To Appear in CVPR 2001

Nine Points of Light: Acquiring Subspaces for Face Recognition under Variable Lighting *

Kuang-Chih Lee † Jeffrey Ho ‡ David Kriegman §

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

Abstract

Previous work has demonstrated that the image variations of many objects (human faces in particular) under variable lighting can be effectively modeled by low dimensional linear spaces. Basis images spanning this space are usually obtained in one of two ways: A large number of images of the object under different conditions is acquired, and principal component analysis (PCA) is used to estimate a subspace. Alternatively, a 3-D model (perhaps reconstructed from images) is used to render virtual images under either point sources from which a subspace is derived using PCA or more recently under diffuse synthetic lighting based on spherical harmonics. In this paper, we show that there exists a configuration of nine point light source directions such that by taking nine images of each individual under these single sources, the resulting subspace is effective at recognition under a wide range of lighting conditions. Since the subspace is generated directly from real images, potentially complex intermediate steps such as PCA and 3D reconstruction can be completely avoided; nor is it necessary to acquire large numbers of training images or physically construct complex diffuse (harmonic) light fields. We provide both theoretical and empirical results to explain why these linear spaces should be good for recognition.

1. Introduction

To build a robust and efficient face recognition system, the problem of lighting variation is one of the main technical challenges facing system designers. In the past few years, many appearance-based methods have been proposed to handle this problem, and new theoretical insights as well as good recognition results have been reported in various publications, e.g. [1, 2, 3, 4, 5, 6]. The main insight gained from these results is that there are both empirical and analytical justifications for using low dimensional linear subspaces to model image variations of human faces under different lighting conditions. Early work showed that the variability of images of a Lambertian surface in fixed pose, but under variable lighting where no surface point is shadowed, is a three-dimensional linear subspace [1, 7, 8]. What has been perhaps more surprising is that even with cast and attached shadows, the set of images is still well approximated by a relatively low dimensional subspace, albeit with a bit higher dimension [2].

Under the Lambertian assumption, the set of images of an object under all possible lighting conditions forms a polyhedral cone, the illumination cone, in the image space [3]. In a follow-up paper [9], it was reported that the illumination cones of human faces can be approximated well by low-dimensional linear subspaces. More recently, 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 [6]. The magic number of nine comes from the number of spherical harmonics with degree less than three. 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 linear subspace, a straightforward recognition scheme can be developed and results obtained in [6] are excellent.

The starting point of our paper is to understand this nine-dimensional space proposed by Basri and Jacobs in the context of the illumination cone and face recognition and to explore other ways to construct this linear subspace. The

*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

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

importance of this 9-dimensional linear subspace for face recognition is confirmed by the good recognition results reported in [6]. In their approach, the 9-dimensional linear subspace H, the harmonic plane, is formed by simulating nine harmonic images: the images of the model under the lighting condition specified by spherical harmonics. These nine harmonic images form the basis of H. In order to simulate harmonic images, the model's 3D structure, or at least its normals and albedos, has to be known in advance. On the other hand, simple linear algebra tells us that any set of nine linearly independent vectors (or images) in H is sufficient to recover the plane. This hints at the possibility of an easier way to obtain the linear subspace: that is, can we find a set of nine real images such that the linear subspace spanned by them coincides with the harmonic plane? For all practical purpose, the answer to this question is 'no'. Since any real image in H requires a smooth lighting condition specified by a linear combination of the first nine spherical harmonics, it is very difficult to reproduce the exact lighting condition in a common laboratory environment. However, one can ask a different but related question: is there another 9-dimensional linear subspace R which can also provide a good representation for face recognition? Can R be constructed in some canonical fashion, perhaps with nine physically and easily realized lighting conditions?

