1996 European Conference on Computer Vision

Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection

Peter N. Belhumeur* João P. Hespanha** David J. Kriegman***

Dept. of Electrical Engineering, Yale University, New Haven, CT 06520-8267

Abstract. We develop a face recognition algorithm which is insensitive to gross variation in lighting direction and facial expression. Taking a pattern classification approach, we consider each pixel in an image as a coordinate in a high-dimensional space. We take advantage of the observation that the images of a particular face under varying illumination direction lie in a 3-D linear subspace of the high dimensional feature space – if the face is a Lambertian surface without self-shadowing. However, since faces are not truly Lambertian surfaces and do indeed produce self-shadowing, images will deviate from this linear subspace. Rather than explicitly modeling this deviation, we project the image into a subspace in a manner which discounts those regions of the face with large deviation. Our projection method is based on Fisher's Linear Discriminant and produces well separated classes in a low-dimensional subspace even under severe variation in lighting and facial expressions. The Eigenface technique, another method based on linearly projecting the image space to a low dimensional subspace, has similar computational requirements. Yet, extensive experimental results demonstrate that the proposed "Fisherface" method has error rates that are significantly lower than those of the Eigenface technique when tested on the same database.

1 Introduction

Within the last several years, numerous algorithms have been proposed for face recognition; for detailed surveys see [4, 24]. While much progress has been made toward recognizing faces under small variations in lighting, facial expression and pose, reliable techniques for recognition under more extreme variations have proven elusive.

In this paper we outline a new approach for face recognition – one that is insensitive to extreme variations in lighting and facial expressions. Note that lighting variability includes not only intensity, but also direction and number of

* P. Belhumeur was supported by ARO Grant DAAH04-95-1-0494.

** J. Hespanha was supported by NSF Grant ECS-9206021, AFOSR Grant F49620-94-1-0181, and ARO Grant DAAH04-95-1-0114.

*** D. Kriegman was supported by NSF under an NYI, IRI-9257990 and by ONR N00014-93-1-0305

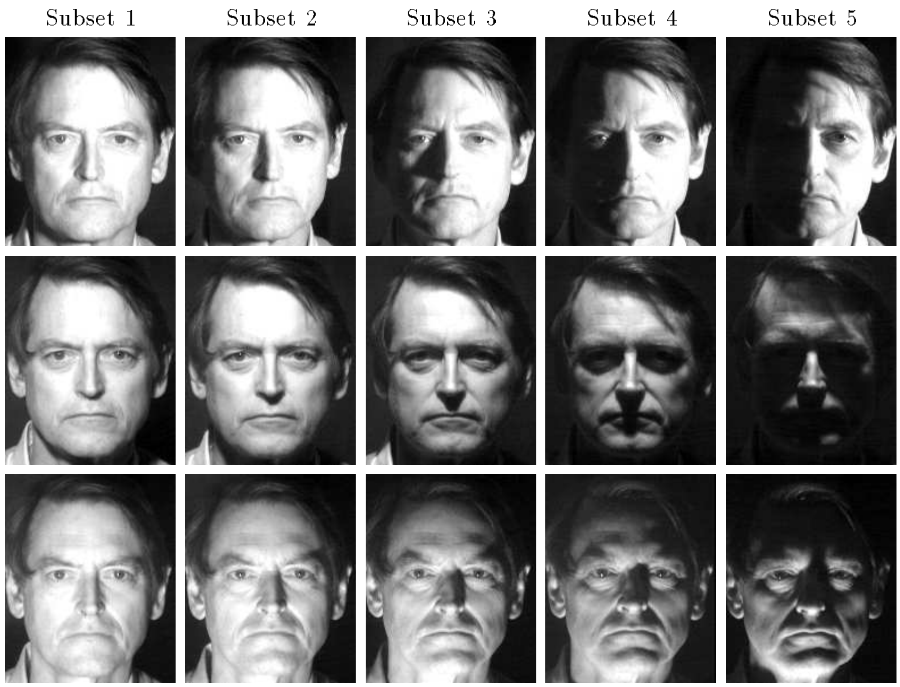

Fig. 1. The same person seen under varying lighting conditions can appear dramatically different. These images are taken from the Harvard database which is described in Section 3.1.

light sources. As seen in Fig. 1, the same person, with the same facial expression, seen from the same viewpoint can appear dramatically different when light sources illuminate the face from different directions.

Our approach to face recognition exploits two observations:

1. For a Lambertian surface without self-shadowing, all of the images of a particular face from a fixed viewpoint will lie in a 3-D linear subspace of the high-dimensional image space [25].
2. Because of expressions, regions of self-shadowing and specularity, the above observation does not exactly apply to faces. In practice, certain regions of the face may have a variability from image to image that often deviates drastically from the linear subspace and, consequently, are less reliable for recognition.

