ sampled at equiangular grid: $\theta_j = \frac{\pi(j+1/2)}{6}$ $(j = 0, \ldots, 5)$, $\phi_k = \frac{2\pi k}{6}$ $(k = 0, \ldots, 5)$. Namely, 36 input images of an object taken by moving a point light source to the directions specified with (θ_j, ϕ_k) on a sphere around the object are required to compute harmonic images of the object.

Note that what needs to be satisfied is not the number of samplings $(4N^2)$ of the function but rather, the sampling intervals $(\Delta\theta = \pi/2N, \Delta\phi = 2\pi/2N)$ that this sampling can provide. For instance, even when a large number of input images of an object taken under a point light source are available, there is no guarantee that this set of images

4

can produce correct harmonic images that can span a low-dimensional linear subspace representing the appearance of the object under arbitrary illumination unless those intervals are satisfied.

4.2 Complex Appearance beyond Lambertian

Ramamoorthi and Hanrahan analytically derived the bandwidth of reflection kernels of objects that have more complex appearance than convex Lambertian surfaces such as the Phong reflection model and the Microfacet BRDF [10].[3] For instance, this study shows that the bandwidth of the Microfacet BRDF is approximately $N \approx \sigma^{-1}$. Thus, if the surface roughness of an object is predicted even roughly, it should help us to find the bandwidth of the reflection kernel of the object.

There are a large number of previous studies on BRDF measurements and BRDF databases which show reflectance parameters of various kinds of surface materials. This knowledge should be useful for estimating bandwidths of reflection kernels of various objects based on the analysis presented in [10].

4.3 Aliasing Caused by Insufficient Sampling

There certainly is a situation where bandwidth of reflection kernels of an object is difficult to predict, or where only a limited number of samplings of the function are obtainable depending on a particular hardware set-up used for acquiring input images. Hence it is important to consider what kinds of artifacts we should expect when the function $R(\theta, \phi)$ has a lower or a higher bandwidth than N determined from the number of samplings $4N^2$ based on the sampling theorem.

Let B denote actual bandwidth of the function. Driscoll and Healy proved that the error in spherical harmonics transformation, generally known as aliasing, is confined to coefficients R_l^m of degree greater than or equal to $|B - 2N|$ in [3]. This can be confirmed by replacing the bandwidth N of the reflection function with B in (11).

From this, in the case where the function has bandwidth B lower than N ($B < N$), there is no aliasing for degree less than $N + k$ for $B = N - k$. Since we compute coefficients R_l^m for $l < N$, those coefficients are correctly computed. In this case, all of the coefficients R_l^m for $B \leq l < N$ become 0.

In the case where the function $R(\theta, \phi)$ has a bandwidth B higher than N ($B > N$), error due to insufficient sampling of the function is confined to the coefficients of degree $l \geq$

$N - k$ for $B = N + k$. In this case, the coefficients R_l^m for $l < N - k$ degree are correctly computed. Therefore, closer to B we select N, smaller error in R_l^m we expect.

Next we show what happens when we have sufficient or insufficient samplings of the function $R(\theta, \phi)$ by using synthetic data. The function $R(\theta, \phi)$ used here is a reflection kernel for the Torrance-Sparrow reflection model with known reflection parameters ($K_d = \{.62, .71, .62\}, K_s = 0.29, \sigma = 0.1$).[4] The surface normal and viewing direction for this reflection kernel are set at the direction ($\theta = 0, \phi = 0$).

We computed the coefficients R_l^m up to the degree $B = \sigma^{-1} \approx 10$ for this reflection kernel from discrete samplings of $R(\theta, \phi)$ at different sampling intervals: $N = 200$ (160000 sampling), $N = 11$ (484 samplings), $N = 5$ (100 samplings), $N = 3$ (36 samplings). The computed coefficients are shown in Figure 1a)∼d). Spherical harmonic coefficient for given degree l and order m are represented using a single index $r = l^2 + l + m$. The horizontal axis represents the index r, and the vertical axis represents the computed coefficients R_l^m. The number of samplings is shown for each graph, and a red arrow represents the upper bound N determined from the number of samplings.

