On Reducing the Complexity of Illumination Cones for Face Recognition *

Jeffrey Ho † Kuang-Chih Lee ‡ David Kriegman §

Beckman Institute and Computer Science Department
University of Illinois at Urbana-Champaign
Urbana, IL 61801

Abstract

In our previous work [1], we have demonstrated how to arrange physical lighting so that the acquired images of a human face can be directly used as the basis vectors of a low-dimensional linear space. The "universal configuration" computed in [1] has been shown to provide a good basis for face recognition under a wide range of lighting conditions. In this paper, we propose a more rigorous formulation of the problem of computing the universal configuration. We propose a new objective function Ob *and define the domain of* Ob*, in which we seek our universal configuration. Recognition experiments and other results reported in the paper have validated this new approach.*

1 Introduction

It is known that the set of n-pixel images of any convex Lambertian surface in fixed pose under all lighting conditions is a polyhedral cone, the illumination cone, in the image space $\mathbb{R}^n$ [2]. Therefore, the illumination cone contains all the image variations of an object (with a fixed pose), and it provides a powerful tool for recognizing objects across a wide range of illumination variations. Indeed, successful work on applying this theory to face recognition has been reported, e.g. [3]. For most objects, the exact illumination cone is very difficult to compute due to the large number of extreme rays that make up the cones, e.g. for a convex, Lambertian surface, there are $O(n^2)$ extreme rays. This complicates both quantitative and qualitative studies of the illumination cones.

However, several recent results have indicated that, although it provides a theoretical basis for discussions on illumination problems, the computation of the full illumination cone may not be necessary. Using spherical harmonics and techniques from signal-processing, Basri and Jacobs have shown that for a convex Lambertian surface, its illumination cone can be accurately approximated by a 9-dimensional linear subspace, the harmonic plane [4, 5, 6]. The major contribution of their work is to treat Lambertian reflection as a convolution process between two spherical harmonics representing the lighting condition and the Lambertian kernel. By observing that the Lambertian kernel contains only low-frequency components, they deduce that the first nine (low frequency) spherical harmonics capture more than 99% of the reflection energy. Using this nine-dimensional harmonic plane, a straightforward face recognition scheme can be developed and results obtained in [4] are excellent. One drawback of using the harmonic plane is that the surface normals and albedos must be known before the harmonic plane can be computed. In our previous work [1], by exploiting the relationship between the illumination cone and the harmonic plane, we demonstrated that there exists a configuration (the universal configuration) of nine point light source directions such that by taking nine images of each individual under these single sources, the resulting linear subspace R spanned by these images is effective at recognition under a wide range of lighting conditions. That is, it is not necessary to know the surface normals or albedos before computing R.

Although the recognition results reported in [1] are good, the way that this universal configuration was computed is somewhat ad hoc. In this paper, we delve deeper into the computational aspect of our earlier work by proposing a more systematic formulation of the problem of computing the universal configuration. We propose a new objective function Ob and define the domain of Ob, in which we seek our universal configuration. We validate this new approach by explicitly computing Ob for a simplified problem. Surprisingly, the results we obtained for this simplfied problem echo the earlier results of [7].

This paper is structured as follows. In the next section, we briefly summarize the ideas of illumination cone [2] and harmonic lighting [4, 5]. The description of our main al-

*Support was provided by the National Science Foundation EIA 00-04056 and CCR 00-86094, National Institute of Health R01-EY 12691-01 and Honda Fundamental Research Laboratory

†j-ho1@uiuc.edu
‡klee10@uiuc.edu
§kriegman@uiuc.edu

gorithms and the report of our experiment results are the subjects of Section 3 and 4, respectively. The final section contains a brief summary and conclusion of this paper.

2 Preliminaries

2.1 Illumination Cone

Let $\mathbf{x} \in \mathbb{R}^n$ denote an image with n pixels of a convex object with a Lambertian reflectance function illuminated by a single point source at infinity, represented by a vector $\mathbf{s} \in \mathbb{R}^n$ such that its magnitude $|\mathbf{s}|$ represents the intensity of the source and the unit normal $\mathbf{s}/|\mathbf{s}|$ represents the direction. Let $B \in \mathbb{R}^{n\times 3}$ be a matrix where each row $\mathbf{b}(\mathbf{x}, \mathbf{y})$ in B is the product of the albedo with unit normal for a point on the surface projecting to a particular pixel in the image. Under the Lambertian assumption, $\mathbf{x}$ is given by

$$\mathbf{x} = \max(\mathrm{Bs}, 0), . \tag{1}$$

where $\max(Bs, 0)$ sets to zero all negative components of the vector Bs. If the object is illuminated by k light sources at infinity, then the image is given by the superposition of the images which would have been produced by the individual light sources, i.e.