Since we know that H is good for face recognition, it is probably a good idea to find a plane R close to H. From the recognition standpoint, it is also preferable to require that the intersection between R and the illumination cone C is as large as possible. That is, we want to find a 9-D linear subspace R generated by elements in the illumination cone C such that the distance between R and H is minimized (in some way) while the (unit)-volume $R \cap C$ is maximized. In Section 3, we formulate the problem in terms of maximizing an objective function defined on the set of extreme rays on the cone. Our end result is a set of nine extreme rays that span R and the nine source directions corresponding to these nine extreme rays. It turns out that the resulting nine light source directions are qualitatively very similar for different individuals. By averaging the objective functions for different individuals and maximizing this new objective function, we obtain a configuration of nine light source directions, the universal configuration, such that on average, the linear space spanned by the corresponding extreme rays is a good approximation to the illumination cone. We demonstrate that by using this universal configuration of nine directional sources, good face recognition results can be obtained using the linear subspace spanned by the resulting nine images.

From a practical standpoint, acquiring image under a single distant and isotropic light source is much easier and less costly than alternatives. That is, the linear subspace R is lot easier to obtain than the harmonic plane H or an illumination cone. This is particularly applicable for acquiring training images of individuals in a controlled environment such as a driver's license office, a bank, or a security office.

This paper is organized as follows. In the next section, we briefly summarize the idea of [6] using a harmonic plane H for face recognition. The relationship between H and the illumination cone [3] is explained. Our algorithms for computing R and the universal configuration are described in Section 3 and 4, respectively. The final section contains a brief summary and conclusion of this paper.

2 Harmonic Images and the Illumination Cone

2.1 Lambertian Reflection and Spherical Harmonics

In this section, we briefly summarize the recent work presented in [6, 10, 11]. 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

$$\begin{aligned} r(p) &= \lambda \int\int_S k(\theta) L(\theta, \phi) dA \\ &= \lambda \int_0^{2\pi} \int_0^{\pi} k(\theta) L(\theta, \phi) sin\theta d\theta d\phi \end{aligned} \tag{1}$$

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 $-l \leq m \leq l$:

$$Y_{lm}(\theta, \phi) = N_{lm} P_l^{|m|}(cos(\theta)) e^{im\phi} \tag{2}$$

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 associ-

ated Legendre functions (its precise definition is not important here; however, see [12]). 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 $g \in SO(3)$:

$$Y_{lm}(g(\theta,\phi)) = \sum_{n=-1}^{l} g_{mn}^{l} Y_{ln}(\theta,\phi). \tag{3}$$

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 [6, 10]. 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 1, 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 1 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 3), the same truncated L' will work at every surface point.

2.2 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 [6]: "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

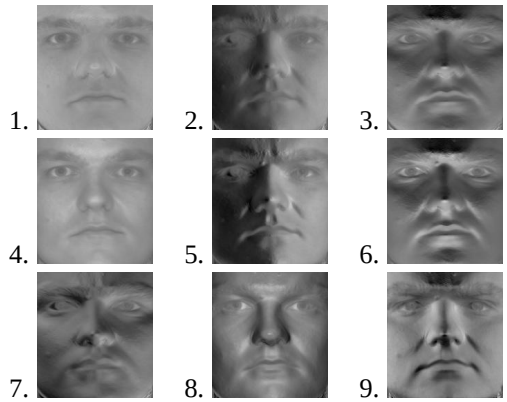

Fig. 1: The nine simulated harmonic images of a face from Yale Database. The light gray and dark gray indicate the positive and negative pixel values. Since the Y_{00} is a constant, the corresponding harmonic image simply scales the albedo values as shown in Picture 1. Pictures 4 is the harmonic image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

$$Y_{00} = 0.2821; \tag{4}$$

$$(Y_{11}; Y_{10}; Y_{1-1}) = 0.4886(x; y; z); \tag{5}$$

$$(Y_{21}; Y_{2-1}; Y_{2-2}) = 1.093(xz; yz; xy); \tag{6}$$

$$Y_{20} = 0.3154(3z^2 - 1); \tag{7}$$

$$Y_{22} = 0.5462(x^2 - y^2); \tag{8}$$

Figure 1 shows the rendered harmonic images for a face taken from the Yale Database. These synthetic images are rendered by sampling 1000 rays on a hemisphere, and the final images are the weighted sum of 1000 ray-traced images. Unlike [6] 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 4–8.

