Illumination Cones for Recognition
Under Variable Lighting: Faces

Athinodoros S. Georghiades   David J. Kriegman   Peter N. Belhumeur

Center for Computational Vision and Control
Department of Electrical Engineering
Yale University
New Haven, CT 06520-8267

Abstract

Due to illumination variability, the same object can appear dramatically different even when viewed in fixed pose. To handle this variability, an object recognition system must employ a representation that is either invariant to, or models this variability. This paper presents an appearance-based method for modeling the variability due to illumination in the images of objects. The method differs from past appearance-based methods, however, in that a small set of training images is used to generate a representation – the illumination cone – which models the complete set of images *of an object with Lambertian reflectance under an arbitrary combination of point light sources at infinity. This method is both an implementation and extension (an extension in that it models cast shadows) of the illumination cone representation proposed in [3]. The method is tested on a database of 660 images of 10 faces, and the results exceed those of popular existing methods.*

1   Introduction

An object's appearance depends in large part on the way in which it is viewed. Often slight changes in pose and illumination produce large changes in an object's appearance. While there has been a great deal of literature in computer vision detailing methods for handling image variation produced by changes in pose, few efforts have been devoted to image variation produced by changes in illumination. For the most part, object recognition algorithms have either ignored illumination variation, or dealt with it by measuring some property or feature of the image – e.g., edges or corners – which is, if not invariant, at least insensitive to the variability. Yet, edges and corners do not contain all of the information useful for recognition. Furthermore, objects which are not simple polyhedra or are not composed of piecewise constant albedo patterns often produce inconsistent edge and corner maps.

Methods have recently been introduced which use low-dimensional representations of images of objects to perform recognition, see for example [5, 11, 16]. These methods, often termed appearance-based methods, differ from the feature-based methods mentioned above in that their low-dimensional representation is, in a least-squared sense, faithful to the original image. Systems such as SLAM [11] and Eigenfaces [16] have demonstrated the power of appearance-based meth-

ods both in ease of implementation and in accuracy. Yet these methods suffer from an important drawback: recognition of an object (or face) under a particular pose and lighting can be performed reliably *provided that object has been previously seen under similar circumstances.* In other words, these methods in their original form have no way of extrapolating to novel viewing conditions.

The "illumination cone" method of [3] is, in spirit, an appearance-based method for recognizing objects under extreme variability in illumination. However, the method differs substantially from previous methods in that a small number of images of each object under small changes in lighting is used to generate a representation, the illumination cone, of all images of the object (in fixed pose) under all variation in illumination. This paper focuses on issues for building the illumination cone representation from training images and using it for recognition.

While the structure of the set of images under variable illumination was characterized in [3] and the relevant results are summarized in Sec. 2, no methods for performing recognition were presented. In this paper, such recognition algorithms are introduced. Furthermore, the cone representation is extended to explicitly model cast shadows produced by objects which have non-convex shapes. This extension is non-trivial, requiring that the surface normals for the objects be recovered up to a *shadow preserving* generalized bas-relief (GBR) transformation.

The effectiveness of these algorithms and the cone representation are validated within the context of face recognition – it has been observed by Moses, Adini and Ullman that the variability in an image due to illumination is often greater than that due to a change in the person's identity [10]. Figure 1 shows the variability for a single individual. It has been observed that methods for face recognition based on finding local image features and using their geometric relation are generally ineffective [4]. Hence, faces provide an interesting and useful class of objects for testing the power of the illumination cone representation.

In this paper, we empirically compare these new methods to a number of popular techniques such as correlation [4] and Eigenfaces [9, 16] as well as more recently developed techniques such as distance to linear subspace [2, 5, 12, 13]; the latter technique has been shown to be much less sensitive to illumination

variation than the former. However, these methods also break down as shadowing becomes very significant. As we will see, the presented algorithm based on the illumination cone outperforms all of these methods on a database of 660 images. It should be noted that our objective in this work is to focus solely on the issue of illumination variation whereas other approaches have been more concerned with issues related to large image databases, face finding, pose, and facial expressions.

2   The Illumination Cone

In earlier work, it was shown that for an object with convex shape and Lambertian reflectance, the set of all images under an arbitrary combination of point light sources forms a convex polyhedral cone in the image space $\mathbb{R}^n$. This cone can be constructed from as few as three images [3]. Here we summarize the relevant results.

