Short Papers

Principal Angles Separate Subject Illumination Spaces in YDB and CMU-PIE

J. Ross Beveridge, *Member*, *IEEE*,
Bruce A. Draper, *Member*, *IEEE*,
Jen-Mei Chang, Michael Kirby,
Holger Kley, and
Chris Peterson

Abstract—The theory of illumination subspaces is well developed and has been tested extensively on the Yale Face Database B (YDB) and CMU-PIE (PIE) data sets. This paper shows that if face recognition under varying illumination is cast as a problem of matching sets of images to sets of images, then the minimal principal angle between subspaces is sufficient to perfectly separate matching pairs of image sets from nonmatching pairs of image sets sampled from YDB and PIE. This is true even for subspaces estimated from as few as six images and when one of the subspaces is estimated from as few as three images if the second subspace is estimated from a larger set (10 or more). This suggests that variation under illumination may be thought of as useful discriminating information rather than unwanted noise.

Index Terms—Face recognition, illumination subspaces, principal angle, set-to-set classification.

1 Introduction

The computer vision community has made great progresses in understanding how the appearance of objects in images changes with illumination. For instance, it is known that the set of images of a convex Lambertian surface acquired under varying illumination conditions forms a convex "illumination cone" [1]. Moreover, it has been shown both empirically and theoretically that this illumination cone is intrinsically low-dimensional [2], [3], [4], in the sense that illumination cones can be embedded in approximately 9- or 10-dimensional linear subspaces, which are sometimes called "illumination subspaces."

To help test these theories, the Yale Face Database B (YDB) [5] and CMU-PIE [6] data sets were created to collect images of subjects under varying illumination, under conditions that match as closely as possible the assumptions underlying the theory of illumination cones. These databases have therefore been used to test many face recognition approaches that either model or remove illumination variations. See [5], [7], and [8] for lists of examples.

This paper makes two empirical claims about the nature of the YDB and PIE data sets and, in particular, about the subspaces spanned by sets of frontal images of any single subject sampled from these data sets. Stated formally, the first claim is given as follows:

Claim 1. *Let $\mathcal{S}^i$ be the complete set of frontal images of subject i in either the YDB or PIE data set. Let $\mathbb{I}^i$ be the set of all linear subspaces spanned by k or more images in $\mathcal{S}^i$ and let $\mathcal{X}^i$ and $\mathcal{Y}^i$ be any two elements of $\mathbb{I}^i$. Finally, define $P(\mathcal{X},\mathcal{Y})$ to be the minimal principal angle between two subspaces $\mathcal{X}$ and $\mathcal{Y}$. Then, for $k=10$ in the YDB and $k=6$ in the PIE data set*

$$P(\mathcal{X}^i,\mathcal{Y}^i) < \min_{j\neq i}\{P(\mathcal{X}^i,\mathcal{Y}^j), P(\mathcal{Y}^i,\mathcal{X}^j)\}.$$

Informally, the minimum principal angle between any two subspaces estimated from a sufficient number of frontal images of a single subject under varying illumination will always be smaller than the minimum principal angles between subspaces estimated from images of different subjects. The connection between the theory of illumination cones and this claim is summarized in Section 3. The empirical evidence for this claim is presented in Sections 4 and 5 and, specifically, Figs. 2 and 3.

Our second claim considers the case where the sets of images being compared may not be of equal size. Stated formally, the second claim is given as follows:

Claim 2. *Let $\mathcal{S}^i$ and P be defined as above. Let $\mathcal{X}^{i,k}$ represent a subspace estimated from exactly k frontal images of subject i and let $\mathcal{Y}^{j,l}$ represent a subspace estimated from exactly l frontal images of subject j, both from PIE. Then,*

$$P(\mathcal{X}^{i,k},\mathcal{Y}^{i,l}) < \min_{j\neq i}\{P(\mathcal{X}^{i,k},\mathcal{Y}^{j,l}), P(\mathcal{Y}^{i,l},\mathcal{X}^{j,k})\},$$

for all $k\geq 3$ and $l\geq 10$.

The empirical evidence supporting this claim is presented in Section 6 and, specifically, Fig. 4.

There are at least three reasons why these claims are important. First, they offer solid measurable support for predictions arising from the theory of illumination subspaces. The importance of this empirical support should not be underestimated given the simplifying assumptions upon which illumination subspace theory rests and the degree to which these assumptions are violated by even well-controlled data collections such as YDB or PIE.

