Analytic PCA construction for theoretical analysis of lighting variability, including attached shadows, in a single image of a convex Lambertian object

Ravi Ramamoorthi
ravir@graphics.stanford.edu

Stanford University

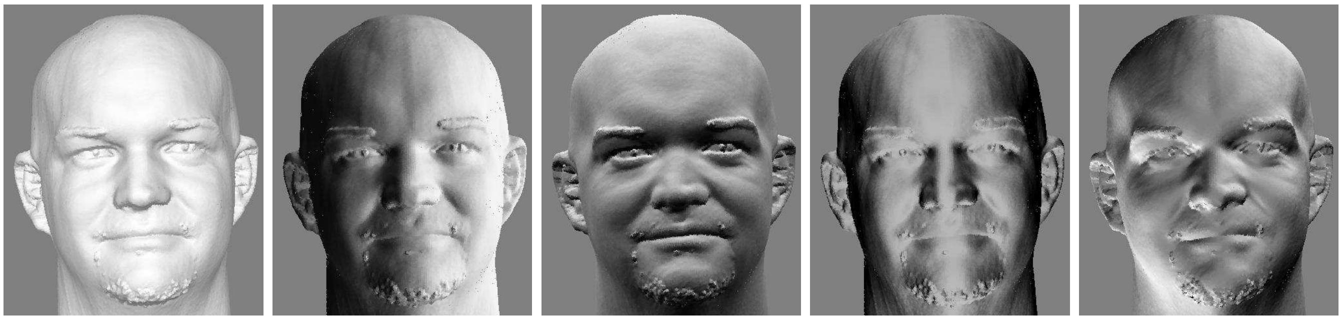

Figure 1: *The first 5 principal components of a face, computed by our method. The form of these eigenmodes is strikingly similar to those derived empirically by previous researchers (see figure 1 in Hallinan [7]). The corresponding eigenvalues are in the ratio of .42, .33, .16, .035, .021 and are in agreement with empirical observation, with the first 5 eigenmodes accounting for 97% of the variance. The principal components contain both positive values (bright regions) and negative values (dark regions), with zero set to the neutral grey of the background. The face is an actual range scan courtesy of Cyberware.*

Abstract

We analyze theoretically the subspace best approximating images of a convex Lambertian object under different distant illumination conditions. Since the lighting is an arbitrary function, the space of all possible images is formally infinite-dimensional. However, previous empirical work by Hallinan [7] and Epstein et al. [4] has shown that images of largely diffuse objects actually lie very close to a 5-dimensional subspace. In this paper, we analytically construct the principal component analysis for images of a convex Lambertian object, explicitly taking attached shadows into account, and find the principal eigenmodes and eigenvalues with respect to lighting variability. Our analysis makes use of an analytic formula for the irradiance in terms of spherical-harmonic coefficients of the illumination [1, 14], and shows, under appropriate assumptions, that the principal components or eigenvectors are identical to the spherical harmonic basis functions evaluated at the surface normal vectors. Our main contribution is in extending these results to the single-image case, showing how the principal eigenmodes and eigenvalues are affected when only a limited subset (the upper hemisphere) of normals is available, and the spherical harmonics are no longer orthonormal over the restricted domain. Our results are very close, both qualitatively and quantitatively, to previous empirical observations and represent the first valid theoretical explanation of these observations. Our analysis is also likely to be of interest in other areas of computer vision and image-based rendering. In particular, our results indicate that using complex illumination for photometric problems in computer vision is not significantly more difficult than using directional sources.

1 Introduction

A robust recognition system must be able to identify an object across variable lighting conditions. This is often a challenging task, since most feature-based methods like edge maps are not robust to large lighting variations [11]. Hence, it is important to derive a good low-dimensional model to explain lighting variability. A classic approach is to consider the principal component analysis (henceforth called *PCA* [8, 18, 19]) of a set of images acquired under different conditions. With respect to lighting variability, however, the space of possible images is infinite-dimensional. Theoretical work by Belhumeur and Kriegman [2] has shown that this is true even when we assume only Lambertian objects under distant illumination—a common simplifying assumption made in many computer vision applications, that we will also make in this paper. Indeed, to represent the illumination exactly, we would need an infinite number of coefficients, corresponding to the intensity for each incident direction.

However, experimental results go against this intuition. Empirical work is reported by Hallinan [7], Epstein et al. [4] and Yuille et al. [20]. They used a human face and other objects, performing PCA on a number of different images acquired by moving a distant point source along a sphere surrounding the object. These classic experiments came to the somewhat counterintuitive conclusion that for largely diffuse objects, the first 5 principal components explain most of the image variation. In other words, images of diffuse objects under varying illumination lie very close to a 5D subspace.

Although these experiments are nearly a decade old, and have been confirmed by other authors [2, 5, 6], there has previously been no satisfactory theoretical explanation. Most previous work ignores the question of visibility, deriving the result that in the absence of shadows, a 3D subspace suffices to describe the set of images of a Lambertian object under distant illumination [12, 16, 21]. This model is too simplistic since it omits attached shadows, which are important, especially when the illumination is complex. Most recently, Basri and Jacobs [1], and Ramamoorthi and Hanrahan [14] have independently derived an analytic formula for the irradiance (and hence, reflected radiance from a convex Lambertian object)

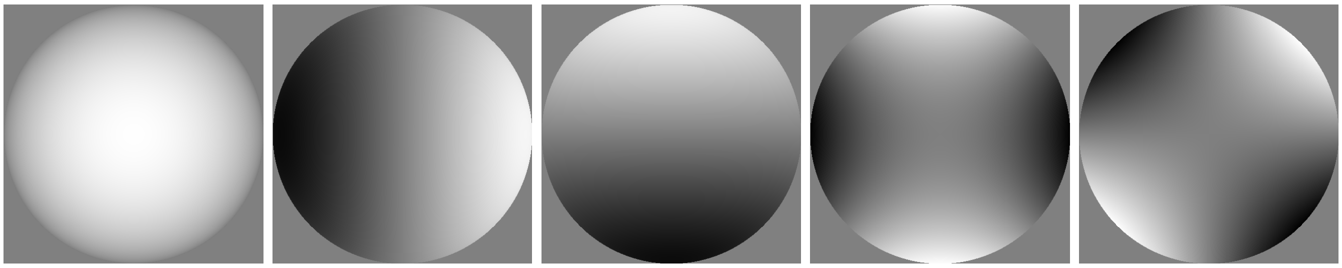

Figure 2: *The first 5 principal components for an image of a sphere, as computed analytically by our method. The corresponding eigenvalues are in the ratio of .43, .24, .24, .023, .023, with the first 5 eigenmodes accounting for 96% of the variance. The principal components contain both positive values (bright regions) and negative values (dark regions), with zero set to the neutral grey of the background.*

under arbitrary distant illumination, explicitly considering attached shadows. They show that the irradiance can be regarded as a *convolution* of the incident illumination with the Lambertian reflectance function (*a clamped cosine*), and express the irradiance in terms of spherical harmonic coefficients of the illumination. A key result of their work is that Lambertian reflection acts as a low-pass filter, so that the irradiance lies very close to a 9D subspace. This model is a significant step towards explaining why images of a diffuse object lie close to a low-dimensional subspace, and provides novel frequency-space tools to understand lighting variability. However, the connection between the spherical harmonic basis and the principal components is not obvious. More importantly, the 9D subspace predicted is somewhat at variance with the 5D subspace observed empirically.