To begin, consider a convex object with a Lambertian reflectance function which is illuminated by a single point source at infinity. Let $\mathbf{x} \in \mathbb{R}^n$ denote an image of this object with n pixels. Let $B \in \mathbb{R}^{n \times 3}$ be a matrix where each row of B is the product of the albedo with the inward pointing unit normal for a point on the surface projecting to a particular pixel in the image. A point light source at infinity can be represented by $\mathbf{s} \in \mathbb{R}^3$ signifying the product of the light source intensity with a unit vector in the direction of the light source. A convex Lambertian surface with normals and albedo given by B, illuminated by $\mathbf{s}$, produces an image $\mathbf{x}$ given by

$$\mathbf{x} = \max(B\mathbf{s}, \mathbf{0}), \tag{1}$$

where $\max(., \mathbf{0})$ sets to zero all negative components of the vector $B\mathbf{s}$. The pixels set to zero correspond to the surface points lying in an *attached shadow*. Convexity of the object's shape is assumed at this point to avoid *cast shadows* (shadows that the object casts on itself). While attached shadows are defined by local geometric condition, cast shadows must satisfy a global condition.

When no part of the surface is shadowed, $\mathbf{x}$ lies in the 3-D subspace $\mathcal{L}$ given by the span of the matrix B. It can be shown that the subset $\mathcal{L}_0 \subset \mathcal{L}$ having no shadows (i.e., falling in the non-negative orthant[1]) forms a convex cone [3].

The illumination subspace $\mathcal{L}$ slices through other orthants as well as the non-negative orthant. Let $\mathcal{L}_i$ be the intersection of the illumination subspace $\mathcal{L}$ with an orthant i in $\mathbb{R}^n$ through which $\mathcal{L}$ passes. Certain components of $\mathbf{x} \in \mathcal{L}_i$ are always negative and others always greater than or equal to zero. Since im-

age intensity is always non-negative, the image corresponding to points in $\mathcal{L}_i$ is formed by the projection P_i given by Equation 1. The projection P_i is such that it leaves the non-negative components of $\mathbf{x} \in \mathcal{L}_i$ untouched, while the negative components of $\mathbf{x}$ become zero. The projected set $P_i(\mathcal{L}_i)$ is also a convex cone. $\mathcal{L}$ intersects at most $n(n-1) + 2$ orthants [3], and so the set of images created by varying the direction and strength of a *single* light source at infinity is given by the union of at most $n(n-1) + 2$ convex cones, each of which is at most three dimensional.

If an 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.,

$$\mathbf{x} = \sum_{i=1}^{k} \max(B\mathbf{s}_i, \mathbf{0}) \tag{2}$$

where $\mathbf{s}_i$ is a single light source. It follows that the set of all possible images $\mathcal{C}$ 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, it is shown in [3] that any image in the cone $\mathcal{C}$ (including the boundary) can be found as a convex combination of *extreme rays* given by

$$\mathbf{x}_{ij} = \max(B\mathbf{s}_{ij}, \mathbf{0}), \tag{3}$$

where

$$\mathbf{s}_{ij} = \mathbf{b}_i \times \mathbf{b}_j. \tag{4}$$

The vectors $\mathbf{b}_i$ and $\mathbf{b}_j$ are the rows of B with $i \neq j$. It is clear that there are at most $m(m-1)$ extreme rays (images) for $m \leq n$ independent surface normals. Since there is a finite number of extreme rays, the convex cone is polyhedral.

3   Constructing the Illumination Cone

Equations 3 and 4 suggest a way to construct the illumination cone for each individual: gather three or more images of the face under varying illumination without shadowing and use these images to estimate the three-dimensional illumination subspace $\mathcal{L}$. One way of estimating this is to normalize the images to be of unit length, and then use singular value decomposition (SVD) to estimate the best three-dimensional orthogonal basis B^* in a least square sense. Note that the basis B^* differs from B by an unknown linear transformation, i.e., $B = B^*A$ where $A \in GL(3)$; for any light source, $\mathbf{x} = B\mathbf{s} = (B^*A)(A^{-1}s)$. Nonetheless from B^*, the extreme rays defining the illumination cone $\mathcal{C}$ can be computed using Equations 3 and 4. This method, introduced in [3], was named the *illumination subspace method*.