Figure 1e)∼h) are visualization of the reflection kernel $R(\theta, \phi)$ reconstructed from the computed coefficients R_l^m in (6) up to the degree $l = 10$ for the cases of $N = 200$, $N = 11$, $N = 5$, and $N = 3$, respectively. Figure 1i) shows a reflection kernel reconstructed from the coefficients up to the degree $l = 2$ for the cases of $N = 3$. In these figures, the upper half of $R(\theta, \phi)$ is visualized in a polar coordinate system with radius indicating $\{\theta | 0 \leq \theta \leq \pi/2\}$, and angle indicating $\{\phi | 0 \leq \phi < 2\pi\}$ as illustrated in Figure 1j).

From Figure 1a) and b), we can see that the coefficients R_l^m obtained from $4N^2 = 484$ samplings are almost the same as those obtained from $4N^2 = 160000$ samplings, and the $R(\theta, \phi)$ is reconstructed correctly in both cases. This shows that a set of samplings for $N = 11$ are sufficient to capture the appearance of a surface with this reflection kernel under varying illumination. This follows what the sampling theorem tells us.

On the other hand, in the cases of $N = 5$ and $N = 3$ where N is smaller than $B = 10$, the computed coefficients suffer from aliasing. It is important to note that the coefficients R_l^m for $l < N$ (the left side of the red arrow in Figure 1c) and d)) closely match the coefficients obtained from sufficient samplings (Figure 1a)) while the coefficients R_l^m for $N \leq l \leq B$ (the right side of the red arrow) differ. As a result, distinct ringing effects are evident in the reflection kernel reconstructed from the coefficients up to the degree $l = 10$ (Figure 1g) and h)). On the other hand, the reflec-

---

[3][14] is also a good reference to get insights into the bandwidth of a reflection kernel of various types of objects in CURET database. This work discusses appropriate sampling resolutions to model appearance of those objects based on their experiments as well.

[4]K_d and K_s are constants for the diffuse and specular reflection components, and σ is the standard deviation of a facet slope of a simplified Torrance-Sparrow reflection model.

5

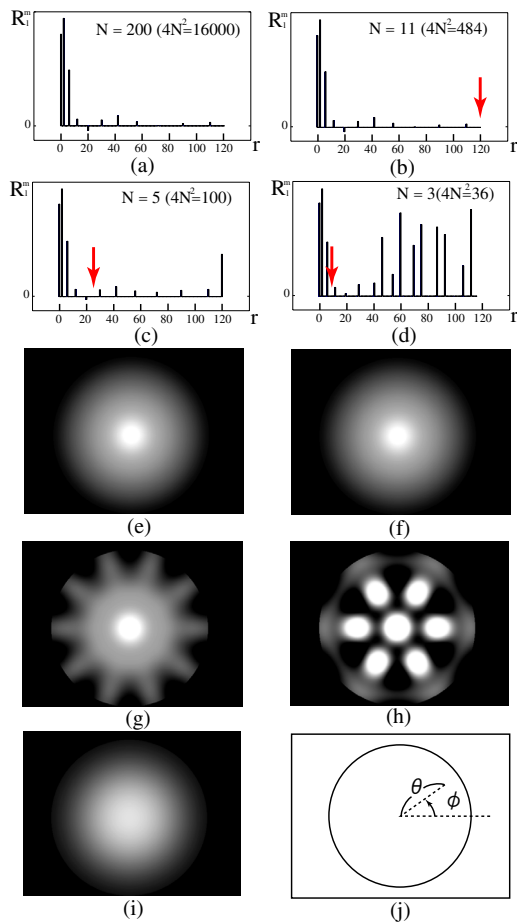(a) ... (b) ... (c) ... (d) ... (e) ... (f) ... (g) ... (h) ... (i) ... (j)

Figure 1: Experimental results using synthetic data: coefficients are computed from a discrete set of $f(\theta, \phi)$ sampled at different sampling intervals.

tion kernel reconstructed from the coefficients up to the degree $l = 2$ for $N = 3$ shown in Figure 1i) does not show such ringing effects while high frequency components such as specularity are missing at the center.