In this paper, we analytically construct the PCA for images of a convex Lambertian object under varying illumination in terms of the spherical harmonic basis functions, uncovering the connections between the spherical harmonic basis and the PCA. We show that under appropriate assumptions, the principal components or eigenvectors are equal to the spherical harmonic basis functions, and the eigenvalues, corresponding to how important a principal component is in explaining image variability, are equal to the spherical harmonic coefficients describing the Lambertian BRDF.

Next, we extend these results even further. Our main contribution is in showing how the eigenvectors (principal components) and eigenvalues are affected by restricting attention to a single image—where we have only the upper hemisphere of normals, instead of the whole sphere of surface orientations. While it is clear qualitatively that restricting the space of surface normals should produce a lower-dimensional approximation, this paper is the first disciplined approach to taking this restriction into account. In particular, we show that the reduced dimensionality stems from the fact that while the spherical harmonics are orthonormal over the entire sphere of normals, this no longer holds over a reduced domain. Besides being of theoretical interest, this analysis gives analytic forms for the modified lighting and appearance basis functions to be used in object recognition. In this way, we extend the 9 parameter Lambertian BRDF model, showing how to modify it when we have only a single viewpoint. Furthermore, the results point to a way of extending other computer vision algorithms to handle complex illumination. Our theoretical predictions, summarized below, are seen to be very close both qualitatively and quantitatively, to the previous empirical observations of Hallinan [7] and Epstein et al. [4].

Orthogonality Matrix: The major mathematical result is that the principal components and associated eigenvalues correspond to the eigenvectors of a scaled version of the *orthogonality matrix*, i.e. a matrix expressing the orthogonality between the spherical harmonic basis functions, when the domain of integration is restricted to the surface normals in the image. If we were considering the entire gaussian sphere of possible normal directions, this matrix would just be the scaled identity matrix. Hence, the eigenvectors or principal components would simply be the spherical harmonic modes. However, if we consider only a single image, the orthogonality matrix is no longer the identity, and the eigenvectors or principal components are linear combinations of the spherical harmonics, with a new eigenvalue spectrum. The orthogonality matrix depends on the distribution of surface normals in the image, and hence depends on object geometry. We are able to derive quantitative results for a number of cases of interest, including images of a human face and a sphere, shown in figures 1 and 2 respectively.

Form of Principal Components: Hallinan [7] characterizes the first 5 principal components for human faces as corresponding to *frontal lighting*, *side lighting*, *lighting from above/below*, *extreme side lighting*, and *lighting from a corner*. Visual inspection of figures 1 and 2 indicates that the principal components predicted by our theoretical model agree well with these labelings. Furthermore, figure 1 looks almost identical to figure 1 in Hallinan [7][1].

Dimensionality of Approximating Linear Subspace: From our derivation, we conclude that 5 eigenvectors or principal components suffice to capture over 95% of the image variance. These results agree with empirical work [7, 4]. In fact, we even have good quantitative agreement. For instance, Epstein et al. [4] report that for an image of a basketball, 3 eigenvectors capture 94% of the variance, while 5 eigenvectors account for 98%. Our corresponding theoretical results for an image of a sphere are 91% and 96%. For a human face, the corresponding empirical numbers [4] for 3 and 5 eigenvectors are 90% and 94% respectively. Our corresponding theoretical predictions are 91% and 97% respectively.

Eigenvalue Spectrum: Hallinan [7] notes that for human faces, the principal components lie in two groups. In a group, the eigenvalues are similar, so the principal components may exchange places. The first group, consisting of two members, corresponds to frontal and side lighting. This is in agreement with our theoretical predictions, illustrated in figure 1, which show these two eigenvalues to be close, having numerical values .42 and .33 respectively, and well separated from the other eigenvalues. The next group consists of the next three principal components. Our theory predicts that the eigenvalues for the fourth and fifth principal components are close (having values .035 and .021), so these may exchange

[1] We would like to propose a correction to his terminology however. Eigenvectors 4 and 5 involve second order or quadratic modes and therefore cannot really be simply classified as corresponding to a single source of illumination. In fact, all the eigenmodes except the first have negative lobes as well as positive and cannot therefore be thought of as corresponding to any physical illumination. Hence, we think it should be emphasized that the principal components do not necessarily correspond to physical lighting conditions since they may take negative values. It may be more appropriate in the future to label the eigenvectors based on how they correspond to the spherical harmonic basis functions.

places, as observed by Hallinan [7]. We predict that the third principal component should always correspond approximately to lighting from above/below, i.e. the Y direction. This appears to be the case in the empirical data also.

2 Mathematics of Analytic PCA Construction

In this section, we give the mathematical details of our approach to analytic PCA construction. We first define the notation and the appropriate matrices. We then show how to reduce the problem to an eigensystem. In the next section, we will compute the eigenmodes for a number of cases of interest. There are two points worth noting about the mathematics. Firstly, our original derivation will be without subtracting out the mean, as opposed to the common practice in PCA analysis. This will make the derivation slightly simpler, and confer some interesting insights. Later, we will redo the analysis with the mean subtracted in order to make a more accurate comparison to empirical work. Secondly, we assume untextured surfaces. This assumption does not significantly affect the validity of our results since we may simply multiply all the principal components by the same texture in order to obtain results for textured objects. However, it should be noted that consideration of texture causes the principal components to have greater amplitude in regions of high albedo. For objects of interest for us, such as human faces, this effect is unlikely to be significant since the low-frequency texture variations are relatively minor.