The first problem that arises with the above procedure is with the estimation of B^*. For even a convex object whose Gaussian image covers the Gauss

Subset 1	Subset 2	Subset 3	Subset 4	Subset 5
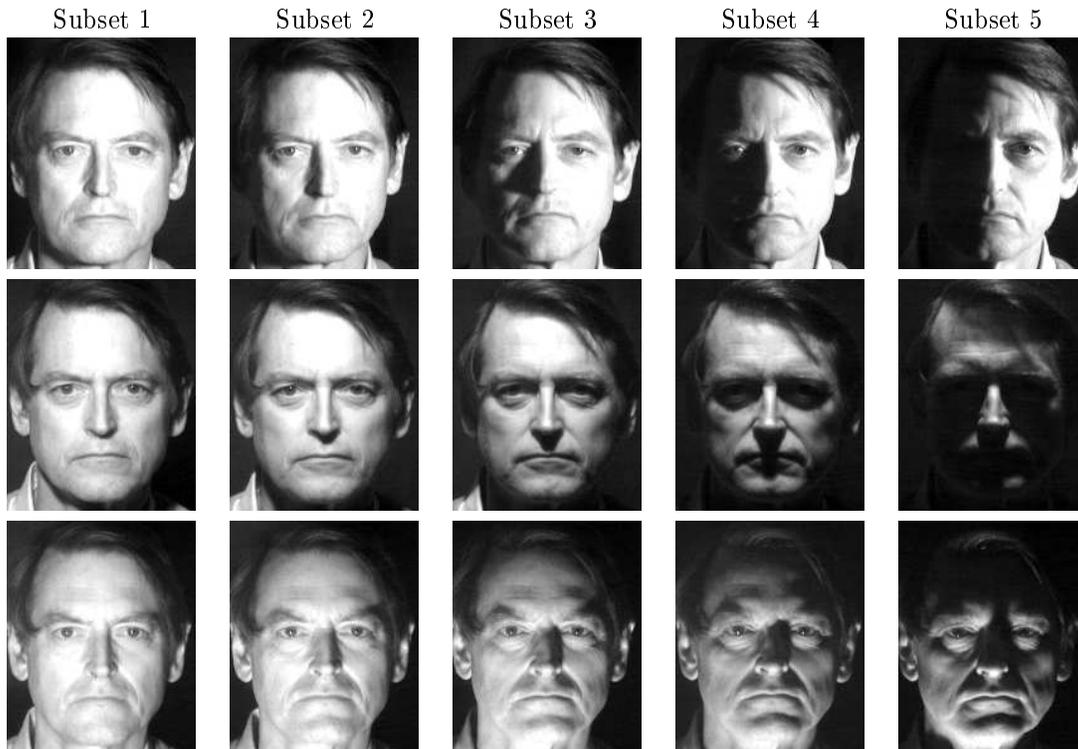
Figure 1: Example images from each subset of the Harvard Database used to test the algorithms.

sphere, there is only one light source direction (the viewing direction) for which no point on the surface is in shadow. For any other light source direction, shadows will be present. For faces, which are not convex, shadowing in the modeling images is likely to be more pronounced. When SVD is used to estimate B^* from images with shadows, these systematic errors can bias the estimate of B^* significantly. Therefore, alternative ways are needed to estimate B^* that take into account the fact that some data values should not be used in the estimation.

The next problem is that usually m, the number of independent normals in B, can be large (more than a thousand) hence the number of extreme rays needed to completely define the illumination cone can run in the millions. Therefore, we must approximate the cone in some fashion; in this work, we choose to use a small number of extreme rays (images). In [3] it was shown empirically that the cone is flat (i.e., elements lie near a low dimensional linear subspace), and so the hope is that a subsampled cone will provide an approximation that leads to good recognition performance. In our experience, around 60-80 images are sufficient, provided that the corresponding light source directions $\mathbf{s}_{ij}$ are more or less uniform on the illumination sphere. The resulting cone $\mathcal{C}^*$ is a subset of the object's true cone $\mathcal{C}$. An alternative approximation to $\mathcal{C}$ can be obtained by directly sampling the space of light source directions rather than generating the samples through

Eq. 4. While the resulting images form the extreme rays of the representation $\mathcal{C}^*$ and lie on the boundary of $\mathcal{C}$, they are not necessarily extreme rays of $\mathcal{C}$. Again $\mathcal{C}^*$ is a subset of $\mathcal{C}$.