$$\mathbf{x} = \sum_{\mathrm{i}=1}^{\mathrm{k}} \max(\mathrm{Bs_i}, 0), . \tag{2}$$

Due to this superposition, the set of all possible images C of a convex Lambertian surface created by varying the direction and strength of an arbitrary number of point light sources at infinity is a convex cone. Furthermore, any image in the illumination cone C (including the boundary) can be determined as a convex combination of extreme rays (images) given by

$$\mathbf{x_{ij}} = \max(\mathrm{Bs_{ij}}, 0) \tag{3}$$

where $s_{ij} = \mathbf{b_i} \times \mathbf{b_j}$ are rows of B with $i \neq j$. It is clear that there are at most $m(m-1)$ extreme rays for $m \leq n$ distinct surface normals [3].

In computer vision, it has been a customary practice to treat the human face as a Lambertian surface. Although human faces are not convex, the degree of non-convexity is not serious enough to render the concept of illumination cone inapplicable [3]. The only difference between the illumination cone of a human face and a convex object is that Equation 3 no longer accounts for all the extreme rays and there are extreme rays which are the result of cast shadows. Therefore, the formula for the upper bound on the number of extreme rays is generally more complicated than the quadratic expression $m(m-1)$ above. This poses a formidable difficulty for computing the exact illumination cone (i.e. specifying all the extreme rays). Instead, a subsampled illumination cone is always computed by sampling lighting directions on the unit sphere, and Equation 1 is accompanied by a ray tracing to account for the cast shadows.

2.2 Lambertian Reflection and Spherical Harmonics

In this section, we briefly summarize the recent work presented in [4, 5, 6]. Consider a convex Lambertian object with uniform albedo illuminated by distant isotropic light sources, and p is a point on the surface of the object. Pick a local (x, y, z) coordinates system F_p centered at p such that the z-axis coincides with the surface normal at p, and let (θ, ϕ) denote the spherical coordinates centered at p. Under the assumption of distant and isotropic light sources, the configuration of lights that illuminate the object can be expressed as a non-negative function $L(\theta, \phi)$. The reflected radiance at p is given by

$$\begin{aligned} r(p) &= \lambda \int\int_S k(\theta) L(\theta, \phi) dA \\ &= \lambda \int_0^{2\pi} \int_0^{\pi} k(\theta) L(\theta, \phi) sin\theta d\theta d\phi \end{aligned} \tag{4}$$

with λ the albedo and $k(\theta) = \max(\cos\theta, 0)$, the Lambertian kernel. A similar integral can be formed for any other point q on the surface to compute the reflected radiance $r(q)$. The only difference between the integrals at p and q is the lighting function L: at each point, L is expressed in a local coordinate system (or coordinate frame F_p) at that point. Therefore, considered as a function on the unit sphere, L_p and L_q differ by a rotation $g \in SO(3)$ that rotates the frame F_p to F_q. That is, $L_p(\theta, \phi) = L_q(g(\theta, \phi))$.

The spherical harmonics are a set of functions that form an orthonormal basis for the set of all square-integrable (L^2) functions defined on the unit sphere. They are the analogue on the sphere to the Fourier basis on the line or circles. The spherical harmonics, Y_{lm}, are indexed by two integers l and m obeying $l \geq 0$ and $-l \leq m \leq l$:

$$Y_{lm}(\theta, \phi) = N_{lm} P_l^{|m|}(cos(\theta)) e^{im\phi} \tag{5}$$

where N_{lm} is a normalization factor guaranteeing that the integral of $Y_{lm} * Y_{l'm'} = \delta_{mm'}\delta_{ll'}$, and $P_l^{|m|}$ is the associated Legendre functions (its precise definition is not important here; however, see [8]). In particular, there are nine spherical harmonics with $l < 3$. One significant property of the spherical harmonics is that the polynomials with fixed l-degree form an irreducible representation of the symmetry group $SO(3)$, that is, a rotated harmonic is the linear superposition of spherical harmonics of same l-degree. For a 3D rotation $g \in SO(3)$:

$$Y_{lm}(g(\theta, \phi)) = \sum_{n=-1}^{l} g_{mn}^{l} Y_{ln}(\theta, \phi). \tag{6}$$

The coefficients g_{nm}^l are real numbers and determined by g.

Expanding the Lambertian kernel $k(\theta)$ in terms of Y_{lm}, one has $k = \sum_{l=0}^{\infty} k_l Y_{l0}$. Because $k(\theta)$ has no ϕ-dependency, its expansion has no Y_{lm} components with $m \neq 0$. An analytic formula for k_l was given in [4, 5]. It can be shown that k_l vanishes for odd values of $l > 1$, and the even terms fall to zero rapidly; in addition, more that 99% of the L^2-energy of $k(\theta)$ is captured by its first three terms, those with $l < 3$. Because of these numerical properties of k_l, by Equation 4, any high-frequency ($l > 2$) component of the lighting function $L(\theta, \phi)$ will be severely attenuated. That is, the Lambertian kernel acts as a low-pass filter. Therefore, for a smooth lighting function L, the result of computing reflected radiance using Equation 4 can be accurately approximated by the same integral with L replaced by L', obtained by truncating the harmonic expansion of L at $l > 2$. Since rotations preserve the l-degree of the spherical harmonics (rf. Equation 6), the same truncated L' will work at every surface point.