We make use of these observations by finding a linear projection of the faces from the high-dimensional image space to a significantly lower dimensional feature space which is insensitive both to variation in lighting direction and facial expression. We choose projection directions that are nearly orthogonal to the within-class scatter, projecting away variations in lighting and facial expression while maintaining discriminability. Our method Fisherfaces, a derivative of Fisher's Linear Discriminant (FLD) [9, 10], maximizes the ratio of between-class scatter to that of within-class scatter.

The Eigenface method is also based on linearly projecting the image space to a low dimensional feature space [27, 28, 29]. However, the Eigenface method, which uses principal components analysis (PCA) for dimensionality reduction, yields projection directions that maximize the total scatter across all classes, i.e. all images of all faces. In choosing the projection which maximizes total scatter, PCA retains some of the unwanted variations due to lighting and facial expression. As illustrated in Fig. 1 and stated by Moses, Adini, and Ullman, "the variations between the images of the same face due to illumination and viewing direction are almost always larger than image variations due to change in face identity" [21]. Thus, while the PCA projections are optimal for reconstruction from a low dimensional basis, they may not be optimal from a discrimination standpoint.

We should point out that Fisher's Linear Discriminant [10] is a "classical" technique in pattern recognition [9] that was developed by Robert Fisher in 1936 for taxonomic classification. Depending upon the features being used, it has been applied in different ways in computer vision and even in face recognition. Cheng *et al.* presented a method that used Fisher's discriminator for face recognition where features were obtained by a polar quantization of the shape [6]. Contemporaneous with our work [15], Cui, Swets, and Weng applied Fisher's discriminator (using different terminology, they call it the Most Discriminating Feature – MDF) in a method for recognizing hand gestures [8]. Though no implementation is reported, they also suggest that the method can be applied to face recognition under variable illumination.

In the sections to follow, we will compare four methods for face recognition under variation in lighting and facial expression: correlation, a variant of the linear subspace method suggested by [25], the Eigenface method [27, 28, 29], and the Fisherface method developed here. The comparisons are done on a database of 500 images created externally by Hallinan [13, 14] and a database of 176 images created at Yale. The results of the tests on both databases shows that the Fisherface method performs significantly better than any of the other three methods. Yet, no claim is made about the relative performance of these algorithms on much larger databases.

We should also point out that we have made no attempt to deal with variation in pose. An appearance-based method such as ours can be easily extended to handle limited pose variation using either a multiple-view representation such as Pentland, Moghaddam, and Starner's View-based Eigenspace [23] or Murase and Nayar's Appearance Manifolds [22]. Other approaches to face recognition that accommodate pose variation include [2, 11]. Furthermore, we assume that the face has been located and aligned within the image, as there are numerous methods for finding faces in scenes [5, 7, 17, 18, 19, 20, 28].

2 Methods

The problem can be simply stated: Given 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*), identify the name of each person in the test images.

In this section, we examine four pattern classification techniques for solving the face recognition problem, comparing methods that have become quite popular in the face recognition literature, i.e. correlation [3] and Eigenface methods

[27, 28, 29], with alternative methods developed by the authors. We approach this problem within the pattern classification paradigm, considering each of the pixel values in a sample image as a coordinate in a high-dimensional space (*the image space*).

2.1 Correlation

Perhaps, the simplest classification scheme is a nearest neighbor classifier in the image space [3]. Under this scheme, an image in the test set is recognized 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 have been normalized to be zero mean and have unit variance, then 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 and the effects of a video camera's automatic gain control.

This procedure, which will subsequently be referred to as correlation, has several well-known disadvantages. First, if the images in the learning set and test set are gathered under varying lighting conditions, then the corresponding points in the image space will not be tightly clustered. So in order for this method to work reliably under variations in lighting, we would need a learning set which densely sampled the continuum of possible lighting conditions. Second, correlation is computationally expensive. For recognition, we must correlate the image of the test face with each image in the learning set; in an effort to reduce the computation time, implementors [12] of the algorithm described in [3] developed special purpose VLSI hardware. Third, it requires large amounts of storage – the learning set must contain numerous images of each person.

2.2 Eigenfaces

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 for dimensionality reduction in computer vision – particularly in face recognition – is principal components analysis (PCA) [13, 22, 27, 28, 29]. PCA techniques, also known as Karhunen-Loeve methods, choose a dimensionality reducing linear projection that maximizes the scatter of all projected samples.