It follows from these examples that it is appropriate to compute R_l^m for $l < N$ only when $4N^2$ samplings of the function are available. In this way, we can avoid annoying ringing effects caused by R_l^m for $N \leq l$ in the reconstructed function $R(\theta, \phi)$.

5 Experimental Results

We have tested the proposed method using real images of several objects taken by moving a point light source to equiangular grid points defined by the sampling theorem. Figure 2 shows an overview of our hardware set-up[5] used
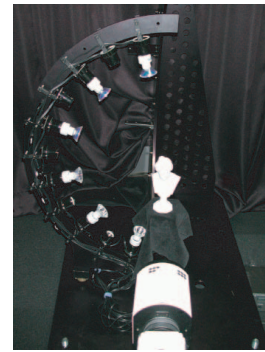
Figure 2: Image acquisition set-up

for obtaining the input images of the objects; an array of light sources is mounted on a turntable. These light sources are equally spaced in elevation, and the set of light sources is rotated around the objects in azimuthal.

In the case of the sheep and the Venus examples, since those objects have an appearance similar to that of a Lambertian surface, 36 input images of them are taken under a point light source positioned at equiangular grid points: $\theta_j = \frac{\pi(j+1/2)}{6}$, $\phi_k = \frac{2\pi k}{6}(j, k = 0, \ldots, 5)$. For the fish example, since it has a more complex appearance, 144 input images are taken under a point light source at equiangular grid points : $\theta_j = \frac{\pi(j+1/2)}{12}$, $\phi_k = \frac{2\pi k}{12}(j, k = 0, \ldots, 11)$. Based on the sampling theorem, coefficients R_l^m are computed up to the degree $l = 2$ from the 36 images, and up to the degree $l = 5$ from the 144 images.

The first nine harmonic images obtained from those input images are shown in Figure 3. In spite of given discrete sets of the appearance of the objects, the obtained harmonic images have complex and smooth shading that reflects the distribution of harmonic lights in these examples.

To evaluate the accuracy of the recovered harmonic images, we took images of those objects under normal lighting conditions in our laboratory, including direct light sources such as fluorescent lamps. Coefficients L_l^m in (5) representing the illumination condition are computed from an omnidirectional image of the scene taken by a camera with a fisheye lens in the same manner as described in [12]. Then appearance of the objects under this illumination condition is synthesized from (7).

In Figure 4, the left column shows the real images of the objects and the right column shows the synthesized appearance. The synthesized appearance of the objects resembles that of the objects in real images, and this shows that the recovered harmonic images provide a good representation of the appearance of the objects. The shoulder of the plaster figure in the middle right image appears darker than that in the real image. This is due to severe cast shadows observed in the input images that do not follow our method's assump-

6

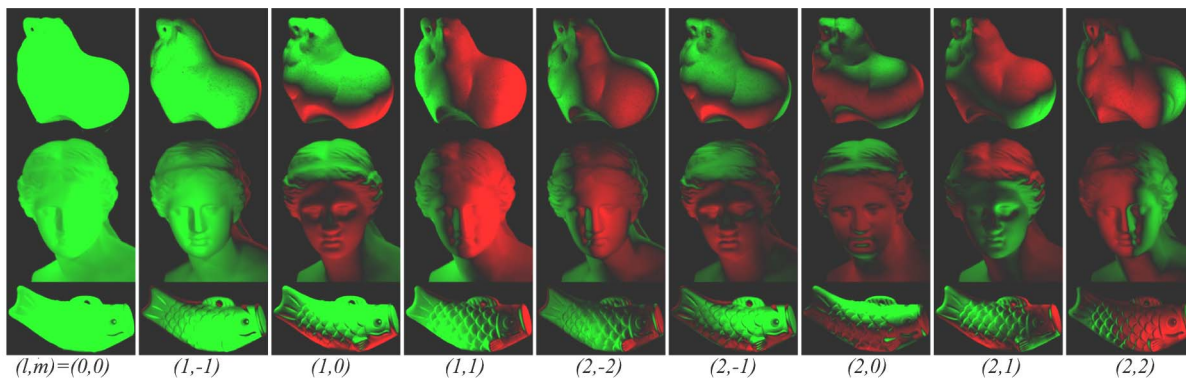|								
(l,m)=(0,0)	(1,-1)	(1,0)	(1,1)	(2,-2)	(2,-1)	(2,0)	(2,1)	(2,2)

Figure 3: Obtained harmonic images: positive values are shown in green, and negative values are shown in red.

tion for convex objects.