2.3 Harmonic Images

From the above discussion, it follows that the set of all possible images of a convex Lambertian object under all lighting conditions can be well approximated by nine 'harmonic images', 'images' formed under lighting conditions specified by the first nine spherical harmonics. Except for the first spherical harmonic (which is a constant), all others have negative values and therefore, they do not correspond to real lighting conditions. The corresponding 'harmonic images' are not real images and as pointed out by [4]: "they are abstractions." Knowing the object's geometry and albedos, these harmonic images can be synthesized using standard techniques, such as the ray-tracing.

For spherical harmonics, the spherical coordinates θ, ϕ are a little bit complicated to work with. Instead, it is usually convenient to write Y_{lm} as a function of x, y, z rather than angles. Each spherical harmonic $Y_{lm}(x, y, z)$ expressed in terms of (x, y, z) is a polynomial in (x, y, z) of degree l. The first nine spherical harmonics in the Cartesian coordinates are

$$Y_{00} = 0.2821; \tag{7}$$

$$(Y_{11}; Y_{10}; Y_{1-1}) = 0.4886(x; y; z); \tag{8}$$

$$(Y_{21}; Y_{2-1}; Y_{2-2}) = 1.093(xz; yz; xy); \tag{9}$$

$$Y_{20} = 0.3154(3z^2 - 1); \tag{10}$$

$$Y_{22} = 0.5462(x^2 - y^2); \tag{11}$$

Figure 1 shows the rendered harmonic images for a face taken from the Yale Database. These synthetic images are rendered by sampling 4098 rays on the unit sphere, and the final images are the weighted sum of these 4098 ray-traced images. Unlike [4] which only accounted for attached shadows, these harmonic images also include the effects of cast shadows arising from non-convex surfaces. Therefore, all nine harmonic images contain 3D information (i.e., the shadows) of the face. The values of the spherical harmonics at a particular point is computed easily using Equations 7–11.

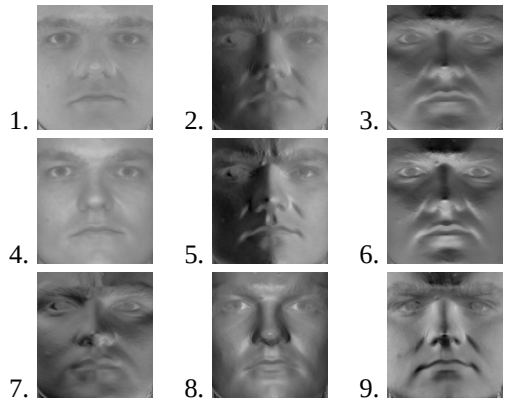

Fig. 1: The nine simulated harmonic images of a face from Yale Database. The light gray and dark gray indicate the positive and negative pixel values. Since the Y_{00} is a constant, the corresponding harmonic image simply scales the albedo values as shown in Picture 1. Pictures 4 is the harmonic image corresponds to $Y_{1-1} = z$, which gives positive values for all pixels. Here, the image plane is defined as the xy-plane.

2.4 Motivations

Equations 1 and 2 can be interpreted geometrically as follows. Let O denote a particular face. O defines a continuous map L_O from S^2 to the image space $\mathbb{R}^n$:

$$L_O(p) = I_O^p, \tag{12}$$

where I_O^p denote the image of O under a point light source specified by the direction $p \in S^2$ with a fixed intensity. The illumination cone C_O is simply the convex hull of the set of rays generated by the image of L_O. The work of Basri and Jacob [4] shows that there exists a good linear approximation of C_O by the harmonic plane H. From its very definition, H can be considered as intrinsic to C_O, since both are completely determined by the B matrix. It is then natural to study the relation between H and C_O; and, in particular, the relation between the continuous map L_O and H. Specifically, we are interested in how the linear space H can be best approximate by vectors in the image of L_O. If a reasonable approximation R of H can indeed be generated by the image of L_O, R can be obtained by taking the images of O under a configuration C_o of nine point sources; therefore,

no information on surface normals or albedos is needed. In the next section, we show that such an R does indeed exist.