2.1 Principal Component Analysis

We first define the notation. We will be interested in images $E(\alpha, \beta)$ corresponding to a light source at direction (α, β) where (α, β) are the global spherical coordinates of the distant light source. A single pixel in the image will be denoted by $E(\alpha, \beta, \theta, \phi)$, where (θ, ϕ) stand for the spherical coordinates of the surface normal. Note that in our model of the world (distant illumination, convex Lambertian surfaces, no cast shadows), these four parameters suffice to determine the shading. Also, since the surface is Lambertian, and we are ignoring the albedo, $E(\alpha, \beta, \theta, \phi)$ can be though of as corresponding directly to the irradiance at orientation (θ, ϕ) due to a unit directional source at (α, β).

Assume now that we sample the space of possible image pixels in some way, i.e. by means of some light source positions (α_j, β_j) and some normal coordinates, i.e. (θ_i, ϕ_i). The matrix Q consisting of all observations then has the form

$$Q_{ij} = E(\alpha_j, \beta_j, \theta_i, \phi_i)$$

Here, the rows (depending on index i) denote image observations, while the columns (depending on index j) correspond to different light source positions. This is the standard definition of the matrix Q. In order to find the principal components or eigenimages, we must find the eigensystem of the matrix $T = QQ^T$. The formula for this matrix is given by

$$\begin{aligned} T_{ij} &= \sum_k Q_{ik} Q^T_{kj} = \sum_k Q_{ik} Q_{jk} \\ &= \sum_k E(\alpha_k, \beta_k, \theta_i, \phi_i) E(\alpha_k, \beta_k, \theta_j, \phi_j) \end{aligned} \quad (1)$$

To help make matters concrete, assume that in an experimental situation, there are n image pixels and d light positions. In most experiments, $n > d$. Then, matrix Q will be $n \times d$ while matrix T will be $n \times n$. The eigenvectors of T will correspond to principal component images, while the eigenvalues will measure their importance. We can get an idea of how well the matrix T is approximated by some number of eigenvectors by considering the fraction of the total sum of eigenvalues given by the sum of the eigenvalues corresponding to the eigenvectors in question. It should be noted that in practical applications, it is possible to write down an equivalent $d \times d$ eigensystem that is computationally simpler. However, for analytic PCA construction, we work directly with the matrix T.

To proceed further, we will need a formula for E. Of course, this can be obtained simply by the dot product of the vectors corresponding to (α, β) and (θ, ϕ), yielding the standard expression:

$$E(\alpha, \beta, \theta, \phi) = \max[0, \sin\alpha \sin\theta \cos(\beta - \phi) + \cos\alpha \cos\theta] \quad (2)$$

Note that $E(\alpha, \beta, \theta, \phi)$ here is simply the irradiance at orientation (θ, ϕ) due to a unit directional light source at (α, β).

The angular-space formula above is difficult to manipulate because of the $\max$ expression. Instead, we will use a frequency-space analytic formula for the irradiance derived independently by Ramamoorthi and Hanrahan [14] as well as Basri and Jacobs [1].

$$E(\alpha, \beta, \theta, \phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \hat{A}_l L_{lm}(\alpha, \beta) Y_{lm}(\theta, \phi)$$

In the equation above, Y_{lm} are the spherical harmonics [3, 9, 17]. Spherical harmonics with $l \geq 0$ and $-l \leq m \leq l$ are the appropriate signal processing tools in angular space and are the analog of the conventional Fourier basis. The coefficients L_{lm} are the spherical harmonic coefficients of the incident illumination, while $\hat{A}_l$ is a constant which vanishes for odd $l > 1$ and decays for even l as $l^{-5/2}$. Because of this rapid decay, a very good approximation can be obtained by limiting $l \leq 2$. In fact, 99% of the energy of the BRDF filter, i.e. of $\hat{A}_l$, is contained by $l \leq 2$. The appendix lists values for $\hat{A}_l$ and Y_{lm} for $l \leq 2$.

The final ingredient to complete the frequency-space description is to find the coefficients $L_{lm}(\alpha, \beta)$ in terms of the angular coordinates (α, β) of the light source. Since the light source is described by a delta function, we simply evaluate the spherical harmonics at the light vector, deriving $L_{lm}(\alpha, \beta) = Y_{lm}(\alpha, \beta)$. Note that we use the real form of the spherical harmonics, so there is no need to worry about complex conjugation. Now,

$$E(\alpha, \beta, \theta, \phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \hat{A}_l Y_{lm}(\alpha, \beta) Y_{lm}(\theta, \phi)$$

Plugging this into equation 1, we get

$$\begin{aligned} T_{ij} = \; & \textstyle\sum_{l,m,l',m'} \hat{A}_l \hat{A}_{l'} Y_{lm}(\theta_i, \phi_i) Y_{l'm'}(\theta_j, \phi_j) \cdot \\ & \textstyle\sum_k Y_{lm}(\alpha_k, \beta_k) Y_{l'm'}(\alpha_k, \beta_k) \end{aligned}$$

We now proceed to do the summation over k. Clearly this will depend in practice on the specific sampling pattern used for moving the light source. However, our goal is to get an analytic understanding. It is most reasonable to assume for mathematical purposes that the light source samples are infinitely dense, sampling the sphere of directions equally for all differential solid angles. This assumption also makes physical sense since we wish to make no *a priori* assumptions about the lighting distribution in the scene, so we should assume the illumination is equally likely to come from any direction. The summation over the index k may then be replaced by an integral over the angular coordinates (α, β), and we obtain

$$\begin{aligned} T_{ij} = \; & \textstyle\sum_{l,m,l',m'} \hat{A}_l \hat{A}_{l'} Y_{lm}(\theta_i, \phi_i) Y_{l'm'}(\theta_j, \phi_j) \times \\ & \textstyle\int_{\alpha=0}^{\pi} \int_{\beta=0}^{2\pi} Y_{lm}(\alpha, \beta) Y_{l'm'}(\alpha, \beta) \, \sin\alpha \, d\alpha \, d\beta \end{aligned}$$

It is now straightforward to use orthonormality of the spherical harmonics to set $l = l'$ and $m = m'$ and write

$$T_{ij} = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \left(\hat{A}_l\right)^2 Y_{lm}(\theta_i, \phi_i) Y_{lm}(\theta_j, \phi_j) \tag{3}$$

We have just derived an analytic form for the elements T_{ij} of the matrix used for principal component analysis. This analytic form has been derived under fairly weak assumptions—a convex Lambertian surface under distant illumination, with a uniform probability for the lighting over the entire sphere of incident directions. We now show how to derive the corresponding eigensystem.