More formally, let us consider a set of N sample images $\{x_1, x_2, \ldots, x_N\}$ taking values in an n-dimensional feature space, and assume that each image belongs to one of c classes $\{\chi_1, \chi_2, \ldots, \chi_c\}$. Let us also consider a linear transformation mapping the original n-dimensional feature space into an m-dimensional feature space, where $m < n$. Denoting by $W \in \mathbb{R}^{n \times m}$ a matrix with orthonormal columns, the new feature vectors $y_k \in \mathbb{R}^m$ are defined by the following linear transformation:

$$y_k = W^T x_k\,, \qquad k = 1, 2, \ldots, N.$$

Let the total scatter matrix S_T be defined as

$$S_T = \sum_{k=1}^{N} (x_k - \mu)(x_k - \mu)^T$$

where $\mu \in \mathbb{R}^n$ is the mean image of all samples.

Note that after applying the linear transformation, the scatter of the transformed feature vectors $\{y_1, y_2, \ldots, y_N\}$ is $W^T S W$. In PCA, the optimal projection W_{opt} is chosen to maximize the determinant of the total scatter matrix of the projected samples, i.e.

$$W_{opt} = \arg\max_W |W^T S_T W| = [\, w_1 \; w_2 \; \ldots \; w_m \,] \qquad (1)$$

where $\{w_i \mid i = 1, 2, \ldots, m\}$ is the set of n-dimensional eigenvectors of S_T corresponding to the set of decreasing eigenvalues. Since these eigenvectors have the same dimension as the original images, they are referred to as Eigenpictures in [27] and Eigenfaces in [28, 29].

A drawback of this approach is that the scatter being maximized is not only due to the between-class scatter that is useful for classification, but also the within-class scatter that, for classification purposes, is unwanted information. Recall the comment by Moses, Adini and Ullman [21]: Much of the variation from one image to the next is due to illumination changes. Thus if PCA is presented with images of faces under varying illumination, the projection matrix W_{opt} will contain principal components (i.e. Eigenfaces) which retain, in the projected feature space, the variation due lighting. Consequently, the points in projected space will not be well clustered, and worse, the classes may be smeared together.

It has been suggested that by throwing out the first several principal components, the variation due to lighting is reduced. The hope is that if the first principal components capture the variation due to lighting, then better clustering of projected samples is achieved by ignoring them. Yet it is unlikely that the first several principal components correspond solely to variation in lighting; as a consequence, information that is useful for discrimination may be lost.

2.3 Linear Subspaces

Both correlation and the Eigenface method are expected to suffer under variation in lighting direction. Neither method exploits the observation that for a Lambertian surface without self-shadowing, the images of a particular face lie in a 3-D linear subspace.

Consider a point p in a Lambertian surface and a collimated light source characterized by a vector $s \in \mathbb{R}^3$, such that the direction of s gives the direction of the light rays and $||s||$ gives the intensity of the light source. The irradiance at the point p is given by

$$E(p) = a(p) < n(p), s > \tag{2}$$

where $n(p)$ is the unit inward normal vector to the surface at the point p, and $a(p)$ is the albedo of the surface at p [16]. This shows that the irradiance at the point p, and hence the gray level seen by a camera, is linear on $s \in \mathbb{R}^3$. Therefore, in the absence of self-shadowing, given three images of a Lambertian surface from the same viewpoint taken under three known, linearly independent light source directions, the albedo and surface normal can be recovered; this is the well known method of photometric stereo [26, 30]. Alternatively, one can reconstruct the image of the surface under an arbitrary lighting direction by a linear combination of the three original images, see [25].

For classification, this fact has great importance: It shows that for a fixed viewpoint, all images of a Lambertian surface lie in a 3-D linear subspace embedded in the high-dimensional image space. This observation suggests a simple classification algorithm to recognize Lambertian surfaces – invariant under lighting conditions.

For each face, use three or more images taken under different lighting directions to construct a 3-D basis for the linear subspace. Note that the three basis vectors have the same dimensionality as the training images and can be thought of as basis images. To perform recognition, we simply compute the distance of a new image to each linear subspace and choose the face corresponding to the shortest distance. We call this recognition scheme the Linear Subspace method. We should point out that this method is a variant of the photometric alignment method proposed in [25] and, although it is not yet in press, the Linear Subspace method can be thought of as special case of the more elaborate recognition method described in [14].

If there is no noise or self-shadowing, the Linear Subspace algorithm would achieve error free classification under *any* lighting conditions, provided the surfaces obey the Lambertian reflectance model. Nevertheless, there are several compelling reasons to look elsewhere. First, due to self-shadowing, specularities, and facial expressions, some regions of the face have variability that does not agree with the linear subspace model. Given enough images of faces, we should be able to learn which regions are good for recognition and which regions are not. Second, to recognize a test image we must measure the distance to the linear subspace for each person. While this in an improvement over a correlation scheme that needs a large number of images for each class, it is still too computationally expensive. Finally, from a storage standpoint, the Linear Subspace algorithm must keep three images in memory for every person.