The last problem comes from the fact that faces are non-convex, and so cast shadows cover significant portions of the face under extreme illumination (See the images from Subsets 4 and 5 in Fig. 1). The image formation model (Eq. 1 used to develop the illumination cone does not account for cast shadows. For the light source directions of the extreme rays given by Equation 4, we can predict which pixels will be in cast shadows.

It has been shown [1, 17] that from multiple images where the light source directions are unknown, one can only recover a Lambertian surface up to a three-parameter family given by the generalized bas-relief (GBR) transformation. This family scales the relief (flattens or extrudes) and introduces an additive plane. Consequently, when computing $\mathbf{s}_{ij}^*$ from B^*, the light source direction differs from the true light source by a GBR transformation. Since shadows are preserved under these transformation [1], images synthesized from a surface whose normal field is given by B^* under light source $\mathbf{s}_{ij}^*$ will have correct shadowing. Thus, in constructing the extreme rays of the cone, we first reconstruct a surface (a height function) and then use ray-tracing techniques to determine which points lie in a cast shadow. It should be noted that the vec-

tor field B^* estimated via SVD may not be integrable, and so prior to reconstructing the surface up to GBR, integrability of B^* is enforced.

This leads to the following steps for constructing a representation of the illumination cone of an individual from a set of images taken under unknown lighting. Details of these steps are given below.

1. Estimate B^* from training images.
2. Enforce integrability of B^*.
3. Reconstruct the surface up to GBR.
4. For a set of light source directions that uniformly sample the sphere, synthesize extreme rays (images) of the cone that account for cast and attached shadows.

3.1 Estimating B^*

Using singular value decomposition directly on the images leads to a biased estimate of B^* due to shadows. In addition, portions of some of the images from the Harvard database were saturated. Both shadows formed under a single light source and saturations can be detected by thresholding and labeled as "missing" – these pixels do not satisfy the linear equation $\mathbf{x} = B\mathbf{s}$. Thus, we need to estimate the 3-D linear subspace B^* with known missing values.

Define the data matrix for c images of an individual to be $X = [\mathbf{x}_1 \ldots \mathbf{x}_c]$. If there were no shadowing, X would be rank 3 and we could use SVD to decompose X into $X = B^* S^*$ where S^* is a $3 \times c$ matrix of the light source direction for all c images. To estimate a basis B^* for the 3-D linear subspace $\mathcal{L}$ from image data with missing elements, we have implemented a variation of the algorithm presented by Shum, Ikeuchi, and Reddy [14]; see also the methods of Tomasi and Kanade [15] and Jacobs [8].

The overview of this method is as follows: without doing any row or column permutations sift out all the full rows (with no invalid data) of matrix X to form a full sub-matrix $\tilde{X}$. Perform SVD on $\tilde{X}$ and get an estimate of S^*. Fix S^* and estimate each of the row of B^* independently using least squares. Then, fix B^* and estimate each of the light source direction $\mathbf{s}_i$ independently. Repeat last two steps until estimates converge. The inner workings of the algorithm are given as follows: Let $\mathbf{b}_i$ be the ith row of B^*, let $\mathbf{x}_i$ be the ith row of X. Let p be the indices of non-missing elements in $\mathbf{x}_i$, and let $\mathbf{x}_i^p$ be the row obtained by taking only the non-missing elements of $\mathbf{x}_i$, and let S^p similarly be the submatrix of S^* consisting of rows with indices in p. Then, each row of B^* is given by

$$\mathbf{b}_i = (S^p)^\dagger (\mathbf{x}_i^p)^T$$

where $(S^p)^\dagger$ is the pseudo-inverse of S^p. With the new estimate of B^* at hand, we now let $\mathbf{x}_j$ be the jth column of X, let p be the indices of non-missing elements in $\mathbf{x}_j$, and let $\mathbf{x}_j^p$ be the column obtained by taking only the non-missing elements of $\mathbf{x}_j$. Let B^p similarly be the submatrix of B^* consisting of rows with indices in p. Then, the light source directions are given by,

$$\mathbf{s}_j = (B^p)^\dagger (\mathbf{x}_j^p)$$

After the new set of light source S^* has been calculated, the last two steps can be repeated until the estimate of B^* converges. The algorithm is quite well behaved converging to the global minimum within 10-15 iterations. Though it is possible to converge to a local minimum, we never observed this in simulation or in practice.