The map L_O can be extended to a continuous universal map $\mathbb{L}$ such that $\mathbb{L}$ maps $S^2 \times \mathbb{F}$ to the image space defined as (for $O \in \mathbb{F}$ and $p \in S^2$),

$$\mathbb{L}(\mathrm{O}, \mathrm{p}) = \mathrm{I}^{\mathrm{p}}_{\mathrm{O}}, \tag{13}$$

where $\mathbb{F}$ is the "space of faces". We assume that a distance function can be defined on $\mathbb{F}$ such that for two similar faces o and o', o is close to o' in $\mathbb{F}$. The continuity of the universal map $\mathbb{L}$ would imply that the sets $\mathbb{L}(\mathrm{o}, \mathrm{S}^2)$ and $\mathbb{L}(\mathrm{o}', \mathrm{S}^2)$ should have similar shape and share similar intrinsic properties for neighboring o and o'. In particular, the configurations C_o and C'_o, considered as subsets of S^2, should be very close. This allows us, instead of computing the configuration C_o and $C_{o'}$ separately, to "transplant" the configuration C_o directly to the illumination cones of its neighbors such that the resulting linear space is still a good approximation for the illumination cone of o' "sufficiently" close to o. In Section 4, we show qualitatively that this is indeed the case by using the configuration computed for one person as the fixed configuration and computing all the transplanted linear spaces for different faces. We demonstrate that these "transplanted" linear spaces are good approximations of their respective illumination cones by showing good recognition results using these "transplanted" linear spaces.

3 Linear Approximation of Illumination Cone

In this section, we detail our algorithm for computing R. Given a model (human face), we assume that we have the detailed knowledge of its surface normals and albedos. Using the methods outlined in the previous section, we can construct its harmonic plane H. Let C and EC denote the model's illumination cone and the set of (normalized) extreme rays in the cone, respectively. By a normalized extreme ray, we mean the unique point on the extreme ray with magnitude 1. For notational reason, we will not make any distinction between a (normalized) extreme ray (which is an image) and the direction of the corresponding light source; therefore, depending on the context, an element of EC can denote either an image or a direction.

3.1 Computing the Linear Subspace R

Since R is meant to provide a good approximation of the illumination cone, it is reasonable for R to satisfy the following two conditions:

1. R should be close to H
2. The normalized volume $C \cap R$ should be large.

Since H has been shown to be a good approximation of C, it is reasonable to assume that any subspace close to H would likewise be a good approximation of C; hence the first condition. The second condition is also intuitively clear, since one prefers R to contain as many real images as possible. Our goal and challenge here is to formulate a tractable computational problem for computing R starting with these two simple conditions. Our first task is to define the domain in which R will be found and the second task is to make the two conditions above mathematically precise.

Let $GR(n, 9)$ denote the space of 9-dimensional linear subspaces of $\mathbb{R}^n$, the Grassmannian. Let $\mathbb{D}$ denote the subset of $GR(n, 9)$ consisting of 9-dimensional linear subspaces generated by the extreme rays of C. We define $\mathbb{D}$ to be our domain. Unlike $GR(n, 9)$, the space $\mathbb{D}$ is discrete and contains at most $C(e, 9)$ points e is the number of extreme rays. While the harmonic plane H is generally in the set $GR(n, 9) \backslash \mathbb{D}$, the linear space R that we are after always resides in $\mathbb{D}$. To extend our notation a bit further, we define, for any k, $\mathbb{D}^{\mathrm{k}}$ as the subset of $GR(n, k)$ consisting of k-dimensional linear subspaces generated by the extreme rays of C. Clearly, $\mathbb{D}^9 \equiv \mathbb{D}$ and $\mathbb{D}^1 \equiv \mathrm{EC}$

3.2 How to compute the volume $R \cap C$?

Suppose the vectors in EC are all linearly independent and R is generated by the extreme rays $\{v_1, \cdots, v_9\}$. It is then simple to show that

Proposition 3.1 *With the assumption above, the intersection $R \cap C$ is simply the cone R_C in R generated by $\{v_1, \cdots, v_9\}$.*

Proof The proof is simple. Since if there exists a $x \in R \cap C$ but $x \notin R_C$, x can be written in terms of a collection of extreme rays $\{u_1, \cdots, u_k\}$, which is different from the collection $\{v_1, \cdots, v_9\}$:

$$x = \sum_{i=1}^{k} a_i u_i = \sum_{i=1}^{9} b_i v_i \tag{14}$$

where a_i are all non-negative. This immediately implies a non-trivial linear relation $\sum_{i=1}^{k} a_i u_i + \sum_{i=1}^{9} b_i v_i = 0$. Hence x must be in R_C. ∎

We define the normalized volume $R \cap C$ as the volume of the the hyper-cube R_S generated by $\{v_1, \cdots, v_9\}$, that is $R_S = \{x \in R | x = \sum a_i v_i, 0 \leq a_i \leq 1\}$.