2.3 Relation to the Illumination cone

From the discussion above, we can conclude a few things about the relationship between the linear subspace H generated by the harmonic images and the illumination cone C [3]. We let P denote the interior of the positive orthant: $P = \{(x_1, \ldots, x_n) | x_i > 0 \text{ for } 1 \leq i \leq n\}$, where n is the dimension of the image space, i.e. the number of pixels.

First, H approximates well the images obtained under smooth diffuse lighting. Images resulting from this type of lighting typically lie in P. That is, every pixel is illuminated. In fact, The polynomial $f(x,y,z) = \sqrt{3} - (x+y+z)$

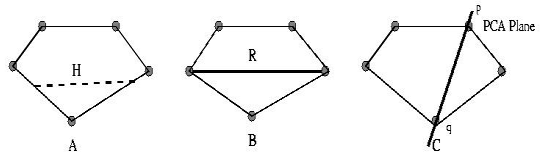

Fig. 2: Cross section of the illumination cone C. The solid circles denote the extreme rays of the cone. A) The intersection $C \cap H$ is shown as the dashed line. Notice that the intersection does not contain extreme rays and H is parallel to the direction in which C is the thickest. B). A possible 9-dimensional space which is good for face recognition. C) A PCA plane obtained by choosing extreme rays p and q as samples.

is a non-negative function on the unit sphere. The image of a human face resulted from a lighting condition specified by f has no pixel with zero intensity value. The proof is simple since the the zero set of f are lines on the unit sphere. That is, the intersection of H with the interior of the cone $C \cap P$ is non-empty : the necessary condition for H to be a good approximation of C. Second, the good recognition results reported by [6] suggest the following two possibilities.

1. The volume of the intersection $C \cap H$ is large.

2. The intersection $C \cap P$ is 'concentrated' near $C \cap H$.

That is the illumination cone C is thick in the directions parallel to H while it is thin in directions perpendicular to H. If none of the above conditions is true, it will be difficult to explain the good recognition results reported in [6]. A general picture emerged from these observations is depicted in Figure2(a). It is then natural to ask the questions: Is there another 9-dimensional linear subspace R which is also good for face recognition? Can R be constructed so that it is, in some way, intrinsic to the illumination cone? Expressed differently, is there a canonical procedure to determine R directly from the illumination cone? And considering the complexity of the illumination cone, how do we compute R, if it exists? This is depicted in Figure 2(b). In the following sections, we provide our (partial) solutions to these questions.

Of course, there are many ways to arrive at a 9-dimensional linear subspace. The most common and straightforward way is to sample images in the cone and use the principal component analysis. However, principal component analysis depends heavily on the sample images used to define the correlation matrix, whose eigenvectors define the resulting PCA plane. A biased set of samples (e.g. small number of samples) would produce a PCA plane that is not effective for face recognition, as shown in Figure 2.

3 Low Dimensional Linear Approximation of Illumination Cone

In this section, we detail our algorithm for computing R. Recall that our overall aim is to find a 9-dimensional linear subspace R which can provide a basis for a good face recognition method; in addition, we would like to have some canonical procedure that can determine R directly from the illumination cone. 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 basis for a good face recognition method, we require R to satisfy the following two conditions:

1. The angular distance between R and H should be minimized.

2. The (unit) volume $C \cap R$ should be maximized. (The unit volume is defined as the volume of the intersection of $C \cap R$ with the unit ball).

Note that $C \cap R$ is a subcone of C. Therefore, the second condition is equivalent to maximizing the angle subtended by the subcone $C \cap R$. Since we know that H is good for face recognition. It is reasonable to assume that any subspace close to H would likewise be good for recognition; hence the first condition.

In [3], it was shown that the number of extreme rays is $m(m-1)$ where m is the number of distinct surface normals - m is typically greater than 1000. Therefore, in most cases, the full illumination cone C is too difficult to compute. This implies that a linear subspace R satisfying the two conditions above is also likely to be difficult to compute. Instead, 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

$$x_i = \arg \max_{x \in EC_{i-1}} \frac{\mathrm{dist(x, R_{i-1})}}{\mathrm{dist(x, H)}}. \quad (9)$$

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 since H is the harmonic plane, $\mathrm{dist(x, H)}$ is always non-zero for all $x \in EC_i$. When computing R_1, we define $\mathrm{dist(x, R_0)} = \mathrm{dist(x, \emptyset)}$ to be 1. Therefore, the first element x_1 is the extreme ray in C that is closest to the harmonic plane H.