2.2 Reduction to Eigensystem

We now wish to find an eigenvector u, which is a vector that satisfies $Tu = \lambda u$ where λ is the eigenvalue. This can also be written as $\sum_j T_{ij} u_j = \lambda u_i$. Note that u corresponds to a principal component image. Now, we expand u in terms of spherical harmonics,

$$u_j = \sum_{p=0}^{\infty} \sum_{q=-p}^{p} c_{pq} Y_{pq}(\theta_j, \phi_j)$$

Our goal is to find the coefficients c_{pq}. Plugging into equation 3, we obtain

$$\begin{aligned} \sum_j T_{ij} u_j &= \sum_{l,m,p,q,j} \left(\hat{A}_l\right)^2 Y_{lm}(\theta_i, \phi_i) Y_{lm}(\theta_j, \phi_j) c_{pq} Y_{pq}(\theta_j, \phi_j) \\ \lambda u_i &= \lambda \sum_{l,m} c_{lm} Y_{lm}(\theta_i, \phi_i) \end{aligned} \tag{4}$$

We require the first and second lines above to be equal. First, we do the summation over the index j in the first line above. Define

$$M_{lm;pq} = \sum_j Y_{lm}(\theta_j, \phi_j) Y_{pq}(\theta_j, \phi_j) \tag{5}$$

The matrix M will be fundamental in the ensuing discussion. It indicates the orthogonality relation between the various spherical harmonics when the domain of integration is taken as the pixels of the image. In the special case where image pixels correspond uniformly to the entire sphere of surface normals, the orthonormality relation for the spherical harmonics will hold and the matrix M will simply be the identity with $M_{lm;pq} = \delta_{lp}\delta_{mq}$. However, this condition is never satisfied in practice since we never see the normals facing away from us. Later, in the next section, we will consider various matrices M and determine the resulting eigenmodes.

In terms of the matrix M, it is straightforward to write

$$\sum_j T_{ij} u_j = \sum_{l,m,p,q} \left(\hat{A}_l\right)^2 M_{lm;pq} c_{pq} Y_{lm}(\theta_i, \phi_i)$$

This expression may be compared to the first line of equation 4. The right-hand side must therefore equal the right-hand side in the second line. Since the spherical harmonics are linearly independent, the coefficients of the Y_{lm} must match, and we obtain

$$\sum_{p,q} \left(\hat{A}_l\right)^2 M_{lm;pq} c_{pq} = \lambda c_{lm} \tag{6}$$

Thus, we have reduced the problem of computing the principal components to an eigenvalue problem involving the matrix M.

To proceed further, we will make a number of notational changes. First, let us collapse the double indices lm and pq into a single index in the standard way, i.e. $r = l^2 + l + m$ and $s = p^2 + p + q$. r and s simply impose an absolute ordering on the spherical harmonics, first on the index l or p, and then on m or q. Finally, define $\hat{M}_{rs} = \left(\hat{A}_r\right)^2 M_{rs}$, where $\hat{A}_r = \hat{A}_l$. It should be noted that while the matrix M is symmetric, the matrix $\hat{M}$ is no longer symmetric, because of the premultiplication factor. It is now straightforward to write equation 6 as

$$\sum_s \hat{M}_{rs} c_s = \lambda c_r \tag{7}$$

But this is simply an eigenvalue problem with $\hat{M}c = \lambda c$. Thus, we have reduced the principal component analysis problem to an eigenvalue problem for $\hat{M}$. However, remember that the matrix $\hat{M}$ is no longer symmetric. It often helps to work with symmetric matrices, since their eigensystems have a number of nice properties. It is straightforward to rescale the matrices and vectors for symmetry. We first define a new symmetric matrix $\tilde{M}$ and vector d by

$$\begin{aligned} c_r &= \hat{A}_r d_r \\ \tilde{M}_{rs} &= \hat{A}_r \hat{A}_s M_{rs} \end{aligned} \tag{8}$$

Now, starting with the basic matrix eigensystem in equation 7, we make the following substitutions:

$$\begin{aligned} & \textstyle\sum_s \hat{A}_r \hat{A}_r M_{rs} c_s && = \lambda c_r \\ \Rightarrow \quad & \textstyle\sum_s \hat{A}_r \left(\hat{A}_r M_{rs} \hat{A}_s\right) d_s && = \lambda \hat{A}_r d_r \\ \Rightarrow \quad & \tilde{M} d && = \lambda d \end{aligned}$$

Thus, we now have a symmetric eigensystem that can be solved for the vectors d. To find the eigenvectors of the original problem, we must find the corresponding vectors c using equation 8. The eigenvalues remain the same.

3 Computation of Eigenmodes

Our goals now are to actually compute the eigenvectors and eigenvalues for various different values of the matrix $\tilde{M}$. However, this matrix is infinite dimensional since both r and s can be arbitrarily large corresponding to all possible spherical harmonic coefficients. However, it has been shown [1, 14] that 99% of the energy of the Lambertian BRDF is captured by $l \leq 2$, i.e. by $r, s < 9$. In other words, only the first 9 spherical harmonic terms are important and there is little error introduced by truncating the series to order 2. Therefore, for numerical calculations in the rest of this section, we will use the 9 term approximation, reducing matrix $\tilde{M}$ to a 9×9 matrix. We analyze the resulting eigensystem for many choices of the matrix $\tilde{M}$. It should be emphasized that the analytic formulas we derive are general, and hold for all values of r and s. It is only the computed numerical values of the eigenvectors and eigenvalues that depend on the 9 term approximation.

Now, we find the principal components for several special cases of interest by considering the corresponding values of $\tilde{M}$ and computing the eigensystem. We first explicitly write down the formula for $\tilde{M}$ by rewriting equation 5 as per our modified notation:

$$\tilde{M}_{rs} = \hat{A}_r \hat{A}_s \sum_j Y_r(\theta_j, \phi_j) Y_s(\theta_j, \phi_j) \tag{9}$$

3.1 Pixels equally distributed over sphere

The first special case of interest is when the pixels j are distributed in such a way that the sum can be replaced with an integral over the entire sphere. As already mentioned, this is unrealistic, but is nevertheless an insightful special case. In that case, orthonormality of the spherical harmonics yields

$$\begin{aligned}\tilde{M}_{rs} &= \hat{A}_r\hat{A}_s\int_{\theta=0}^{\pi}\int_{\phi=0}^{2\pi}Y_r(\theta,\phi)Y_s(\theta,\phi)\sin\theta\,d\theta d\phi \\ &= \hat{A}_r\hat{A}_s\delta_{rs} = \left(\hat{A}_r\right)^2\delta_{rs}\end{aligned}$$