The definition works exactly when the vectors in EC are linearly independent. Since the number of extreme rays is usually considerably larger than the dimension of the image space, there are many non-trivial linear relations between elements of EC. Nevertheless, we will still define the normalized volume of $R \cap C$ as the volume of the simplex R_S

in this case. There are two reasons for this. First, the definition is very simple and can be easily evaluated. Second, in every computation involving EC e.g [3], EC is always replaced by a set of k images such that $k \ll n$. In this case, the elements of EC are almost always linearly independent.

The volume of R_S can be computed as the determinant of $\{v_1, \cdots, v_9\}$, when v_i are expressed in some orthonormal frame. More precisely, applying the Gram-Schmidt to $\{v_1, \cdots, v_9\}$, let Q be the resulting matrix with orthonormal columns; and let V be the matrix with columns v_i, the volume of R_S is then computed as

$$vol(R_S) = |\det(V^t Q)| \tag{15}$$

The absolute value is necessary when V and Q have different orientations and give negative determinant values.

3.3 How to compute the distance between R and H?

Since both R and H are in $GR(n, 9)$, one can use the natural notion of distance on $GR(n, 9)$. Being a homogeneous space of a compact Lie group $O(n)$, the orthogonal group $O(n)$ naturally induces a metric and hence a notion of distance on $GR(n, 9)$. However, we propose a distance measure on $GR(n, 9)$ that ties well with both conditions above. Let P_H denote the orthogonal projection onto the harmonic plane H in the image space $\mathbb{R}^n$. In a nutshell, our distance measure records the effect of the projection P_H on the volume of R_S. If R_S^H is the projection of R_S onto the harmonic plane H, we consider R to be far away from H if P_H shrinks the volume of R_S significantly and on the other hand, R is consider to be close to H if P_H does not have significant effect on the volume of R_S. Note that being an orthogonal projection, P_H never increases the projected volume. Hence $vol(R_S) \geq vol(R_S^H)$.

Therefore, we define a distance measure $d(R, H)$ between R and H as

$$d(R, H) = \arctan(\frac{vol(R_S)}{vol(R_S^H)}) - \frac{\pi}{4} \tag{16}$$

Since $\frac{vol(R_S)}{vol(R_S^H)} \geq 1$, $d(R, H)$ is always non-negative. Notice that in the definition of $d(R, H)$, it is not necessary for R to be of dimension 9. We let $d^i(R, H)$ denote the (extended) distance measure when R has dimension i.

With both the volume and distance defined, we can define an objective function Ob on $\mathbb{D}$ as

$$\mathrm{Ob(R)} = \mathrm{vol(R_S)/d(R, H)}. \tag{17}$$

We define R as the plane in $\mathbb{D}$ at which Ob takes its maximal values

$$R \equiv \arg \max_{R_k \in \mathbb{D}} \mathrm{Ob(R_k)} \tag{18}$$

With our definitions, R always exists since the set $\mathbb{D}$ is finite. Again, the objective function $\mathrm{Ob(R)}$ can be extended to linear spaces R with dimension different from 9: $\mathrm{Ob^i} \equiv \mathrm{vol(R_S)/d_i(R, H)}$ is a function defined on $\mathbb{D}^i$. With this definition, we have

Proposition 3.2 *The maximum of* Ob^1 *occurs at* $x \in EC$ *such that the* L^2*-distance between* x *and* H *is minimal among the elements of EC.*

Proof The proof is a simple walk through various definitions. By definition, Ob^1 is a function on EC. Since $x \in EC$ has magnitude 1, $vol(x_S) = 1$. Therefore, if x maximizes Ob^1, x must minimize $d_1(x, H)$. This implies $vol(x_S^H)$ is maximal. This condition says that the magnitude of the projection of x on H is maximal among elements of EC. By Pythagorean theorem, the L^2-distance between x and H must be minimal. ∎

3.4 Discussion

To satisfy Condition One, it is tempting to find the nine extreme rays which are closest to H and define R as the linear space spanned by these rays. We have observed that these nine extreme rays are generally clustered around the direct frontal direction and the resulting linear space R is a poor approximation of the illumination cone. The explanation, according to Condition Two, is because the resulting intersection $R \cap C$ has small volume. Geometrically, using nearby (with respect to H) rays is no guarantee that R will be a good approximate of H. Indeed, one can easily create a counter-example in dimensional three showing the peril of choosing nearby rays for this purpose, and the situation becomes trickier when R has a large co-dimension (which is our case).