3.2 Enforcing Integrability

To predict cast shadows, we must reconstruct a surface and to do this, the vector field B^* must correspond to an integrable normal field. Since no method has been developed to enforce integrability during the estimation of B^*, we enforce it afterwards. That is, given B^* computed as described above, we estimate a matrix $A \in GL(3)$ such that $B^* A$ corresponds to an integrable normal field; the development follows [17].

Consider a continuous surface defined as the graph of a function $z(x, y)$, and let $\mathbf{b}$ be the corresponding normal field scaled by an albedo (scalar) field. The integrability constraint for a surface is $z_{xy} = z_{yx}$ where subscripts denote partial derivatives. In turn, $\mathbf{b}$ must satisfy:

$$\left(\frac{b_1}{b_3} \right)_y = \left(\frac{b_2}{b_3} \right)_x$$

To estimate A such that $\mathbf{b}^T(x, y) = \mathbf{b}^{*^T}(x, y) A$ corresponds to a surface, we expand this out. Letting the columns of A be denoted by A_1, A_2, A_3 yields

$$(\mathbf{b}^{*^T} A_3)(\mathbf{b}_x^{*^T} A_2) - (\mathbf{b}^{*^T} A_2)(\mathbf{b}_x^{*^T} A_3) =$$
$$(\mathbf{b}^{*^T} A_3)(\mathbf{b}_y^{*^T} A_1) - (\mathbf{b}^{*^T} A_1)(\mathbf{b}_y^{*^T} A_3)$$

which can be expressed as

$$\mathbf{b}^{*^T} S_1 \mathbf{b}_x^* = \mathbf{b}^{*^T} S_2 \mathbf{b}_y^* \qquad (5)$$

where $S_1 = A_3 A_2^T - A_2 A_3^T$ and $S_2 = A_3 A_1^T - A_1 A_3^T$.

S_1 and S_2 are skew-symmetric matrices, and so they each have three degrees of freedom. Equation 5 is linear in the six elements of S_1 and S_2. From the estimate of B^* obtained using the method in Section 3.1, discrete approximations of the partial derivatives ($\mathbf{b}_x^*$ and $\mathbf{b}_y^*$) are computed, and then SVD is used to solve for the six elements of S_1 and S_2. In [17], it was shown that the elements of S_1 and S_2 are cofactors of A, and a simple method for computing A from the cofactors was presented. This procedure only determines six degrees of freedom of A. The other three correspond to the generalized bas relief (GBR) transformation [1] and can be chosen arbitrarily since GBR preserves integrability. The surface corresponding to $B^* A$ differs from the true surface by GBR, i.e., $z^*(x, y) = \lambda z(x, y) + \mu x + \nu y$ for arbitrary λ, μ, ν with $\lambda \neq 0$.

3.3 Generating the Height Function

Having estimated the matrix B^* and then enforcing integrability, we now calculate the height function $z(x, y)$ of the face so that cast shadows can be predicted. Note that the reconstruction of the height is not Euclidean, but a representative element of the orbit under a GBR transformation. For each normal $\mathbf{b}_i$ the derivatives of $z(x, y)$ with respect to x and y are given by the following equations

$$p = \frac{\partial z}{\partial x} = -\frac{b_{i1}}{b_{i3}}, \qquad q = \frac{\partial z}{\partial y} = -\frac{b_{i2}}{b_{i3}}.$$

In order to find $z(x, y)$, we use the variational approach presented in [7]. A surface $z(x, y)$ is fit to the given components of the gradient p and q by minimizing the functional

$$\int \int_\Omega (z_x - p)^2 + (z_y - q)^2 \, dx \, dy,$$

whose Euler equation reduces to $\nabla^2 z = p_x + q_y$. We need to constrain the solution of the Euler equation, and this is achieved by the following natural boundary conditions $(z_x, z_y) \cdot \mathbf{n} = (p, q) \cdot \mathbf{n}$ where $\mathbf{n} = (-dy/ds, dx/ds)$ is a normal vector to the boundary curve $\partial \Omega$, and s is the arc-length along the boundary [7]. Thus, the component of (z_x, z_y) normal to the chosen boundary curve must match the normal component of (p, q). An iterative scheme using a discrete approximation of the Laplacian can be used to generate a height function of the face [7].

Once the height function has been determined, it is a simple matter to modify the illumination cone representation to incorporate *cast* shadows. Using ray-tracing techniques, we can determine the cast shadow regions and correct the extreme rays of $\mathcal{C}^*$.