2.4 Fisherfaces

The Linear Subspace algorithm takes advantage of the fact that under ideal conditions the classes are linearly separable. Yet, one can perform dimensionality reduction using linear projection and still preserve linear separability; error free classification under any lighting conditions is still possible in the lower dimensional feature space using linear decision boundaries. This is a strong argument in favor of using linear methods for dimensionality reduction in the face recognition problem, at least when one seeks insensitivity to lighting conditions.

Here we argue that using class specific linear methods for dimensionality reduction and simple classifiers in the reduced feature space one gets better recognition rates in substantially less time than with the Linear Subspace method. Since the learning set is labeled, it makes sense to use this information to build a more reliable method for reducing the dimensionality of the feature space. Fisher's Linear Discriminant (FLD) [10] is an example of a *class specific method*, in the sense that it tries to "shape" the scatter in order to make it more reliable for classification. This method selects W in such a way that the ratio of the between-class scatter and the within-class scatter is maximized. Let the between-class scatter matrix be defined as

$$S_B = \sum_{i=1}^{c} |\chi_i| \, (\mu_i - \mu)(\mu_i - \mu)^T$$

and the within-class scatter matrix be defined as

$$S_W = \sum_{i=1}^{c} \sum_{x_k \in \chi_i} (x_k - \mu_i)(x_k - \mu_i)^T$$

where μ_i is the mean image of class χ_i, and $|\chi_i|$ is the number of samples in class χ_i. If S_W is nonsingular, the optimal projection W_{opt} is chosen as that which maximizes the ratio of the determinant of the between-class scatter matrix of

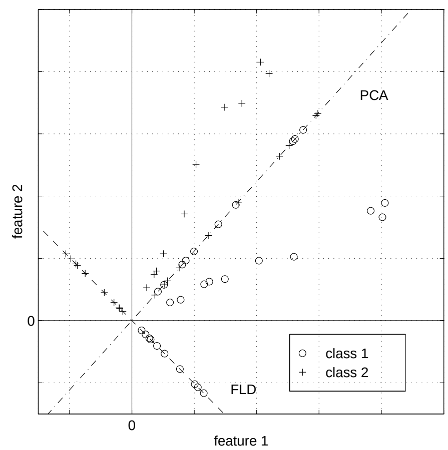

Fig. 2. A comparison of principal component analysis (PCA) and Fisher's linear discriminant (FLD) for a two class problem where data for each class lies near a linear subspace.

the projected samples to the determinant of the within-class scatter matrix of the projected samples, i.e.

$$W_{opt} = \arg\max_{W} \frac{|W^T S_B W|}{|W^T S_W W|} = [\, w_1 \; w_2 \; \dots \; w_m \,] \tag{3}$$

where $\{w_i \mid i = 1, 2, \dots, m\}$ is the set of generalized eigenvectors of S_B and S_W corresponding to set of decreasing generalized eigenvalues $\{\lambda_i \mid i = 1, 2, \dots, m\}$, i.e.

$$S_B w_i = \lambda_i S_W w_i \, , \qquad i = 1, 2, \dots, m.$$

Note that an upper bound on m is $c-1$ where c is the number of classes. See [9].

To illustrate the benefits of the class specific linear projections, we constructed a low dimensional analogue to the classification problem in which the samples from each class lie near a linear subspace. Figure 2 is a comparison of PCA and FLD for a two-class problem in which the samples from each class are randomly perturbed in a direction perpendicular to the linear subspace. For this example $N = 20$, $n = 2$, and $m = 1$. So the samples from each class lie near a line in the 2-D feature space. Both PCA and FLD have been used to project the points from 2-D down to 1-D. Comparing the two projections in the figure, PCA actually smears the classes together so that they are no longer linearly separable in the projected space. It is clear that although PCA achieves larger total scatter, FLD achieves greater between-class scatter, and consequently classification becomes easier.

In the face recognition problem one is confronted with the difficulty that the within-class scatter matrix $S_W \in \mathbb{R}^{n \times n}$ is always singular. This stems from the fact that the rank of S_W is less than $N - c$, and in general, the number of pixels in each image (n) is much larger than the number of images in the learning set (N). This means that it is possible to chose the matrix W such that the within-class scatter of the projected samples can be made exactly zero.

In order to overcome the complication of a singular S_W, we propose an alternative to the criterion in Eq. 3. This method, which we call Fisherfaces, avoids this problem by projecting the image set to a lower dimensional space so that the resulting within-class scatter matrix S_W is nonsingular. This is achieved by using PCA to reduce the dimension of the feature space to $N - c$ and then, applying the standard FLD defined by Eq. 3 to reduce the dimension to $c - 1$. More formally, W_{opt} is given by

$$W_{opt} = W_{fld} \, W_{pca} \tag{4}$$

where

$$W_{pca} = \arg\max_{W} |W^T S_T W|$$

$$W_{fld} = \arg\max_{W} \frac{|W^T W_{pca}^T S_B W_{pca} W|}{|W^T W_{pca}^T S_W W_{pca} W|}.$$

Note that in computing W_{pca} we have thrown away only the smallest c principal components.