The second reason these claims are important is that the measure proposed for comparing linear subspaces derived from image samples, namely, the minimum principal angle, is about as close as we can come to a simple objective universal metric for comparing linear subspaces. In our other work [9], we have considered alternative measures of geodesic distance on the Grassmann manifold. However, this additional complexity is not necessary here, and our claims as stated above are readily understood simply in terms of the minimum principal angle: a concept we will both review and illustrate in Section 3.

The third reason these claims are important is that they represent a basic intrinsic property of the data sets themselves and not a claim about a particular algorithm. In other words, on these data sets, there is a simple comparative metric that determines whether two subspaces were estimated from the same subject or different subjects. As a result, any face recognition system that compares randomly selected sets of frontal images of one subject to sets of images of another should be expected to perform perfectly when using a sufficient number of frontal images, namely, 10 images for YDB and six images for PIE.

There is also a larger message implied by the image set separability of both the YDB and PIE imagery summarized in Claim 1 and Claim 2. While the common view is that variation in illumination is a confounding nuisance, under the correct

- *J.R. Beveridge and B.A. Draper are with the Computer Science Department-1873, Colorado State University, Fort Collins, CO 80523-1873. E-mail: {ross, draper}@cs.colostate.edu.*
- *J.-M. Chang is with the Department of Mathematics and Statistics, California State University, Long Beach, Long Beach, CA 90840. E-mail: jchang@csulb.edu.*
- *M. Kirby, H. Kley, and C. Peterson are with the Department of Mathematics, Colorado State University, Fort Collins, CO 80523. E-mail: {kirby, kley, peterson}@math.colostate.edu.*

Manuscript received 6 Aug. 2007; revised 25 Apr. 2008; accepted 15 July 2008; published online 29 July 2008.
Recommended for acceptance by A. Martinez.
For information on obtaining reprints of this article, please send e-mail to: tpami@computer.org, and reference IEEECS Log Number TPAMI-2007-08-0482.
Digital Object Identifier no. 10.1109/TPAMI.2008.200.

circumstances, the variation in appearance of an individual's face under different illumination conditions is highly discriminatory and potentially useful. Indeed, future face recognition systems might do well to acquire many images under intentionally varied lighting and then approach face recognition as a subspace-to-subspace, rather than image-to-image, matching problem.

2 Background

The YDB and PIE data sets have been widely used to study face recognition algorithms and to develop new face recognition techniques that will work under varying illumination. The most common approach is to try to remove the effects of illumination before matching, as in [7], [10], and [11]. Closer to this paper are methods that preserve illumination variations, although most of these define recognition as the task of matching either single images to single images or single images to a set of images, for example, [5], [12], and [8].

There are relatively fewer works on set-to-set image comparisons. Srivastava et al. [13] looks for an optimal linear subspace representation on the Grassmann manifold to represent a set of images. Similar to many other recognition schemes, this type of comparison requires a significant amount of training a priori before recognition can take place. Although perfect recognition after extensive training is good, it is even better to be able to use general subspace representations to achieve perfect recognition without extensive training.

At the heart of our findings is a well-known method for comparing two linear subspaces, namely, principal angles. Many others have used principal angles in the context of face recognition. For example, Yamaguchi et al. [14] used the minimal/first principal angle (or maximum correlation) between training and testing subspaces to capture the similarity between the two sets and called their method the *Mutual Subspace Method (MSM)*. Since then, the concept of canonical correlation has been widely used. For example, the central idea in the *Constraint Mutual Subspace Method (CMSM)* [15] is that by projecting the probe and gallery subspaces to a constrained subspace (generated by considering the difference subspaces), the new principal angle preserves the difference between people while excluding unnecessary components for recognition, namely, undesirable variations.

In the *Multiple Constrained Mutual Subspace Method (MCMSM)* [16], multiple constrained subspaces are created using methods of ensemble learning (bagging and boosting), and MSM is used to classify. The combined similarity between two subspaces is given by combining the similarities calculated on each constrained subspace. In [17] and [18], the authors proposed a new feature extraction method and a new feature fusion strategy based on the generalized canonical correlation analysis (GCCA). Kim et al. proposed in [19] a method that maximizes the canonical correlations of within-class sets and minimizes the canonical correlations of between-class sets that is inspired by *Linear Discriminative Analysis*.

Other extensions of Canonical Correlation Analysis include Kernel Principal Angles (KPAs) by Wolf and Shashua [20] and Incremental Kernel SVD by Chin et al. [21]. Kim et al. [22] addressed one of the major shortcomings of MSM-based methods: ad hoc fusion of information contained in different principal angles. They proposed using principal angles to build simple weak classifiers, which are then combined using the AdaBoost algorithm [23].