On the other hand, the collection of images which are produced by extreme lighting conditions (lighting from the sides, up/down or behind) generally produce large intersection volume $C \cap R$, see Figure 2. This is because, these images are already (or close to being) orthonormal: the sets of pixels illuminated in each image are disjoint. Therefore, the volume $vol(R_S)$, according to our definition, will be close to the maximal possible value of 1. However, the resulting intersection $R \cap C$ is only on the *boundary* of C and does not contain the interior of C. To correct these pathological cases, we need Condition One to "pull the plane inside". Heuristically, the first condition favors lighting directions which are nearly frontal while the second condition favors extreme lighting conditions. In this sense, the two conditions above complement each other.

3.5 Solving R Iteratively

Having stated that the solution R always exists does not necessarily guarantee that R can be found easily. In fact,

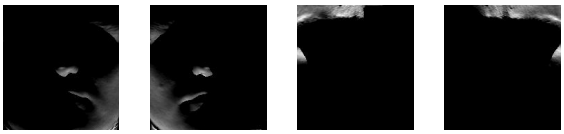

Fig. 2: If R is the plane generated by these four images, the intersection volume $vol(R_S)$ is very close to 1. Note that these four images are almost mutually orthogonal in the sense that their mutual L^2-inner product is very close to 0. This is quite obvious since their non-shadow parts almost never intersect.

only when the size of $\mathbb{D}$ is sufficiently small, we can compute R directly from Equation 18 by enumerating every element of $\mathbb{D}$. In [2], it was shown that the number of extreme rays e is $m(m-1)$ where m is the number of distinct surface normals - m is typically greater than 1000. To further complicate the matter, the size of $\mathbb{D}$ is in the order of $O(e^9)$. Therefore, in most cases, the direct application of Equation 18 is not viable. We propose a straightforward approach, using a greedy algorithm, to compute a good (but less than optimal) solution R. We compute a R as a sequence of nested linear subspaces $R_0 \subseteq R_1 \subseteq \ldots \subseteq R_i \cdots \subseteq R_9 = R$ with $R_i, i > 0$ a linear subspace of dimension i and $R_0 \equiv \emptyset$ as follows. First, we let EC_i denote the set obtained by deleting i extreme rays from EC. It follows that $EC_0 = EC$. We will define R_i and EC_i inductively. Assume that R_{i-1} and EC_{i-1} have been defined (or computed). The sets EC_i and R_i are defined iteratively as follows:

Let x_i denote the element in EC_{i-1} such that

$$x_i = \arg \max_{x_k \in EC_{i-1}} \mathrm{Ob^i(R_k)} \tag{19}$$

R_k is defined as the space spanned by x_k and R_{i-1}. The space R_i is defined as the space spanned by x_i and R_{i-1}, and the set EC_i is defined as $EC_{i-1} \backslash x_i$. The algorithm terminates after $R_9 \equiv R$ is computed. Note that by Proposition 3.2, the first element $x_1 = \arg\max_{x_k \in EC_0} \mathrm{Ob^1(R_k)}$ is simply the extreme ray in EC that is closest to the harmonic plane H.

4 Experiments and Results

In this section, we report the results of two experiments we have done to validate the new approach we have outlined in the previous section. First, as mentioned earlier, it is generally very difficult to apply Equation 18 directly since the size of $\mathbb{D} \equiv \mathbb{D}^9$ is in the order of $O(e^9)$ where e the size of the set EC. Therefore, we have decided in the first experiment to drastically simplify the problem by choosing the set EC to be a small set of sample points on the sphere. Furthermore, we have reduce the domain from $\mathbb{D}^9$ to $\mathbb{D}^5$. In this simplified problem, Equation 18 is applied directly to produce a solution. The second experiment is a direct application of the iterative steps we have outlined in Section 3.5. The linear subspace R is computed for one person in the Yale Face Database and we "transplant" the resulting configuration of nine directions to form the transplanted linear subspaces for all faces in the database. We show that the transplanted linear subspaces are good approximations of their respective illumination cones by reporting good face recognition results using these subspaces.

4.1 A Toy Example

Here, we will compute a 5-dimensional linear subspace R generated by rays taken from two collections of sampled points on S^2, $\mathbb{U}$ and $\mathbb{U}'$. The coordinates frame of the experiment is defined as such that the face is facing the positive z-axis and we use the standard spherical coordinates (θ, ϕ) on S^2 to represent points on the sphere. The first collection $\mathbb{U}$ contains 35 points. We place 4, 8, 10, 8 and 4 points uniformly on the circles defined by $\phi = 45°, 65°, 90°, 105°$ and $115°$, respectively. In addition, we place a point at $\phi = 0°$, which is the direct frontal position with respect to the face. We also define the smaller collection $\mathbb{U}'$ of 21 points, by placing 1 4, 6, and 10 points uniformly on the circles defined by $\phi = 0°, 45°, 90°$, and $125°$, respectively. Note that both collections contain lighting directions from behind the face (those with $\phi > 90°$) as well as frontal directions.