Figure 2 demonstrates the process of constructing the cone $\mathcal{C}^*$. Figure 2.a shows the training images for one individual in the database. Figure 2.b shows the columns of the matrix B^*. Figure 2.c shows the reconstruction of the surface up to a GBR transformation. The left column of Fig. 2.d shows sample images in the database; the middle column shows the closest image in the illumination cone without cast shadows; and the right column shows the closest image in the illumination cone with cast shadows.

4 Recognition

The cone $\mathcal{C}^*$ can be used in a natural way for face recognition, and in experiments described below, we compare three recognition algorithms to the proposed method. From a set of face images labeled with the person's identity (*the learning set*) and an unlabeled set of face images from the same group of people (*the test set*), each algorithm is used to identify the person in the test images. For more details of the comparison algorithms, see [2]. We assume that the face has been located and aligned within the image.

a.

b.

c.

d.
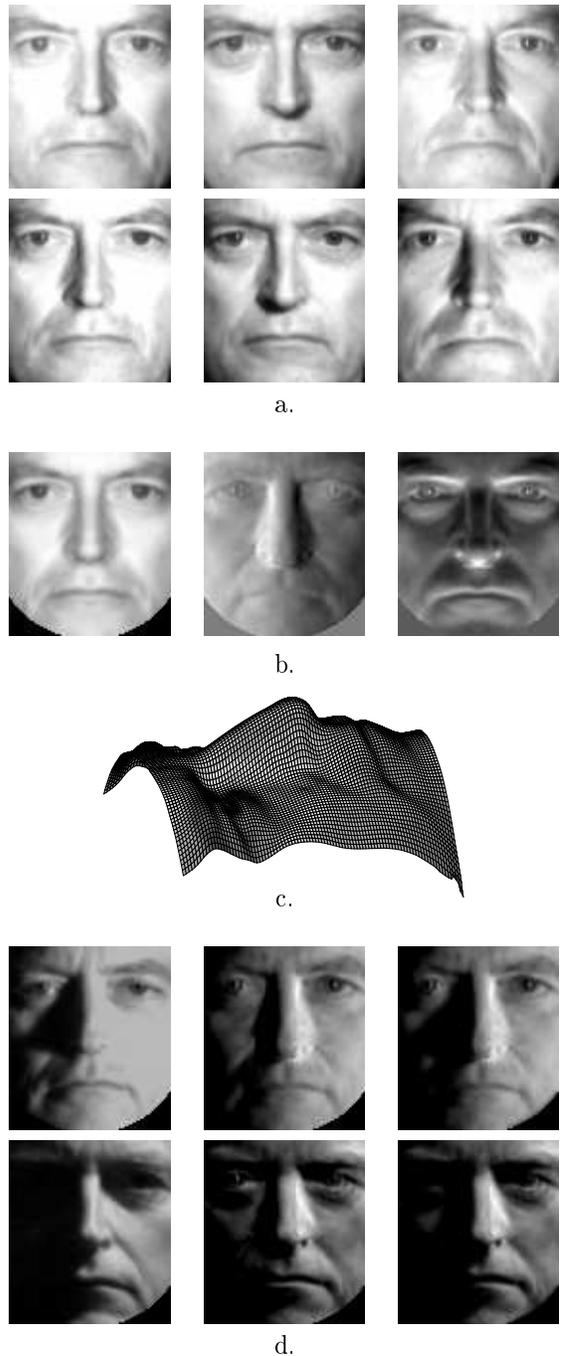
Figure 2: The figure demonstrates the process of constructing the cone $\mathcal{C}^*$. a) the training images. b) matrix B^*. c) reconstruction up to a GBR transformation. d) sample images from database (left column); closest image in illumination cone without cast shadows (middle column); and closest image in illumination cone with cast shadows (right column).

The simplest recognition scheme is a nearest neighbor classifier in the image space [4]. An image in the test set is recognized (classified) by assigning to it the

label of the closest point in the learning set, where distances are measured in the image space. If all of the images are normalized to have zero mean and unit variance, this procedure is equivalent to choosing the image in the learning set that best *correlates* with the test image. Because of the normalization process, the result is independent of light source intensity.