In other words, $\tilde{M}$ is the scaled identity matrix. The eigenvectors or principal components (for both d and c) are simply the spherical harmonics themselves with eigenvalues $\left(\hat{A}_r\right)^2$. The amount of variance accounted for by some number of eigenvectors is simply the sum of the corresponding eigenvalues divided by the sum of all the eigenvalues. This corresponds to the case previously studied by Basri and Jacobs [1], and Ramamoorthi and Hanrahan [14]. As noted by Basri and Jacobs [1], 37.5% of the variance is accounted for by the constant term $l = 0; r, s < 1$, 87.5% of the variance is accounted for when also considering the linear terms $l \leq 1; r, s < 4$, and over 99% of the variance is accounted for when considering the quadratic terms, i.e. all 9 terms in our approximation: $l \leq 2; r, s < 9$. Thus, we see that our formulation agrees with the previous subspace results. It is worth pointing out an important special case or corollary here. Shashua [16] has considered the situation where there are no attached shadows. In this case, the images of a Lambertian object lie exactly in a three-dimensional subspace. We can consider this case in our formulation by removing the threshold to 0 in our definition of the Lambertian BRDF intensity of equation 2. In terms of the above formulation, $\hat{A}_r$ would vanish unless $l = 1$ (the linear terms), i.e. $r = 1, 2, 3$. Thus, the eigenvectors would be exactly the linear spherical harmonics, and as noted by Shashua [16], a 3D subspace would suffice.

3.2 Pixels equally distributed over hemisphere

We now consider a more realistic case. Since only front facing normals are visible in a single image, we allow the pixels to be equally distributed over the hemisphere with $z > 0$. In the standard spherical coordinates this corresponds to $\theta < \pi/2$. The matrix M now encodes the orthogonality relation between the spherical harmonics when the domain of integration is restricted to the upper hemisphere. While linear independence of the spherical harmonics guarantees that no linear combination can have norm 0, we will see that the norm of certain linear combinations comes very close to 0, i.e. most of the norm is concentrated over the unseen lower hemisphere. Thus, these lighting configurations have negligible impact on the irradiance of the upper hemisphere and may be neglected.

It is now straightforward to define

$$\tilde{M}_{rs} = \hat{A}_r\hat{A}_s\int_{\theta=0}^{\pi/2}\int_{\phi=0}^{2\pi}Y_r(\theta,\phi)Y_s(\theta,\phi)\sin\theta\,d\theta d\phi \qquad (10)$$

It should be noted that $\tilde{M}_{rs} \neq 0$ only for terms having the same m index, i.e. those with $m = 0$: Y_0, Y_2, Y_6 (corresponding in two-index notation to Y_{00}, Y_{10}, Y_{20}), those with $m = -1$: Y_1, Y_5 (corresponding to Y_{1-1} and Y_{2-1}), and those with $m = 1$: Y_3, Y_7 (corresponding to Y_{11} and Y_{21}). We have not included $m = \pm 2$ since there is only one term with that value of m (Y_4, Y_8 corresponding to Y_{2-2} and Y_{22} respectively). All other cross terms will vanish. Also, cross terms involving odd ($l + m$ odd) and even ($l + m$ even) spherical harmonics only vanish. Hence, the cross terms between Y_0 and Y_6 vanish. Thus, the effects of rearranging the integration to lie over the hemisphere rather than the full sphere will be the intermingling of the basis functions corresponding to $m = 0$, and those corresponding to $m = 1$ and $m = -1$. Y_4 and Y_8, corresponding to $m = \pm 2$, will remain eigenvectors of $\tilde{M}$ and are not affected by this intermingling.

#	Eigenvector d	Eigenvector c	λ	VAF
1	.85 Y_0 + .53 Y_2 + .03 Y_6	.92 Y_0 + .39 Y_2 + .01 Y_6	.51	.51
2	.95 Y_3 + .31 Y_7	.99 Y_3 + .12 Y_7	.18	.69
3	.95 Y_1 + .31 Y_5	.99 Y_1 + .12 Y_5	.18	.88
4	-.42 Y_0 + .63 Y_2 + .65 Y_6	-.68 Y_0 + .68 Y_2 + .26 Y_6	.05	.93
5	Y_8	Y_8	.023	.95
6	Y_4	Y_4	.023	.98
7	-.31 Y_3 + .95 Y_7	-.66 Y_3 + .75 Y_7	.006	.98
8	-.31 Y_1 + .95 Y_5	-.66 Y_1 + .75 Y_5	.006	.99
9	.32 Y_0 - .57 Y_2 + .76 Y_6	.61 Y_0 - .71 Y_2 + .36 Y_6	.0008	.99

Table 1: *Eigenvectors and eigenvalues for the hemisphere*

It is now simple to compute the matrix $\tilde{M}$ and its eigensystem. We used Mathematica for this purpose. The eigenvectors are normalized to have unit norm and are determined only up to sign. The eigenvalues are normalized to be the percentage of the total sum of the 9 eigenvalues. They thus correspond to the fraction of variance accounted for by that particular eigenvector. We also show the cumulative sum of eigenvalues. This indicates how well the image is approximated using only a number of the most important principal components. To account for components not considered by us, i.e. terms higher than order 2, we multiply the eigenvalues by 0.99, corresponding to the amount of energy captured by the first 9 terms, i.e. modes up to order 2. It should be noted that this is only an approximation. However, exact bounds on the error for any physical lighting distribution can be derived from the fact that the lighting must be everywhere positive, using an approach similar to that of Basri and Jacobs [1].

The four columns in table 1 stand for the eigenvectors d of the matrix $\tilde{M}$, the corresponding eigenvectors c of the original PCA (as per equation 8)—note that these have more energy in the lower frequencies—, the normalized eigenvalues λ, and the cumulative sum of eigenvalues, corresponding to the variance accounted for (VAF).

It can be seen that 98% of the variance is accounted for using only the first 6 principal components or eigenvectors. The observation that fewer than 9 eigenvectors suffice for an accurate approximation is not really surprising. In fact, one can show from first principles that there must be at least one eigenvector with a negligible eigenvalue. Our numerical calculations are based on the observation that the *half-cosine* function can be well approximated using spherical harmonics up to order 2, which allows us to use the 9×9 matrix approximation for numerical work. But this also means that the *backwards half-cosine function*, which has its positive lobe in the lower hemisphere and is zero over the visible upper hemisphere, is well approximated using spherical harmonics up to order 2. This approximation has negligible norm over the visible upper hemisphere. Indeed, we see that the ninth eigenvector c in table 1 corresponds directly to the normalized *backwards half-cosine*, with the corresponding eigenvalue being nearly three orders of magnitude smaller than the largest eigenvalue.