One major distinction between the prior works and the proposed method is the requirement of training. This could be subspace training [15], [16], [24], feature extraction [18], [19], or classifier training [22]. In contrast, face recognition algorithms based on principal angles may not require training, and Claims 1 and 2 suggest that they may obtain excellent results when lighting variations are observable.

Our findings strongly suggest that the minimum principal angle is useful in the context of face recognition, and it is reasonable to ask about the maximum principal angle. The maximum or largest principal angle is related to the notion of distance between equidimensonal subspaces that are defined based on the orthogonal projections onto the subspaces [25], [26]. While the largest principal angle is useful in a lot of statistical applications, our experience suggests that it conveys little useful information for set-to-set face recognition.

3 Illumination Subspaces and Principal Angles

The theory of illumination cones and the illumination subspaces they lie in is well developed. It assumes that the geometry of the scene is invariant and that the surfaces are convex and Lambertian. Care was taken in the construction of both the YDB and PIE data sets to try to match, as much as possible, these assumptions. Thus, empirical evidence from these data sets suggests that the theory is a good predictor of how empirically estimated illumination subspaces behave.

The theory predicts that for any given person and pose, a collection of images drawn from a wide variety of different lighting conditions are well approximated by a 9- or 10-dimensional linear subspace. This prediction, in turn, suggests that the linear subspace spanned by an appropriately distributed and modest-sized set of sample images taken under varying illumination should span a reasonable portion of the entire illumination subspace. Moreover, if for a given person's face, the process is repeated using two sets of, say, 10 sample images each, then the two linear subspaces spanned by these sample images should approximate the same illumination subspace and should therefore be similar. It is through this connection that the empirical results below, which compare the similarity of linear subspaces derived from images under varying illumination, reveal just how well the implications of the theory are manifested in face images from the YDB and PIE data sets.

3.1 Measuring Principal Angles

To find the minimum principal angle between pairs of subspaces that are basis independent, we use a robust method developed by Björck and Golub [27]. Formally, let X and Y be collections of k and l images:

$$X = [x_1, x_2, \ldots, x_k] \quad \text{and} \quad Y = [y_1, y_2, \ldots, y_l], \tag{1}$$

where column vectors x_i and y_i are understood to contain n pixels from a face image that has been geometrically normalized to place the eyes at a fixed location. In addition, pixels outside the face oval have been cropped.

It is also understood that all the images in X are of the same person, and all the images in Y are of the same person. Sets X and Y may contain images of the same or different people. The two image sets X and Y span two vector subspaces $\mathcal{X}$ and $\mathcal{Y}$ in $\mathbb{R}^n$, respectively.

Given that the YDB and PIE data set images sample variations in illumination, as k grows beyond approximately 10, much of the total energy/variation in $\mathcal{X}$ and $\mathcal{Y}$ is captured by the first 10 dimensions. However, there are other sources of variation, and we have never observed an $\mathcal{X}$ or $\mathcal{Y}$ that is not full rank. Consequently, the dimensionality of these subspaces is equal to the number of sample face images, specifically k for $\mathcal{X}$ and l for $\mathcal{Y}$.

The *minimal principal angle*, θ, between the vector subspaces $\mathcal{X}$ and $\mathcal{Y}$ is then defined as follows:

$$\cos(\theta) = \max_{\substack{u \in \mathcal{X} \\ \|u\|_2 = 1}} \max_{\substack{v \in \mathcal{Y} \\ \|v\|_2 = 1}} u^T v. \tag{2}$$

While our claims are only concerned with the minimum/first principal angle, an additional principal angle is defined by searching in the complement of spaces spanned by u and v. In