As correlation methods are computationally expensive and require great amounts of storage, it is natural to pursue dimensionality reduction schemes. A technique now commonly used in computer vision – particularly in face recognition – is principal components analysis (PCA) which is popularly known as *Eigenfaces* [5, 11, 9, 16]. Given a collection of training images $\mathbf{x}_i \in \mathbb{R}^n$, a linear projection of each image $\mathbf{y}_i = W\mathbf{x}_i$ to an f-dimensional feature space is performed. A face in a test image $\mathbf{x}$ is recognized by projecting $\mathbf{x}$ to the feature space, and nearest neighbor classification is performed in $\mathbb{R}^f$. The projection matrix W is chosen to maximize the scatter of all projected samples. It has been shown that when f equals the number of training images, the Eigenface and Correlation methods are equivalent (See [2, 11]). One proposed method for handling illumination variation in PCA is to discard from W the three most significant principal components; in practice, this yields better recognition performance [2].

A third approach is to model the illumination variation of each face as a three-dimensional linear subspace $\mathcal{L}$ as described in Section 2. To perform recognition, we simply compute the distance of the test image to each linear subspace and choose the face corresponding to the shortest distance. We call this recognition scheme the *Linear Subspace* method [1]; it is a variant of the photometric alignment method proposed in [13], and related to [6, 12]. While this is expected to model the variation in intensity when the surface is completely illuminated, it does not model shadowing.

Finally, given a test image $\mathbf{x}$, recognition using *illumination cones* is performed by first computing the distance of the test image to each cone, and then choosing the face that corresponds to the shortest distance. Since each cone is convex, the distance can be found by solving a convex optimization problem. In particular, the non-negative linear least squares technique contained in Matlab was used in our implementation, and this algorithm has computational complexity $O(n\,e^2)$ where n is the number of pixels and e is the number of extreme rays.

5 Experimental Results

To test the effectiveness of these recognition algorithms, we performed a series of experiments on a database from the Harvard Robotics Laboratory in which lighting had been systematically varied [5, 6]. In each image in this database, a subject held his/her head steady while being illuminated by a dominant light source. The space of light source directions,
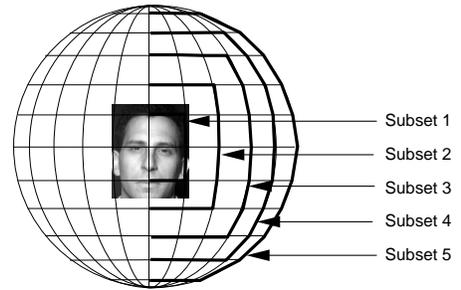
Figure 3: The highlighted lines of longitude and latitude indicate the light source directions for Subsets 1 through 5. Each intersection of a longitudinal and latitudinal line on the right side of the illustration has a corresponding image in the database.

which can be parameterized by spherical angles, was then sampled in 15° increments. See Figure 3. From this database, we used 660 images of 10 people (66 of each). We extracted five subsets to quantify the effects of varying lighting. Sample images from each subset are shown in Fig. 1. Subset 1 (respectively 2, 3, 4, 5) contains 30 (respectively 90, 130, 170, 210) images for which both of the longitudinal and latitudinal angles of light source direction are within 15° (respectively $30°, 45°, 60°, 75°$) of the camera axis.

All of the images were cropped within the face so that the contour of the head was excluded. For the Eigenface and correlation tests, the images were normalized to have zero mean and unit variance, as this improved the performance of these methods. For the Eigenface method, we used twenty principal components – recall that performance approaches correlation as the dimension of the feature space is increased [2, 11]. Since the first three principal components are primarily due to lighting variation and since recognition rates can be improved by eliminating them, error rates are also presented when principal components four through twenty-three are used. For the cone experiments, we tested two variations: in the first variation (cones-attached), the representation was constructed ignoring cast shadows and so extreme rays were generated directly from Eq. 3. In the second variation (Cones-cast), the representation was constructed as described in Section 3.

Mirroring the extrapolation experiment described in [2], each method was trained on samples from Subset 1 and then tested using samples from Subsets 2, 3, 4 and 5. (Note that when tested on Subset 1, all methods performed without error). Figure 4 shows the result from this experiment.