To demonstrate how well the recovered harmonic images represent the appearance of those objects under varying illumination, we synthesize their appearance from the recovered harmonic images under several natural illumination conditions provided by high-dynamic range light probe measurements by [2]. Figure 5 shows the results. In this figure, synthesized appearance changes dynamically depending on characteristics of the illumination, and one can say from this that the proposed method succeeded in providing a set of basis images representing appearance of those objects under varying illumination.

6 Summary and Conclusions

In this study, we have presented a novel method for analytically obtaining a set of basis images of an object for arbitrary illumination from input images of the object taken under a point light source. The main contribution of our work is that we have shown that a set of lighting directions can be determined for sampling images of an object depending on the spectrum of the object's BRDF in the angular frequency domain such that a set of harmonic images can be obtained analytically based on the sampling theorem on spherical harmonics. The future directions of this study include extending our method for: (1) modeling appearance of object seen from arbitrary viewing directions, and (2) extending our method further for reducing aliasing effects caused by insufficient sampling of appearance of objects.

Acknowledgments

Real Images   Synthesized Images

Figure 4: Comparison between real images and synthesized images under complex illumination

References

[1] R. Basri and D. Jacobs, "Lambertian Reflectance and Linear Subspaces," *Proc. IEEE Intl. Conf. Computer Vision 01*, pp.383-389, 2001.

[2] P. Debevec, "Rendering synthetic objects into real scenes: Bridging traditional and image-based graphics with global illumination with high dynamic range photography," *Proc. ACM SIGGRAPH 98*, pp. 189-198, 1998.

[3] J. Driscoll and D. Healy, Jr, "Computing Fourier transforms and convolutions on the 2-sphere," *J. Advanced in Applied Mathematics*, Vol.15, pp. 202-250, 1994.
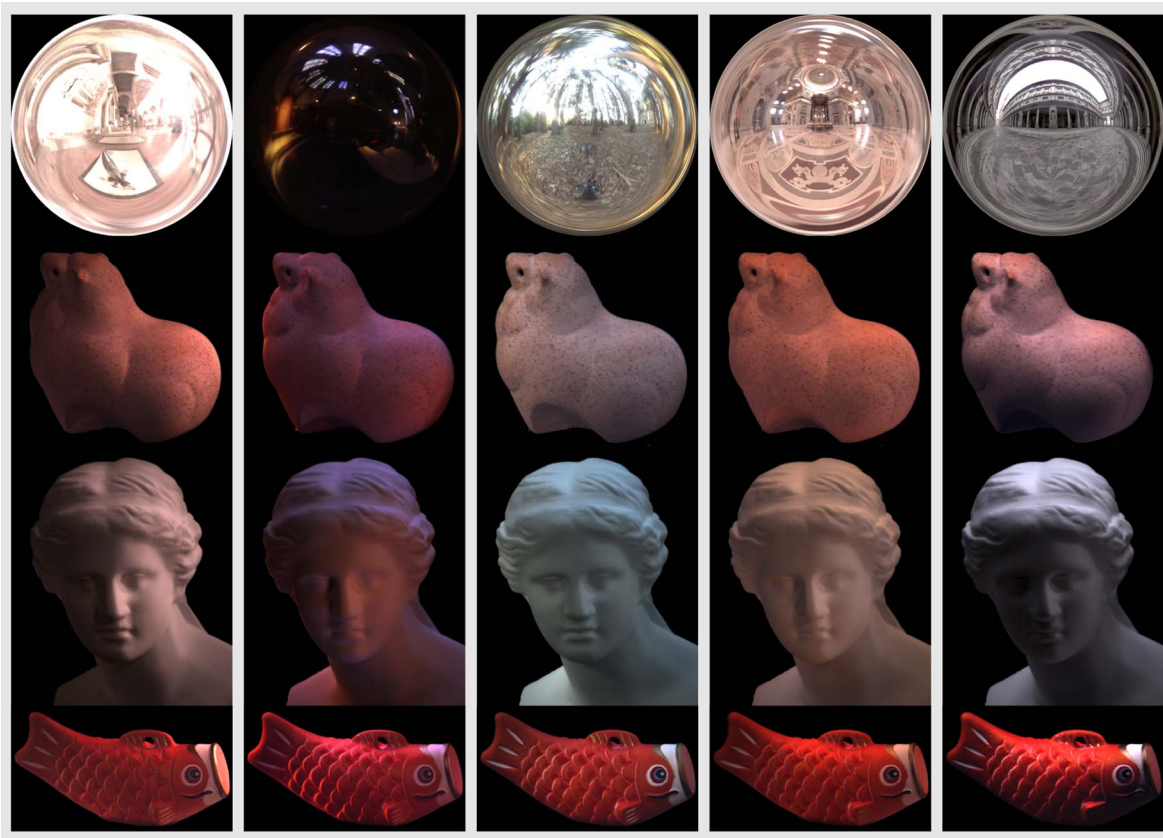
Figure 5: Synthesized images of objects under natural illumination. The first row shows illumination maps. The second, third and forth rows shows synthesized appearance of objects under the corresponding illumination map.