There are certainly other ways of reducing the within-class scatter while preserving between-class scatter. For example, a second method which we are currently investigating chooses W to maximize the between-class scatter of the projected samples after having first reduced the within-class scatter. Taken to an extreme, we can maximize the between-class scatter of the projected samples subject to the constraint that the within-class scatter is zero, i.e.

$$W_{opt} = \arg\max_{W \in \mathcal{W}} |W^T S_B W| \tag{5}$$

where $\mathcal{W}$ is the set of $n \times m$ matrices contained in the kernel of S_W.

3 Experimental Results

In this section we will present and discuss each of the aforementioned face recognition techniques using two different databases. Because of the specific hypotheses that we wanted to test about the relative performance of the considered algorithms, many of the standard databases were inappropriate. So we have used a database from the Harvard Robotics Laboratory in which lighting has been systematically varied. Secondly, we have constructed a database at Yale that includes variation in both facial expression and lighting.[4]

3.1 Variation in Lighting

The first experiment was designed to test the hypothesis that under variable illumination, face recognition algorithms will perform better if they exploit the fact that images of a Lambertian surface lie in a linear subspace. More specifically, the recognition error rates for all four algorithms described in Section 2 will be compared using an image database constructed by Hallinan at the Harvard Robotics Laboratory [13, 14]. 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, which can be parameterized by spherical angles, was then sampled in 15° increments. From a subset of 225 images of five people in this database, we extracted five subsets to quantify the effects of varying lighting. Sample images from each subset are shown in Fig. 1.

[4] The Yale database is available by anonymous ftp from daneel.eng.yale.edu.

Subset 1 contains 30 images for which both of the longitudinal and latitudinal angles of light source direction are within 15° of the camera axis.
Subset 2 contains 45 images for which the greater of the longitudinal and latitudinal angles of light source direction are 30° from the camera axis.
Subset 3 contains 65 images for which the greater of the longitudinal and latitudinal angles of light source direction are 45° from the camera axis.
Subset 4 contains 85 images for which the greater of the longitudinal and latitudinal angles of light source direction are 60° from the camera axis.
Subset 5 contains 105 images for which the greater of the longitudinal and latitudinal angles of light source direction are 75° from the camera axis.

For all experiments, classification was performed using a nearest neighbor classifier. All training images of an individual were projected into the feature space. The images were cropped within the face so that the contour of the head was excluded.[5] 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, results are shown when ten principal components are used. Since it has been suggested that the first three principal components are primarily due to lighting variation and that recognition rates can be improved by eliminating them, error rates are also presented using principal components four through thirteen. Since there are 30 images in the training set, correlation is equivalent to the Eigenface method using 29 principal components.

We performed two experiments on the Harvard Database: extrapolation and interpolation. In the extrapolation experiment, each method was trained on samples from Subset 1 and then tested using samples from Subsets 1, 2 and 3.[6] Figure 3 shows the result from this experiment.

In the interpolation experiment, each method was trained on Subsets 1 and 5 and then tested the methods on Subsets 2, 3 and 4. Figure 4 shows the result from this experiment.

These two experiments reveal a number of interesting points:

1. All of the algorithms perform perfectly when lighting is nearly frontal. However as lighting is moved off axis, there is a significant performance difference between the two class-specific methods and the Eigenface method.
2. It has also been noted that the Eigenface method is equivalent to correlation when the number of Eigenfaces equals the size of the training set [22], and since performance increases with the dimension of the Eigenspace, the Eigenface method should do no better than correlation [3]. This is empirically demonstrated as well.
3. In the Eigenface method, removing the first three principal components results in better performance under variable lighting conditions.
4. While the Linear Subspace method has error rates that are competitive with the Fisherface method, it requires storing more than three times as much information and takes three times as long.

---

5 We have observed that the error rates are reduced for all methods when the contour is included and the subject is in front of a uniform background. However, all methods performed worse when the background varies.

6 To test the methods with an image from Subset 1, that image was removed from the training set, i.e. we employed the "leaving-one-out" strategy [9].