Our experiment is straightforward: for each collection, we enumerate all the possible 5-dimensional linear planes generated by rays in the collection. For each ray in the collection, we render its corresponding image by using the normals and albedo values provided by the Yale Face Database B. We simplify the problem further, by assuming that R contains the frontal direction $\phi = 0°$. Therefore, there are 4847 and 46376 different planes for $\mathbb{U}$ and $\mathbb{U}'$, respectively. The final result is the plane that gives the largest $\mathrm{OB^5}$ values.

There are two reasons why we have restricted to this much simplified problem:

1. Using Matlab, it takes almost three hours to compute R using collection $\mathbb{U}$ for one individual on a 1GHz PC. Our image size is 168-by-192 and the image space has dimension 32256. Any larger collection of samples would obviously make the computation even longer.

2. In an earlier work [7], Hallinan has shown empirically that there exists a reasonably good 5-dimensional approximation of the illumination cone C and a detailed characterization of these five basis images (the eigenfaces) were given. It is interesting and informative to compare his results with ours.

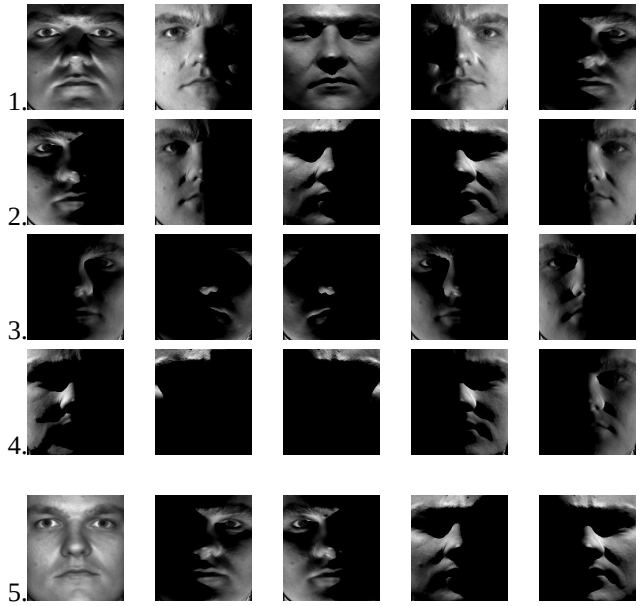

Fig. 3: The first four rows contain the 20 images from the collection $\mathbb{IU}'$ (without the image under the direct frontal lighting, i.e. $\phi = 0$). The bottom row lists the configuration obtained by our algorithm. The configuration of five lighting directions picked by our algorithm is $\{(0,0),(90,-60),(90,-120),(90,120),(90,60)\}$

For $\mathbb{IU}'$, we have computed all the ten persons in the Yale Face Database B with the normal and albedos values provided therein. Out of 4845 different configurations, our algorithm consistently picks one particular configuration for all ten persons in the Yale Database. This configuration of five directions is (in spherical coordinates (θ,ϕ)): $\{(0,0),(90,-60),(90,-120),(90,120),(90,60)\}$. This configuration is symmetric with respect to the symmetric plane of the human face. It contains frontal, sides and top/down directions. For the collection $\mathbb{IU}$, we have run our algorithm on four different individuals in the Yale Database and the result is again quite consistent. Except in one case, our algorithm picks the same configuration for all the test models. See Figure 3.

Our results are in good agreements with Hallinan's earlier result [7]. In [7], the five eigenimages are characterized as by the lightings from above/below, sides and corners. This pattern agrees with the configuration we have obtained. It has to be noted here that our results follow directly from the computational problem we defined in Section 3 while Hallinan's result is obtained empirically by acquiring large number of images and applying the principal component analysis.

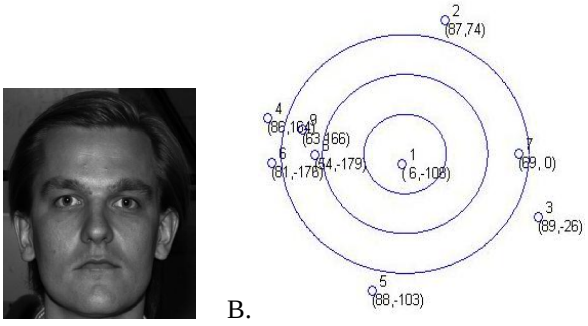

Fig. 4: A. The face used in computing the new configuration of nine lighting directions. B). The plot of the nine directions produced by our algorithm. The polar axis is the elevation angle ϕ and the azimuth angle θ goes to the usual θ in the 2D polar coordinates. The circles represent the circles with $\phi = 25^\circ, 50^\circ$ and 75°, respectively.