3.2 Discussion

From the recognition standpoint, the optimal linear plane R should be the plane that maximizes the unit volume of the intersection $R \cap C$. To formulate this problem more precisely, let $GR(n, 9)$ denote the space of 9-dimensional linear subspaces of $\mathbb{R}^n$, the Grassmannian. The unit volume of the intersection $S \cap C$ of each $S \in GR(n, 9)$ with the illumination cone C defines a continuous function vol on $GR(n, 9)$ and the optimal linear plane R is simply a global maximum of vol. Since the dimension of $GR(n, 9)$ is $9(n - 9)$ and n is the number of pixels, direct computations on $GR(n, 9)$ is out of question. Instead, we restrict our domain to a subset $\mathbb{D}$ of $GR(n, 9)$ consisting of 9-dimensional linear subspaces generated by the extreme rays of C. That is, each linear space in $\mathbb{D}$ has a basis consisting of only extreme rays of the cone. The maximization problem is now equivalent to maximizing the "solid angles" subtended by the extreme ray basis. The space $\mathbb{D}$ is discrete and contains at most $C(e, 9)$ points (e is the number of extreme rays). If e is small, we can enumerate every point in $\mathbb{D}$ and compute the maximal intersection. This, again, is not possible in general.

A straightforward approach, using a greedy algorithm, is similar to the iterative steps we outlined above. However, the greedy algorithm can not guarantee that a global maximum is reached at the end; nevertheless, one expects the resulting linear space R should have "large" intersection $R \cap C$. Starting with any extreme ray, we can iteratively find a sequence of extreme rays such that the successive linear spaces spanned by these rays intersect the cone in some 'maximal way'. We formulate this in terms of maximizing the distance between a normalized extreme ray x and the current linear space R_i as in the numerator of Equation 9. Of course, the resulting linear subspace generated by this process depends on the initial extreme ray. Figure 3 illustrates two possibilities. If the number of extreme rays is small, we can perform the same iterative process with every extreme ray as the initial ray. Unfortunately, this is computationally impossible in our case because of the large number of extreme rays. Therefore, it is important to have a good initial ray to start with. In our case, our starting ray x_1 is the extreme ray that is closest to the harmonic plane H. According to our observation, H should be parallel to the directions in which the cone is thickest and more than likely, there is an almost optimal R which is close to it. This almost proves the fact that an almost optimal plane is very likely to contain our initial ray x_1. That is, our iterative process will start at the right initial value.

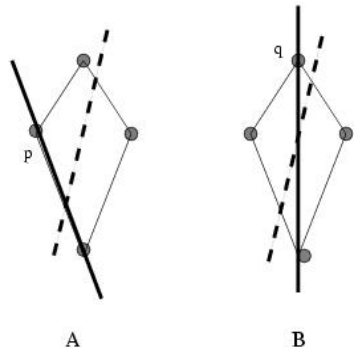

Fig. 3: Both figures depict a cross section of the illumination cone with the dashed line indicating the harmonic plane. A) If p is chosen as the initial ray, the resulting linear subspace intersects the cone only on the boundary. B) If q is chosen instead, the resulting linear space is the optimal one.

In a sense, condition 1 is not necessary if the number of extreme rays is small. However, if this is not the case, for computational reason, it is necessary to have a good starting ray. Condition 1 or more precisely, the harmonic plane H provides us with a good guess.

3.3 Experiments and results

In our implementation, EC, the full set of extreme rays of C, is replaced by a subset of 200 extreme rays. Following [5] we obtain these 200 rays by uniformly sampling the hemisphere. For each sampled direction, we produce the corresponding extreme ray by rendering an image under a single directional source emanating from this direction (with intensity set to 1). The azimuth ϕ and elevation θ angles are defined on the hemisphere such that for these 200 images, θ varies between $-180°$ and $180°$ and ϕ goes from $0°$ to $87°$. This set of 200 sampled extreme rays is used to define the domain for the maximization procedure specified by Equation 9. We have implemented our algorithm for computing the linear subspace R using the Yale Face Database B. For ten individuals, the Yale database contains a 3D model and 45 images under different lighting conditions of each person.