6 Discussion

From the results of this experiment, we draw the following conclusions:

- The illumination cone representation outperforms all of the other techniques.
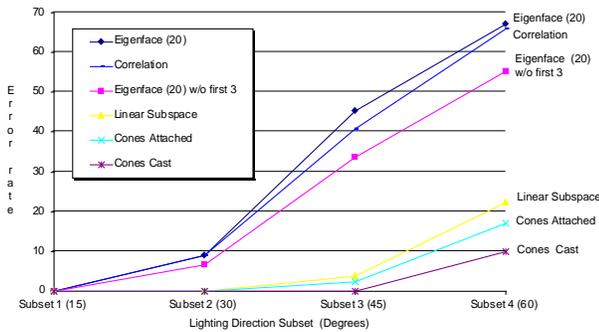
Extrapolating from Subset 1				
Method	Error Rate (%)			
	Subset 2	Subset 3	Subset 4	Subset 5
Correlation	8.9	40.8	65.9	84.1
Eigenface	8.9	45.4	67.1	84.1
Eigenface w/o 1st 3	6.7	33.8	55.3	78.1
Linear subspace	0.0	3.8	22.4	50.7
Cones-attached	0.0	2.3	17.1	44.8
Cones-cast	0.0	0.0	10.0	37.0

Figure 4: **Extrapolation:** When each of the methods is trained on images with near frontal illumination (Subset 1), the graph and corresponding table show the relative performance under extreme light source conditions.

- When cast shadows are included in the illumination cone, error rates are improved.

- For very extreme illumination (Subset 5), the Correlation and Eigenface methods completely break down, and exhibit results that are slightly better than chance (90% error rate). The cone method performs significantly better, but certainly not well enough to be usable in practice. At this point, more experimentation is required to determine if recognition rates can be improved by either using more sampled extreme rays or by improving the image formation model.

The experiment described above was limited to the available dataset from the Harvard Robotics Laboratory. To perform more extensive experimentation, we are constructing a geodesic lighting rig that supports 64 xenon strobes. Using this rig, we will be able to modify the illumination at frame rates and gather an extensive image database covering a broader range of lighting conditions including multiple sources. The speed of acquisition will also permit us to readily obtain images of a large number of individuals. We will then perform more extensive experimentation with this newly gathered database.

References

[1] P. Belhumeur, D. Kriegman, and A. Yuille. The basrelief ambiguity. In *Proc. IEEE Conf. on Comp. Vision and Patt. Recog.*, pages 1040–1046, 1997.

[2] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman. Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection. *IEEE Trans. Pattern Anal. Mach. Intelligence*, 19(7):711–720, 1997. Special Issue on Face Recognition.

[3] P. N. Belhumeur and D. J. Kriegman. What is the set of images of an object under all possible lighting conditions. In *Proc. IEEE Conf. on Comp. Vision and Patt. Recog.*, pages 270–277, 1996.

[4] R. Brunelli and T. Poggio. Face recognition: Features vs templates. *IEEE Trans. Pattern Anal. Mach. Intelligence*, 15(10):1042–1053, 1993.

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

[6] P. Hallinan. *A Deformable Model for Face Recognition Under Arbitrary Lighting Conditions*. PhD thesis, Harvard University, 1995.

[7] B. Horn and M. Brooks. The variational approach to shape from shading. *Computer Vision, Graphics and Image Processing*, 35:174–208, 1992.

[8] D. Jacobs. Linear fitting with missing data: Applications to structure-from-motion and to characterizing intensity images. In *CVPR97*, pages 206–212, 1997.

[9] L. Sirovitch and M. Kirby. Low-dimensional procedure for the characterization of human faces. *J. Optical Soc. of America A*, 2:519–524, 1987.

[10] Y. Moses, Y. Adini, and S. Ullman. Face recognition: The problem of compensating for changes in illumination direction. In *European Conf. on Computer Vision*, pages 286–296, 1994.

[11] H. Murase and S. Nayar. Visual learning and recognition of 3-D objects from appearence. *Int. J. Computer Vision*, 14(5–24), 1995.

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

[13] A. Shashua. On photometric issues to feature-based object recognition. *Int. J. Computer Vision*, 21:99–122, 1997.

[14] H. Shum, K. Ikeuchi, and R. Reddy. Principal component analysis with missing data and its application to polyhedral object modeling. *PAMI*, 17(9):854–867, September 1995.

[15] C. Tomasi and T. Kanade. Shape and motion from image streams under orthography: a factorization method. *International Journal of Computer Vision*, 9(2):134–154, 1992.

[16] M. Turk and A. Pentland. Eigenfaces for recognition. *J. of Cognitive Neuroscience*, 3(1), 1991.

[17] A. Yuille and D. Snow. Shape and albedo from multiple images using integrability. In *Proc. IEEE Conf. on Comp. Vision and Patt. Recog.*, pages 158–164, 1997.