[4] R. Epstein, P. Hallinan, and A. Yuille, "5+/-2 eigenimages suffice: An empirical investigation of low-dimensional lighting models," *Proc. IEEE Workshop on Physics-Based Modeling in Computer Vision*, pp. 108-116, 1995.

[5] 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.

[6] A. Georghiades, D. Kriegman, and P. Belhumeur, "From few to many: Generative models for recognition under variable pose and illumination," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 23, no. 6, pp. 643-660, 2001.

[7] 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.

[8] 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 01*, pp.519-526, 2001.

[9] H. Murase and S. Nayar, "Visual learning and recognition of 3-D objects from appearance," *Int. J. Computer Vision*, vol. 14, no. 1, pp. 5-24, 1995.

[10] R. Ramamoorthi and P. Hanrahan, "A signal-procession framework for inverse rendering", *Proc. ACM SIGGRAPH 01*, pp.117-128, 2001.

[11] R. Ramamoorthi and P. Hanrahan, "On the relationship between radiance and irradiance: Determining the illumination from images of a convex lambertian object," *J. Optical Society of America A*, vol. 18, no. 10, pp. 2448-2459, 2001.

[12] R. Ramamoorthi and P. Hanrahan, "An efficient representation for irradiance environment maps," *Proc. ACM SIGGRAPH 01*, pp. 497-500, 2001.

[13] R. Ramamoorthi, "Analytic PCA construction for theoretical analysis of lighting variability in images of a lambertian object"' *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 24, no. 10, pp. 1322-1333, 2002.

[14] R. Ramamoorthi and P. Hanrahan, "Frequency space environment map rendering"' *Proc. ACM SIGGRAPH 02*, pp. 517-526, 2002.

[15] A. Shashua, "On photometric issues in 3D visual recognition from a single image," *Int. J. Computer Vision*, vol. 21, pp. 99-122, 1997.

[16] P.-P. Sloan, J. Kautz, and J. Snyder, "Precomputed radiance transfer for real-time rendering in dynamic, low-frequency lighting environments," *ACM SIGGRAPH 02*, pp. 527-536, 2002.

[17] 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. J. Computer Vision*, vol. 35, no. 3, pp. 203-222, 1999.

[18] L. Zhao and Y. Yang, "Theoretical analysis of illumination in PCA-based vision systems," *Pattern Recognition*, vol. 32, pp. 547-564, 1999.