3.3 Image of a sphere

While the previous subsection gave considerable insight, it does not correspond to a realistic situation. This is because, in practice, the number of pixels occupied by regions at oblique angles to the camera is reduced by a cosine factor when projected down into the camera plane. Thus, we should consider that factor when doing the integrations. Hence, we add a factor of $\cos\theta$ in equation 10:

$$\tilde{M}_{rs} = \hat{A}_r \hat{A}_s \int_{\theta=0}^{\pi/2} \int_{\phi=0}^{2\pi} Y_r(\theta,\phi) Y_s(\theta,\phi) \cos\theta \sin\theta \, d\theta d\phi$$

Note that this does not change the azimuthal structure of the matrix M. Thus, our previous discussion regarding the groups for the eigenvectors for different values m continues to hold. The eigensystem is given in table 2 and is largely similar to that in table 1.

#	Eigenvector d	Eigenvector c	λ	VAF
1	.77 Y_0 + .63 Y_2 + .14 Y_6	.88 Y_0 + .48 Y_2 + .04 Y_6	.62	.62
2	.91 Y_3 + .42 Y_7	.99 Y_3 + .17 Y_7	.15	.77
3	.91 Y_1 + .42 Y_5	.99 Y_1 + .17 Y_5	.15	.92
4	-.52 Y_0 + .48 Y_2 + .71 Y_6	-.82 Y_0 + .51 Y_2 + .28 Y_6	.034	.95
5	Y_8	Y_8	.015	.97
6	Y_4	Y_4	.015	.98
7	-.42 Y_3 + .91 Y_7	-.78 Y_3 + .63 Y_7	.004	.99
8	-.42 Y_1 + .91 Y_5	-.78 Y_1 + .63 Y_5	.004	.99
9	.38 Y_0 - .62 Y_2 + .69 Y_6	.65 Y_0 - .70 Y_2 + .30 Y_6	.0004	.99

Table 2: *Eigenvectors and eigenvalues for image of a sphere*

3.4 Image of a face

To more accurately compare our predictions to those of Hallinan [7], we would ideally like to consider the image of a face rather than a sphere. Hence, we took a range scan of a face (courtesy of Cyberware), and raytraced a single frontal image, storing the surface normal, i.e. (θ, ϕ) at each pixel. We could then evaluate the matrix $\tilde{M}$ directly in a numerical fashion using equation 9, by summing over all pixels. The symmetry between side and top directions is now broken because of the asymmetric dimensions of the face. While the azimuthal structure is still nearly preserved, numerically, all elements of the matrix $\tilde{M}$ will be nonzero unlike in the previous subsections. The eigenvectors and eigenvalues are listed in tables 3 and 4. These are somewhat similar to those in table 2, but some important differences are present and will be discussed in the next section. It should be noted the numerical values depend on the specific face model used by us and will differ slightly for other faces.

#	Eigenvector d	Eigenvector c
1	.77 Y_0 + .62 Y_2 + .12 Y_6	.88 Y_0 + .48 Y_2
2	.91 Y_3 + .40 Y_7	.98 Y_3 + .16 Y_7
3	-.10 Y_0 + .89 Y_1 + .44 Y_5	-.15 Y_0+ .97 Y_1 + .18 Y_5
4	-.43 Y_0 + .46 Y_2 + .65 Y_6 -.41 Y_8	-.76 Y_0 + .54 Y_2 + .29 Y_6 - .18 Y_8
5	-.06 Y_3 + .99 Y_4	-.15 Y_3 + .98 Y_4
6	-.24 Y_0 + .16 Y_2 + .30 Y_6 + .90 Y_8	-.68 Y_0 + .30 Y_2 + .20 Y_6 + .63 Y_8
7	-.40 Y_3 + .91 Y_7	-.75 Y_3 + .65 Y_7
8	-.44 Y_1 + .89 Y_5	-.79 Y_1 + .60 Y_5
9	.39 Y_0 - .62 Y_2 + .68 Y_6	.65 Y_0 - .70 Y_2 + .29 Y_6

Table 3: *Eigenvectors for image of a face. For clarity and ease of comparison, only those spherical harmonics with non-negligible coefficient magnitudes (greater than* .10*) are noted.*

4 Results with removal of mean value

One final point in attempting a quantitative comparison with the work of Hallinan [7] and Epstein et al. [4] concerns removal of the mean value before taking the PCA. As is the standard practice when computing principal components, those authors have subtracted the mean value before applying their analysis. We have so far chosen not to do so, because the mathematics is somewhat simpler to explain when retaining the mean value, and we believe there are valuable insights in the previous section. Furthermore, in many applications, the numerical forms of the basis functions derived in the previous section are likely to be useful.

#	λ	VAF
1	.61	.61
2	.21	.82
3	.10	.92
4	.038	.96
5	.013	.97
6	.010	.98
7	.004	.99
8	.002	.99
9	.002	.99

Table 4: *Eigenvalues for image of a face*

However, it is not significantly more difficult to apply our framework with the mean image value removed. In this section, we will extend our results under those conditions. It can be shown that removing the mean simply corresponds to ignoring the constant or ambient term Y_0 or Y_{00}. Alternatively, we may set $\hat{A}_0 = 0$. This makes intuitive sense since the ambient term simply sets the mean value. The same result can also be derived through some simple algebra on the definition of the PCA. We will now numerically use an 8×8 matrix and 8 eigenvectors and eigenvalues, since the mean term will no longer contribute.

We first show the eigenvectors and eigenvalues for the case of an image of a sphere with the constant term subtracted in table 5.