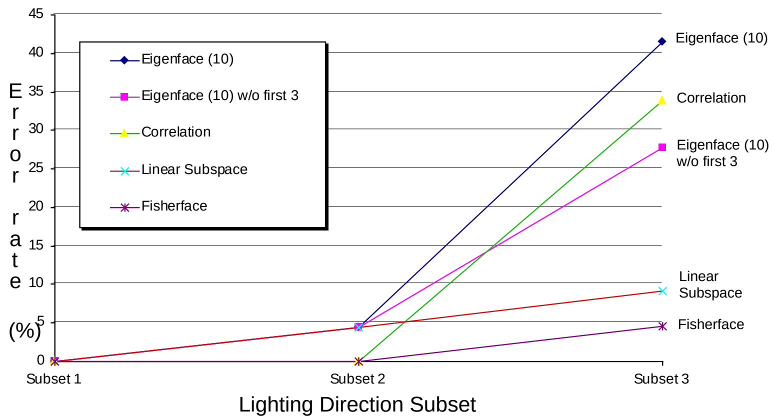

Fig. 3. Extrapolation: When each of the methods is trained on images with near frontal illumination (Subset 1), the graph and corresponding table show their relative performance under extreme light source conditions.

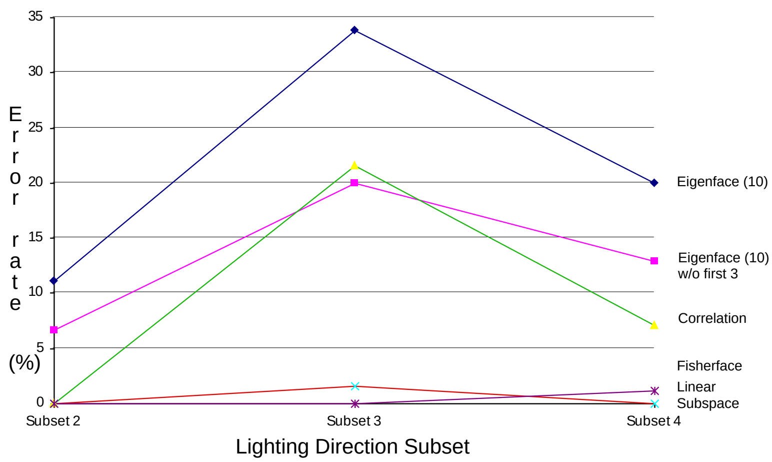

Fig. 4. Interpolation: When each of the methods is trained on images from both near frontal and extreme lighting (Subsets 1 and 5), the graph and corresponding table show the methods' relative performance under intermediate lighting conditions.

5. The Fisherface method had error rates many times lower than the Eigenface method and required less computation time.

3.2 Variation in Facial Expression, Eyewear, and Lighting

Using a second database constructed at the Yale Vision and Robotics Lab, we constructed tests to determine how the methods compared under a different range of conditions. For sixteen subjects, ten images were acquired during one session in front of a simple background. Subjects included females and males (some with facial hair), and some wore glasses. Figure 5 shows ten images of one subject. The first image was taken under ambient lighting in a neutral facial expression, and the person wore his/her glasses when appropriate. In the second image, the glasses were removed if glasses were not normally worn; otherwise,

Fig. 5. The Yale database contains 160 frontal face images covering sixteen individuals taken under ten different conditions: A normal image under ambient lighting, one with or without glasses, three images taken with different point light sources, and five different facial expressions.

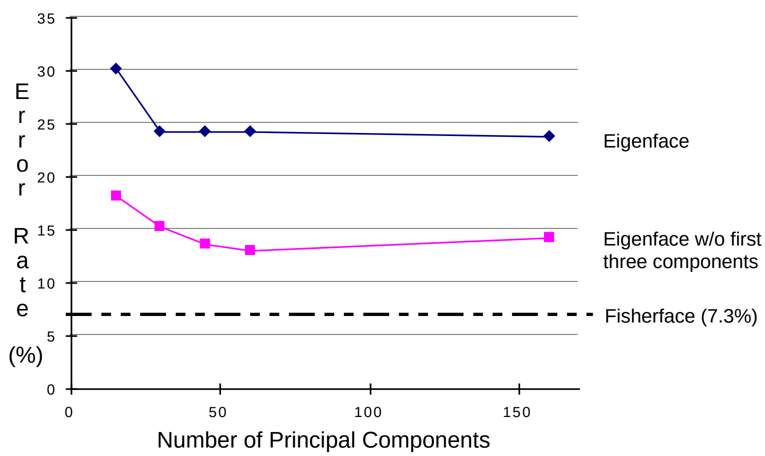

Fig. 6. As demonstrated on the Yale Database, the variation in performance of the Eigenface method depends on the number of principal components retained. Dropping the first three appears to improve performance.

a pair of borrowed glasses were worn. Images 3-5 were acquired by illuminating the face in a neutral expression with a Luxolamp in three position. The last five images were acquired under ambient lighting with different expressions (happy, sad, winking, sleepy, and surprised). 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. The images were manually centered and cropped to two different scales: The larger images included the *full face* and part of the background while the *closely cropped* ones included internal structures such as the brow, eyes, nose, mouth and chin but did not extend to the occluding contour.