Some results of computing the 9-dimensional linear subspace R for each person in the database are shown in Figure 4. For each extreme ray forming a basis vector of R, we plot

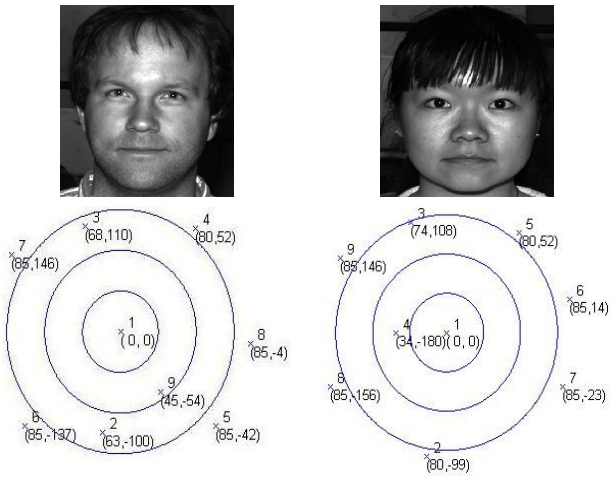

Fig. 4: The first row displays two of the ten uncropped faces in the Yale database. The second row is the corresponding plots of the nine directions produced by our algorithm. The polar axis is the elevation angle ϕ and the azimuth angle θ goes to the usual θ in the 2D polar coordinates. The circles represent the circles with $\phi = 25^\circ, 50^\circ$ and 75°, respectively.

the direction of the corresponding light source. It is worthwhile to note that the set of nine extreme rays chosen by the algorithm has a particular type of configuration. First, the frontal direction (with $\theta = \phi = 0^\circ$) is always present. In fact, it is always the first basis vector x_1, the one closest to the harmonic plane H. Second, besides the frontal image, there are another 2 to 3 "interior" images, i.e. those produced by the lighting directions with $|\theta|, \phi \leq 65^\circ$ Third, the other directions are concentrated on the sides (rather than above or below) and with directions $|\theta|, \phi \geq 65^\circ$. It is well known that these directions produce large shadows on human faces, and makes face recognition more difficult [9]. Our results seem to tell us the obvious: more samples are needed on the part of the hemisphere that is most likely to produce difficult images to recognize. It is important to note that it is by no mean clear a priori that our algorithm based on the two conditions explained in Section 3.1 will favor such type of configurations.

4 Nine Points of Light

The results in the previous section demonstrate that, for each individual, there exists a configuration of nine lighting directions such that the linear subspace spanned by these images is a good linear approximation of the illumination cone. The configurations are qualitatively similar for different individuals with small variations in each lighting direction. It is then logical to seek a fixed configuration of nine lighting directions for all individuals such that for each

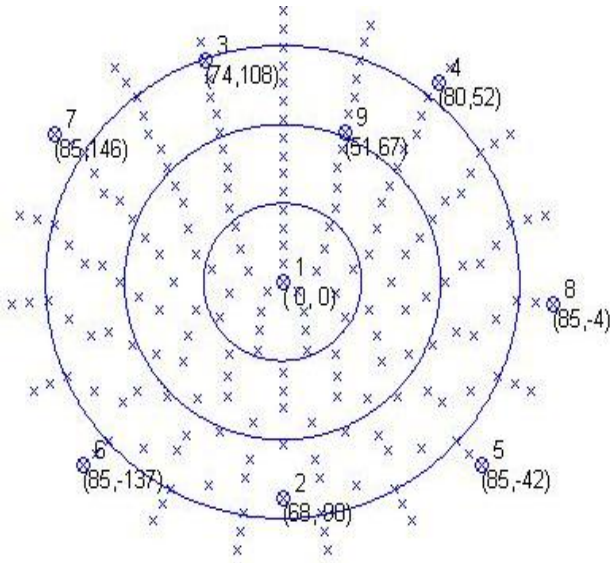

Fig. 5: The plot (the projection from the hemisphere (ϕ, θ) onto the xy-plane in polar coordinates (r, θ) is: $\phi \to r, \theta \to \theta$) of the universal configuration of nine light source directions with all 200 sample points. The circles represent the circles on the hemisphere with ϕ = 25°, 50° and 75°, respectively.