#	Eigenvector d	Eigenvector c	λ	VAF
1	.97 Y_2 + .26 Y_6	.99 Y_2 + .10 Y_6	.43	.43
2	.91 Y_3 + .42 Y_7	.99 Y_3 + .17 Y_7	.24	.67
3	.95 Y_1 + .31 Y_5	.99 Y_1 + .17 Y_5	.24	.91
4	Y_8	Y_8	.023	.94
5	Y_4	Y_4	.023	.96
6	-.26 Y_2 + .97 Y_6	-.59 Y_2 + .81 Y_6	.019	.98
7	-.42 Y_3 + .91 Y_7	-.78 Y_3 + .63 Y_7	.006	.98
8	-.42 Y_1 + .91 Y_5	-.78 Y_1 + .63 Y_5	.006	.99

Table 5: *Eigenvectors and values for image of the sphere with the mean image value subtracted out (i.e. ignoring the ambient term).*

The first three eigenmodes are now very clearly identifiable as frontal, side and top/bottom lighting while the next two are quadratic modes. Three eigenvectors account for 91% of the variance, and 5 eigenvectors for 96%. These results clearly show why the first 5 eigenvectors form a stable group that is a good approximation in a variety of circumstances. Furthermore, our numerical values are quantitatively similar to the empirical results quoted by Epstein et al. [4] for an image of a basketball, wherein 3 eigenvectors captured 94% of the variance, while 5 eigenvectors accounted for 98%. The somewhat better low-dimensional fit obtained by Epstein et al. [4] can be at least partially explained by the details of their experiment. Their lighting conditions sampled only part of the illumination sphere (the top right), while we make the more general assumption of light sources equally sampling the entire sphere of incident directions. It should be noted that the sixth eigenvector, although having eigenvalue similar to eigenvectors 4 and 5, corresponds to a backlighting configuration not well sampled experimentally by Hallinan [7] and Epstein et al. [4].

Finally, we can look at the eigenvectors and values for an image of a face after the mean has been subtracted. As before, the eigenvectors are largely similar to those above, with the specific numerical values dependent on the specific face model used by us.

However, it should be noted that there are some subtle differences between face and sphere eigenvectors, as can be seen by comparing eigenvector 4 (second from right) in figures 1 and 2, or tables 5 and 6. The eigenvectors and eigenvalue spectrum are shown below in tables 6 and 7 as well as figure 3.

#	Eigenvector d	Eigenvector c
1	.15 Y_1 + .95 Y_2 + .24 Y_6	.15 Y_1 + .98 Y_2 + .09 Y_6
2	.91 Y_3 + .40 Y_7	.98 Y_3 + .16 Y_7
3	.88 Y_1 -.17 Y_2 + .44 Y_5	.96 Y_1 -.19 Y_2 + .18 Y_5
4	.09 Y_2 + .13 Y_4 -.58 Y_6 + .80 Y_8	.24 Y_2 + .12 Y_4 -.56 Y_6 + .77 Y_8
5	-.06 Y_3 + .98 Y_4 -.14 Y_8	-.15 Y_3 + .97 Y_4 - .14 Y_8
6	-.23 Y_2 + .77 Y_6 + .58 Y_8	-.53 Y_2 + .67 Y_6 + .50 Y_8
7	-.40 Y_3 + .91 Y_7	-.75 Y_3 + .64 Y_7
8	-.44 Y_1 + .89 Y_5	-.79 Y_1 + .60 Y_5

Table 6: *Eigenvectors for image of a face with the mean image value subtracted out (i.e. ignoring the ambient term). For clarity and ease of comparison, only those spherical harmonics with non-negligible coefficient magnitudes (greater than* .10*) are noted here.*

#	λ	VAF
1	.42	.42
2	.33	.75
3	.16	.91
4	.035	.95
5	.021	.97
6	.011	.98
7	.007	.99
8	.003	.99

Table 7: *Eigenvalues for face image with mean value removed.*

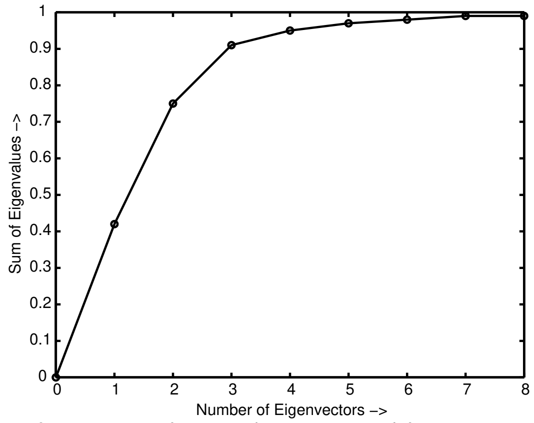

Figure 3: *Fraction of eigenvalues accounted for versus number of eigenvectors or principal components for PCA of a face with mean removed. 5 principal components account for 97% of the variance.*

The significant difference compared to the sphere is that the eigenvalue for side lighting (eigenvalue #2) is substantially higher than that for top/bottom lighting (eigenvalue #3) and is in fact comparable to the eigenvalue for frontal lighting (eigenvalue #1). The spherical symmetry between side and top/bottom (X and Y) directions is broken since human faces are elongated, and are not symmetric about X and Y directions.

These results explain Hallinan's observation that the eigenvectors split into two groups, with the first two eigenvectors corresponding to frontal and side lighting, and the next three to top/bottom lighting and more extreme forms. It does appear from the eigenvalues in table 7 that the eigenvector corresponding to top/bottom lighting should always appear in third position, and the empirical data appear to confirm this. The quantitative predictions for variance accounted for are also in accordance with empirical work. The empirical numbers given by Epstein et al. [4] for the variance accounted for by 3 and 5 eigenvectors in images of a human face are 90% and 94% respectively. Specularity and a small amount of cast-shadowing accounts for the slightly lower numbers as compared to the basketball example. Our corresponding theoretical predictions are 91% and 97% respectively. These are slightly higher than the empirical values because we assume Lambertian surfaces, and do not take cast shadows into account. Finally, figure 1 indicates that the principal components predicted by us are in very good agreement with those from figure 1 in Hallinan [7].

5 Conclusions and Future Work

We have presented a method to analytically construct the principal components for a single view of a convex Lambertian object under varying distant illumination. From this construction, we derive a number of results including good approximating low-dimensional subspaces, the forms of the principal components, and the eigenvalue spectra of the eigenmodes. These results show excellent qualitative and quantitative agreement with previous empirical work.

Besides explaining a number of previous papers on lighting variability, the results are likely to be of considerable interest in computer vision and graphics. It is likely that the principal components derived by us can be used as optimal basis functions in many photometric problems in computer vision such as identifying objects across lighting variability, photometric stereo, and factorization of lighting and texture. In all of these areas, our results indicate that complex illumination is not significantly more difficult to model than point sources, pointing the way to a suite of computer vision algorithms that works under far more general illumination conditions than currently. Indeed, we simply need to consider a few additional basis functions, i.e. 5D or 6D subspaces instead of 3D. Furthermore, our low-dimensional lighting subspaces may also have implications for human vision, and the human ability to identify objects across lighting variability.

Our results are also of theoretical and practical interest in computer graphics for both the forward rendering and inverse rendering problems. For image-based rendering under varying illumination [10], our approach yields optimal basis functions, and easily allows the use of complex illumination. Our work is also a first step in extending the inverse rendering framework of Ramamoorthi and Hanrahan [15] to consider the fundamental limits of what information about the lighting and BRDF can be estimated when the entire reflected light field, corresponding to all surface orientations and outgoing directions, is not available.