In this test, error rates were determined by the "leaving-one-out" strategy [9]: To classify an image of a person, that image was removed from the data set and the dimensionality reduction matrix W was computed. All images in the database, excluding the test image, were then projected down into the reduced space to be used for classification. Recognition was performed using a nearest neighbor classier. Note that for this test, each person in the learning set is represented by the projection of ten images, except for the test person who is represented by only nine.

In general, the performance of the Eigenface method varies with the number of principal components. So, before comparing the Linear Subspace and Fisher-

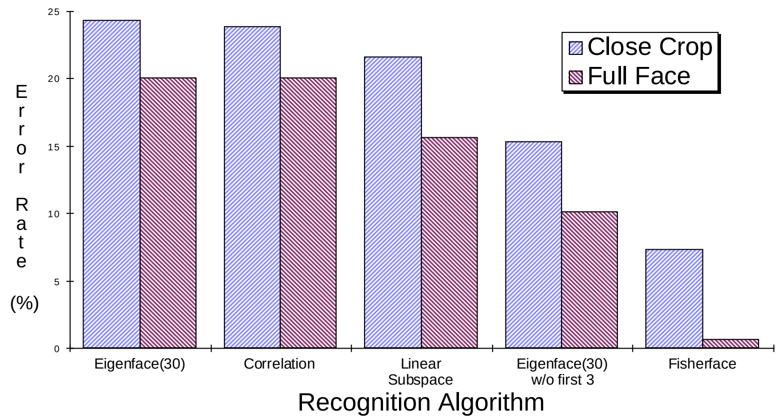

Fig. 7. The graph and corresponding table show the relative performance of the algorithms when applied to the Yale database which contains variation in facial expression and lighting.

face methods with the Eigenface method, we first performed an experiment to determine the number of principal components that results in the lowest error rate. Figure 6 shows a plot of error rate vs. the number of principal components when the initial three principal components were retained and when they were dropped.

The relative performance of the algorithms is self evident in Fig. 7. The Fisherface method had error rates that were better than half that of any other method. It seems that the Fisherface method learns the set of projections which performs well over a range of lighting variation, facial expression variation, and presence of glasses.

Note that the Linear Subspace method faired comparatively worse in this experiment than in the lighting experiments in the previous section. Because of variation in facial expression, the images no longer lie in a linear subspace. Since the Fisherface method tends to discount those portions of the image that are not significant for recognizing an individual, the resulting projections W tend to mask the regions of the face that are highly variable. For example, the area around the mouth is discounted since it varies quite a bit for different facial expressions. On the other hand, the nose, cheeks and brow are stable over the within-class variation and are more significant for recognition. Thus, we conjecture that Fisherface methods, which tend to reduce within-class scatter for all classes, should produce projection directions that are good for recognizing other faces besides the training set. Thus, once the projection directions are determined, each person can be modeled by a single image.

All of the algorithms performed better on the images of the full face. Note that there is a dramatic improvement in the Fisherface method where the error rate was reduced from 7.3% to 0.6%. When the method is trained on the entire face, the pixels corresponding to the occluding contour of the face are chosen as good features for discriminating between individuals. i.e., the overall shape of the face is a powerful feature in face identification. As a practical note however, it is expected that recognition rates would have been much lower for the full face images if the background or hairstyles had varied and may even have been worse than the closely cropped images.

4 Conclusion

The experiments suggest a number of conclusions:

1. All methods perform well if presented with an image in the test set which is similar to an image in the training set.
2. The Fisherface method appears to be the best at extrapolating and interpolating over variation in lighting, although the Linear Subspace method is a close second.
3. Removing the initial three principal components does improve the performance of the Eigenface method in the presence of lighting variation, but does not alleviate the problem.
4. In the limit as more principal components are used in the Eigenface method, performance approaches that of correlation. Similarly, when the first three principal components have been removed, performance improves as the dimensionality of the feature space is increased. Note however, that performance seems to level off at about 45 principal components. Sirovitch and Kirby found a similar point of diminishing returns when using Eigenfaces to represent face images [27].
5. The Fisherface method appears to be the best at simultaneously handling variation in lighting and expression. As expected, the Linear Subspace method suffers when confronted with variation in facial expression.

Even with this extensive experimentation, interesting questions remain: How well does the Fisherface method extend to large databases. Can variation in lighting conditions be accommodated if some of the individuals are only observed under one lighting condition? i.e., how can information about the class of faces be exploited?

Additionally, current face detection methods are likely to break down under extreme lighting conditions such as Subsets 4 and 5 in Fig. 1, and so new detection methods will be needed to support this algorithm. Finally, when shadowing dominates, performance degrades for all of the presented recognition methods, and techniques that either model or mask the shadowed regions may be needed. We are currently investigating models for representing the set of images of an object under *all possible* illumination conditions; details will appear in [1].