4.2 Another Nine Points of Light

In this section, we experiment with the iterative algorithm outlined in Section 3.5. We obtain 592 images by uniformly sampling the sphere and taking those sample points with non-negative z coordinate. We have implemented our algorithm for computing the 9-dimensional linear subspace R for one of the individual in the Yale Face Database B. For each extreme ray forming a basis vector of R, we plot the direction of the corresponding light source as shown in Figure 4.

4.3 Recognition Results

According to the plan outlined in Section 2, we transplant this configuration to the illumination cone of all the individuals in the Yale Face Database. To measure qualitatively how well the resulting linear space approximates each illumination cone, we apply these transplanted linear subspaces in a recognition experiment to see if this configuration of nine directions leads to effective face recognition compared to using either the illumination cone or eigenfaces. Using this set of nine directions, we construct a linear subspace for each of the ten persons by rendering the images of each person under these lighting conditions. These images are our training images. In practice, the nine images should be real; however, due to the lack of samples, we have opted for rendering instead. We call our method the Nine Points of Light prime(9PL') method, as a companion to our earlier Nine Points of Light method reported in [1]. The recognition results of 9PL' using this particular configuration of nine lighting directions given above together with other methods reported previously in [9] are shown in Table 1. Our results clearly indicated that the "transplanted" linear subspaces are indeed good approximations of their illumination cones.

Fig. 5: Images of one of the 10 individuals under the 4 subsets of lighting. See [3] for more examples.

COMPARISON OF RECOGNITION METHODS			
Method	Error Rate (%) vs. Illum.		
	Subset 1&2	Subset 3	Subset 4
Correlation	0.0	23.3	73.6
Eigenfaces	0.0	25.8	75.7
Eigenfaces w/o 1st 3	0.0	19.2	66.4
3D-Linear subspace	0.0	0.0	15.0
Cones-attached	0.0	0.0	8.6
9PL	0.0	0.0	2.8
9PL'	0.0	0.7	1.4
Cones-cast	0.0	0.0	0.0

Table 1: The recognition results using various different methods. Except for the two Nine Points of Light (9PL and 9PL') methods, the data for all other methods were taken from [10].

For the recognition test, real images of ten faces each under 45 different lighting conditions are used, and the test is performed on all of the 450 images. The results are grouped into 4 subsets according to the lighting angle with respect to the camera axis, Figure 5. The first two subsets cover angles $0^\circ - 25^\circ$, third $25^\circ - 50^\circ$, and the fourth $50^\circ - 77^\circ$.

5 Conclusion

Continuing our earlier work on computing the universal configuration, we have proposed a more systematic formulation by defining a new objective function. This new objective function is the quotient of two terms. One measures the distance between a given linear subspace R and the harmonic plane H. The other is simply the normalized volume of the intersection $R \cap C$. We also proposed an iterative scheme that allows us to obtain a good (but not optimal) linear approximation of the illumination cone. We demonstrate the effectiveness of the iterative scheme by producing a lighting configuration which gives good face recognition results.

Acknowledgement

We would like to thank Athos Georghiades for providing us with the Yale Face Database and his face recognition code.

References

[1] K.-C. Lee, J. Ho, and D. Kriegman, "Nine points of light: Acquiring subspaces for face recognition under variable lighting," in *Proc. IEEE Conf. on Comp. Vision and Patt. Recog.*, 2001.

[2] P. Belhumeur and D. Kriegman, "What is the set of images of an object under all possible lighting conditions," in *Int. J. Computer Vision*, vol. 28, 1998, pp. 245–260.

[3] A. Georghiades, D. Kriegman, and P. Belhumeur, "From few to many: Generative models for recognition under variable pose and illumination," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2001, pp. 643–660.

[4] R. Basri and D. Jacobs, "Lambertian reflectance and linear subspaces," in *Int. Conf. on Computer Vision*, vol. 2, 2001, pp. 383–390.

[5] R. Ramamoorthi and P. Hanrahan, "A signal-processing framework for inverse rendering," in *Proceedings of SIGGRAPH*, 2001, pp. 117–228.

[6] R. Ramamoorthi and P. Hanrahan, "An efficient representation for irradiance environment," in *Proceedings of SIGGRAPH*, 2001, pp. 497–500.

[7] P. Hallinan, "A low-dimensional representation of human faces for arbitrary lighting conditions," in *Proc. IEEE Conf. on Comp. Vision and Patt. Recog.*, 1994, pp. 995–999.

[8] W. Strauss, *Partial Differential Equations*. John Wiley & Sons, Inc, 1992.

[9] H. Chen, P. Belhumeur, and D. Jacobs, "In search of illumination invariants," in *Proc. IEEE Conf. on Comp. Vision and Patt. Recog.*, 2000, pp. 1–8.

[10] A. Georghiades, D. Kriegman, and P. Belhumeur, "Illumination cones for recognition under variable lighting: Faces," in *Proc. IEEE Conf. on Comp. Vision and Patt. Recog.*, 1998.