In future work, we would like to extend our derivation to consider non-uniform sampling patterns for the light source, corresponding to preferred directions of illumination. We would also like to incorporate specularity in our analysis. From a practical point of view, we would like to implement some of the algorithms proposed here for problems such as photometric stereo, inverse rendering, and image-based rendering.

Finally, we note that this paper has developed a significantly new set of tools for analyzing lighting variability, and we expect these and more advanced methods to be of increasing significance in object recognition and other areas of computer vision.

Acknowledgements

I am grateful to Pat Hanrahan and Steve Marschner for reviewing early drafts of this manuscript and providing many helpful suggestions. The work described in this paper was supported in part by a Hodgson-Reed Stanford graduate fellowship and NSF ITR grant #0085864 "Interacting with the Visual World."

References

[1] R. Basri and D. Jacobs. Lambertian reflectance and linear subspaces. In *ICCV 01*, pages 383–390, 2001.

[2] P. Belhumeur and D. Kriegman. What is the set of images of an object under all possible illumination conditions? *IJCV*, 28(3):245–260, 1998.

[3] B. Cabral, N. Max, and R. Springmeyer. Bidirectional reflection functions from surface bump maps. In *SIGGRAPH 87*, pages 273–281, 1987.

[4] R. Epstein, P. Hallinan, and A. Yuille. 5 plus or minus 2 eigenimages suffice: An empirical investigation of low-dimensional lighting models. In *IEEE Workshop on Physics-Based Modeling in Computer Vision*, pages 108–116, 1995.

[5] A. Georghiades, P. Belhumeur, and D. Kriegman. From few to many: Generative models for recognition under variable pose and illumination. In *Fourth International Conference on Automatic Face and Guesture Recogniton*, pages 277–284, 2000.

[6] A. Georghiades, D. Kriegman, and P. Belhumeur. Illumination cones for recognition under variable lighting: Faces. In *CVPR 98*, pages 52–59, 1998.

[7] P. Hallinan. A low-dimensional representation of human faces for arbitrary lighting conditions. In *CVPR 94*, pages 995–999, 1994.

[8] M. Kirby and L. Sirovich. Application of the Karhunen-Loeve procedure for the characterization of human faces. *PAMI*, 12(1):103–108, Jan 1990.

[9] T.M. MacRobert. *Spherical harmonics; an elementary treatise on harmonic functions, with applications*. Dover Publications, 1948.

[10] T. Malzbender, D. Gelb, and H. Wolters. Polynomial texture maps. In *SIGGRAPH 01*, pages 519–528, 2001.

[11] Y. Moses, Y. Adini, and S. Ullman. Face recognition: the problem of compensating for changes in illumination direction. In *ECCV 94*, pages 286–296, 1994.

[12] H. Murase and S. Nayar. Visual learning and recognition of 3-D objects from appearance. *IJCV*, 14(1):5–24, 1995.

[13] R. Ramamoorthi and P. Hanrahan. An efficient representation for irradiance environment maps. In *SIGGRAPH 01*, pages 497–500, 2001.

[14] R. Ramamoorthi and P. Hanrahan. On the relationship between radiance and irradiance: Determining the illumination from images of a convex lambertian object. *To appear, Journal of the Optical Society of America A*, 18(10):2448–2459, Oct 2001.

[15] R. Ramamoorthi and P. Hanrahan. A signal-processing framework for inverse rendering. In *SIGGRAPH 01*, pages 117–128, 2001.

[16] A. Shashua. On photometric issues in 3D visual recognition from a single 2D image. *IJCV*, 21:99–122, 1997.

[17] F. X. Sillion, J. Arvo, S. H. Westin, and D. Greenberg. A global illumination solution for general reflectance distributions. In *SIGGRAPH 91*, pages 187–196.

[18] L. Sirovich and M. Kirby. Low-dimensional procedure for the characterization of human faces. *JOSA A*, 4(3):519–524, Mar 1987.

[19] M. Turk and A. Pentland. Eigenfaces for recognition. *Journal of Congitive Neuroscience*, 3(1):71–96, 1991.

[20] 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. *IJCV*, 35(3):203–222, 1999.

[21] L. Zhao and Y. Yang. Theoretical analysis of illumination in PCA-based vision systems. *Pattern Recognition*, 32:547–564, 1999.

Appendix: Spherical Harmonics

Below, we list the values of the constants $\hat{A}_r$ as well as the real forms of the first 9 spherical harmonics, used for computation, in both single and double index form. The spherical harmonics are simply constant linear and quadratic polynomials of the Cartesian components of the surface normal [13]. However, it will be convenient here to write them in trigonometric form.

$$
\begin{aligned}
\hat{A}_0 = &\quad \pi &&: l = 0 \\
\hat{A}_1, \hat{A}_2, \hat{A}_3 = &\quad 2\pi/3 &&: l = 1 \\
\hat{A}_4, \hat{A}_5, \hat{A}_6, \hat{A}_7, \hat{A}_8 = &\quad \pi/4 &&: l = 2
\end{aligned}
$$

$$
\begin{aligned}
Y_0 = Y_{00}(\theta, \phi) &= \sqrt{\frac{1}{4\pi}} \\
Y_1 = Y_{1-1}(\theta, \phi) &= \sqrt{\frac{3}{4\pi}} \sin\theta \sin\phi \\
Y_2 = Y_{10}(\theta, \phi) &= \sqrt{\frac{3}{4\pi}} \cos\theta \\
Y_3 = Y_{11}(\theta, \phi) &= \sqrt{\frac{3}{4\pi}} \sin\theta \cos\phi \\
Y_4 = Y_{2-2}(\theta, \phi) &= \sqrt{\frac{15}{4\pi}} \sin^2\theta \cos\phi \sin\phi \\
Y_5 = Y_{2-1}(\theta, \phi) &= \sqrt{\frac{15}{4\pi}} \sin\theta \cos\theta \sin\phi \\
Y_6 = Y_{20}(\theta, \phi) &= \sqrt{\frac{5}{16\pi}} \left(3\cos^2\theta - 1\right) \\
Y_7 = Y_{21}(\theta, \phi) &= \sqrt{\frac{15}{4\pi}} \sin\theta \cos\theta \cos\phi \\
Y_8 = Y_{22}(\theta, \phi) &= \sqrt{\frac{15}{16\pi}} \left(\sin^2\theta \left(\cos^2\phi - \sin^2\phi\right)\right)
\end{aligned}
$$