Acknowledgements We would like to thank Peter Hallinan for providing the Harvard Database, and Alan Yuille and David Mumford for many useful discussions.

References

1. P. Belhumeur and D. Kriegman. What is the set of images of an object under all possible lighting conditions? In *IEEE Proc. Conf. Computer Vision and Pattern Recognition*, 1996.
2. D. Beymer. Face recognition under varying pose. In *Proc. Conf. Computer Vision and Pattern Recognition*, pages 756–761, 1994.
3. R. Brunelli and T. Poggio. Face recognition: Features vs templates. *IEEE Trans. Pattern Anal. Mach. Intelligence*, 15(10):1042–1053, 1993.
4. R. Chellappa, C. Wilson, and S. Sirohey. Human and machine recognition of faces: A survey. *Proceedings of the IEEE*, 83(5):705–740, 1995.
5. Q. Chen, H. Wu, and M. Yachida. Face detection by fuzzy pattern matching. In *Int. Conf. on Computer Vision*, pages 591–596, 1995.

6. Y. Cheng, K. Liu, J. Yang, Y. Zhuang, and N. Gu. Human face recognition method based on the statistical model of small sample size. In *SPIE Proc.: Intelligent Robots and Computer Vision X: Algorithms and Techn.*, pages 85–95, 1991.
7. I. Craw, D. Tock, and A. Bennet. Finding face features. In *Proc. European Conf. on Computer Vision*, pages 92–96, 1992.
8. Y. Cui, D. Swets, and J. Weng. Learning-based hand sign recognition using SHOSLIF-M. In *Int. Conf. on Computer Vision*, pages 631–636, 1995.
9. R. Duda and P. Hart. *Pattern Classification and Scene Analysis.* Wiley, New York, 1973.
10. R. Fisher. The use of multiple measures in taxonomic problems. *Ann. Eugenics*, 7:179–188, 1936.
11. A. Gee and R. Cipolla. Determining the gaze of faces in images. *Image and Vision Computing*, 12:639–648, 1994.
12. J. Gilbert and W. Yang. A Real–Time Face Recognition System Using Custom VLSI Hardware. In *Proceedings of IEEE Workshop on Computer Architectures for Machine Perception*, pages 58–66, 1993.
13. 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.
14. P. Hallinan. *A Deformable Model for Face Recognition Under Arbitrary Lighting Conditions.* PhD thesis, Harvard University, 1995.
15. J. Hespanha, P. Belhumeur, and D. Kriegman. Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection. Center for Systems Science 9506, Yale University, PO Box 208267, New Haven, CT 06520, May 1995.
16. B. Horn. *Computer Vision.* MIT Press, Cambridge, Mass., 1986.
17. A. Lanitis, C. Taylor, and T. Cootes. A unified approach to coding and interpreting face images. In *Int. Conf. on Computer Vision*, pages 368–373, 1995.
18. T. Leung, M. Burl, and P. Perona. Finding faces in cluttered scenes using labeled random graph matching. In *Int. Conf. on Computer Vision*, pages 637–644, 1995.
19. K. Matsuno, C. Lee, S. Kimura, and S. Tsuji. Automatic recognition of human facial expressions. In *Int. Conf. on Computer Vision*, pages 352–359, 1995.
20. Moghaddam and Pentland. Probabilistic visual learning for object detection. In *Int. Conf. on Computer Vision*, pages 786–793, 1995.
21. 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.
22. H. Murase and S. Nayar. Visual learning and recognition of 3-D objects from appearence. *Int. J. Computer Vision*, 14(5–24), 1995.
23. A. Pentland, B. Moghaddam, and Starner. View-based and modular eigenspaces for face recognition. In *Proc. Conf. Computer Vision and Pattern Recognition*, pages 84–91, 1994.
24. A. Samal and P. Iyengar. Automatic recognition and analysis of human faces and facial expressions: A survey. *Pattern Recognition*, 25:65–77, 1992.
25. A. Shashua. *Geometry and Photometry in 3D Visual Recognition.* PhD thesis, MIT, 1992.
26. W. Silver. *Determining Shape and Reflectance Using Multiple Images.* PhD thesis, MIT, Cambridge, MA, 1980.
27. Sirovitch, L. and Kirby, M. Low-dimensional procedure for the characterization of human faces. *J. Optical Soc. of America A*, 2:519–524, 1987.
28. M. Turk and A. Pentland. Eigenfaces for recognition. *J. of Cognitive Neuroscience*, 3(1), 1991.
29. M. Turk and A. Pentland. Face recognition using eigenfaces. In *Proc. IEEE Conf. on Comp. Vision and Patt. Recog.*, pages 586–591, 1991.
30. R. Woodham. Analysing images of curved surfaces. *Artificial Intelligence*, 17:117–140, 1981.