individual, on average, the linear space spanned by the corresponding extreme rays is a good linear approximation to the illumination cone.

To find such a configuration (or an approximate of it), we can modify our previous method slightly by computing the average of the quotient in Equation 9 over all the available training models. With all the notations defined as above, we find the nested linear subspaces $R_0 \subseteq R_1 \subseteq \ldots \subseteq R_i \cdots \subseteq R_9 = R$ by computing each x_i such that

$$x_i = \arg \max_{x \in EC_{i-1}} \sum_{k=1}^{l} \frac{\mathrm{dist}(\mathrm{x^k}, \mathrm{R^k_{i-1}})}{\mathrm{dist}(\mathrm{x^k}, \mathrm{H^k})}. \qquad (10)$$

Since we are computing Equation 9 for all the available face models simultaneously, the set EC denotes the set of 200 sample points on the hemisphere and for each $x \in EC$, $\mathrm{x^k}$ denotes the image of model k taken under a single light source with direction x. EC_i denotes the set obtained by deleting i elements from EC. k indexes the available face models. H^k denotes the harmonic plane of model k and R^k_{i-1} represents the linear subspace spanned by the images $\{x^k_1, \cdots, x^k_i\}$ of model k under light source directions $\{x_1, \cdots, x_i\}$.

We call the resulting configuration of nine directions the universal configuration. These directions are $\{$ $(0, 0)$, $(68, 90)$, $(74, 108)$, $(80, 52)$, $(85, -42)$, $(85, -137)$, $(85, 146)$, $(85, -4)$, $(51, 67)$ $\}$. They along with the 200 samples on the hemisphere are plotted in Figure 5.

Fig. 6: Images of one of the 10 individuals under the 4 subsets of lighting. See [5] for more examples.

4.1 Recognition Results

Next we apply the previous result in a recognition experiment to see if the 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. 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 (9PL) method. The recognition results of 9PL using this particular configuration of nine lighting directions given above together with other methods reported previously in [13] are shown in Table 4.1.

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. The first two subsets cover angles $0^\circ - 25^\circ$, third $25^\circ - 50^\circ$, and the fourth $50^\circ - 77^\circ$.

All of the other methods reported in the table require considerable amount of off-line processing on training data. For the Nine Points of Light method, there is no training involved !! The work is almost minimal: only nine images are needed. It is also interesting to observe that our method performs much better than the eigenfaces method. It should be pointed out that all the methods listed in Table 4.1 that require off-line processing were trained using all the images of Subset 1 and 2, in particular, the eigenfaces and linear subspace methods. Our nine rendered images are mostly from Subset 4 (seven images per person). Because of this, it is expected that our method should perform better than most of the other methods for Subset 4 which is the most difficult subset with great amount of shadow variations. This is indeed the case. However, for Subsets 1-3, our method still performs equally well compared with all other methods.

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
Linear subspace	0.0	0.0	15.0
Cones-attached	0.0	0.0	8.6
9PL	0.0	0.0	2.8
Cones-cast	0.0	0.0	0.0

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

4.2 Further Dimensional Reduction

We have demonstrated that there is a configuration of nine light source directions which provide a good representation for face recognition. As shown in [3], the actual dimension of an illumination cone is the number of distinct surface normals. Hence, for human faces, the actual dimension of the illumination cone is quite large; nevertheless, the previous results show that the illumination cone for a human face (under a fixed pose) admits a good approximation by a 9-dimensional linear plane in the image space. The natural extension of this conclusion is to further reduce the dimension of the linear approximation and observe the resulting recognition rate.

We experimented with this type of dimensional reduction by successively using each linear subspace in the nested sequence, $R_0 \subseteq R_1 \subseteq \ldots \subseteq R_i \cdots \subseteq R_9 = R$, for face recognition. The results are shown in Figure 7. It is clear that the recognition result is still reasonably good even when the dimension has been reduced down to only five. However, the error rate becomes noticeable when the dimension is further reduced. These results corroborate well with the much earlier results of [1, 2]. They have shown that using 5 ± 2 eigenimages is sufficient to provide a good representation of the images of a human face under variable lighting. The main distinctions between these earlier results and ours are 1) the linear approximations provided by the earlier work have always been characterized in terms of eigenimages. In contrast, our linear approximations are characterized by real images. 2) There is no report of recognition results in these earlier work while we have demonstrated that not only a good low-dimensional linear approximation of the illumination cone is possible but it also provides reasonably good face recognition results.

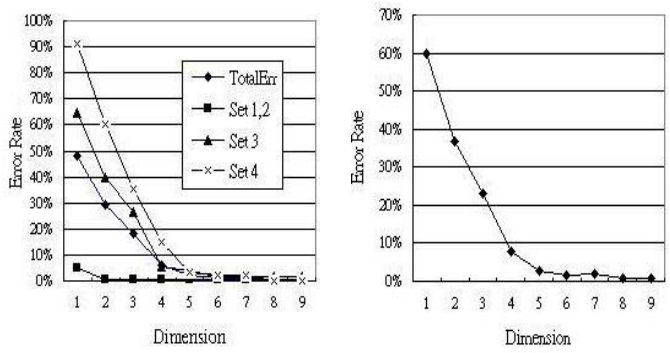

Fig. 7: **Left**: The Error rates for face recognition using successively smaller linear subspaces. The abscissa represents the dimension of the linear subspace while the ordinate gives the error rate. In this experiment, the extended Yale Face Database, containing 1710 images of 38 individuals, was used. **Right**: The total error rate on all 1710 images.

5 Conclusion and Future Work

We have shown that there exists a set of nine single light source directions that is important for face recognition. The linear subspace spanned by the corresponding extreme rays is a good approximation to the illumination, and it provides good face recognition results under variable lighting. We obtain the set by maximizing a function defined on the set of the extreme rays of the illumination cone. Our result provides a recipe for building a simple but robust face recognition system. By taking nine images of each individual with single light sources emanating from these nine directions, our results show that these nine images are already sufficient for the task of recognizing faces under different illumination conditions. The usual complicated intermediate steps, such as the 3D reconstruction of the model, can be completely avoided.

One surprising conclusion of our work is that for modeling the *effect of illumination* on human faces, linear superposition of a few directional sources may very well be as effective as linear superposition of smooth diffuse light sources (the harmonic plane). This is surprising because the directional sources (represented as delta functions) and smooth diffuse sources (represented as smooth functions) are, in some ways, completely opposite of each other. In any function space with any reasonable norm, the delta functions and smooth functions are certainly different. We believe that this seemingly paradoxical conclusion can be attributed to the prominent geometric feature of human faces: human faces are generally flat and the variation in normals are generally small over a large portion of the face. Verifying this claim will be one of the main themes in our future research.

Acknowledgement

We would like to thank David Jacobs for the discussion on harmonic lighting [6]. Thanks also go to Athos Georghiades for providing us with the Yale Face Database and his face recognition code.

References

[1] 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.

[2] R. Epstein, P. Hallinan, and A. Yuille, "5+/-2 eigenimages suffice: An empirical investigation of low-dimensional lighting models," in *PBMCV*, 1995.

[3] 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(3), 1998, pp. 245–260.

[4] Y. Adini, Y. Moses, and S. Ullman, "Face recognition: The problem of compensating for changes in illumination direction," in *IEEE Trans. PAMI*, vol. 19, July 1997, pp. 721–732.

[5] 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.

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

[7] S. Nayar and H. Murase, "Dimensionality of illumination in appearance matching," *IEEE Conf. on Robotics and Automation*, 1996.

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

[9] 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.

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

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

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

[13] 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.