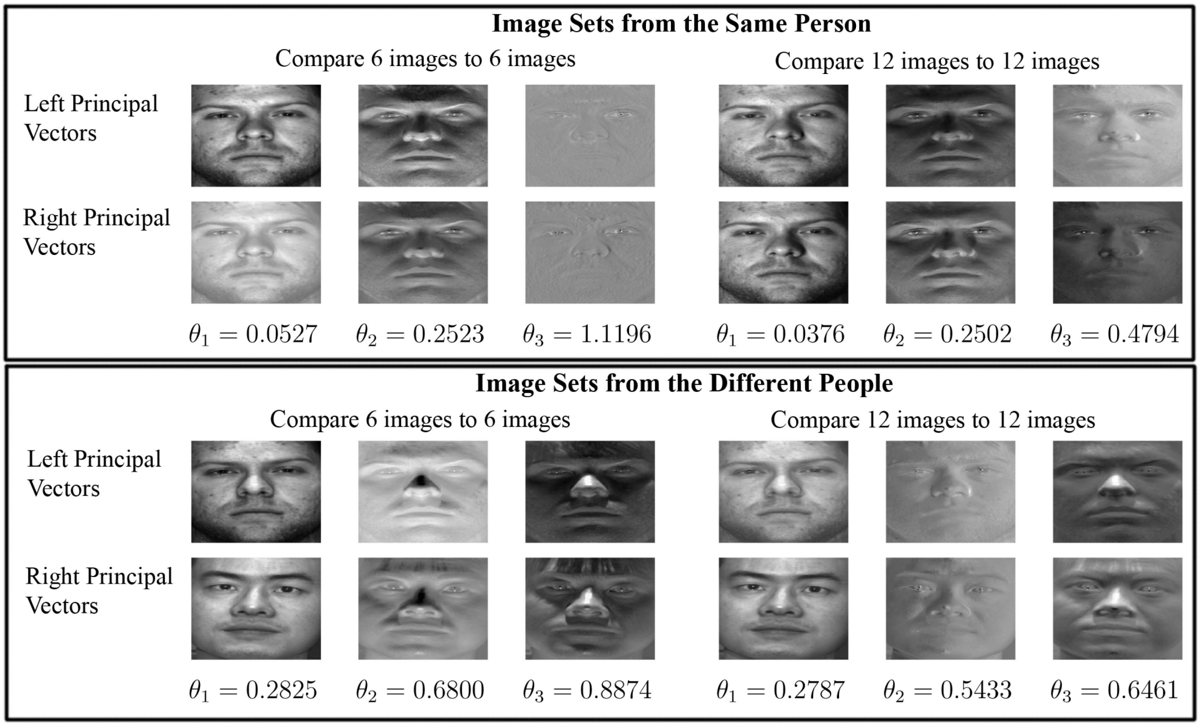

Fig. 1. The first three left and right principal vectors and associated principal angles for comparisons from the YDB. The top comparisons are for sets of images of the same person, while the bottom comparisons are for sets of images of different people. Note that the values for the minimum/first principal angles, θ_1, are notably lower on the top, and the principal vectors are more visually similar. The comparisons are for sets containing 6 (left) and 12 (right) images.

recursive fashion, an entire sequence of principal angles may thus be defined.

The minimum principal angle between two sets of images has an intuitive interpretation. It is essentially a measure of correlation between two images, each of which is defined as some linear combinations of the original images in each of the two sets. Thus, the minimum principal angle may be thought of as answering the question: "What linear combination of images in one set comes closest to a linear combination of images in a second set?"

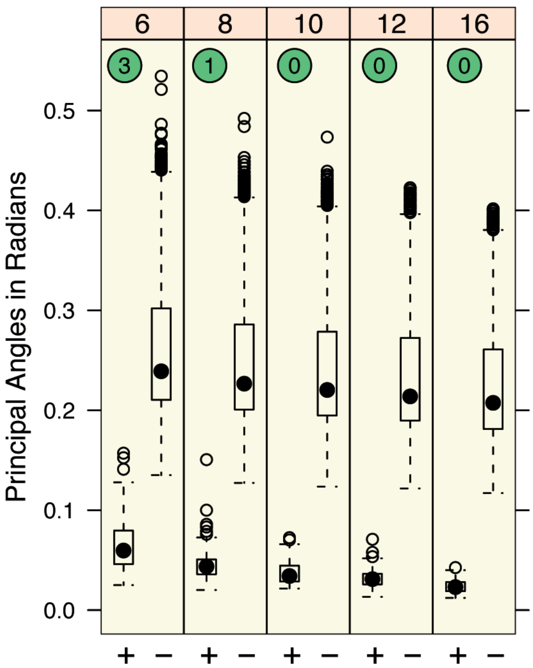

Fig. 2. Box whisker plots of the minimal principal angles of the matching $(+)$ and nonmatching $(-)$ subspaces where image sets are of size 6 to 16 from left to right, accordingly, selected from YDB. Numbers of match angles that are greater than or equal to the smallest nonmatch angles are shown along the top of the distributions. Notice that perfect separation of the matching and nonmatching subspaces is observed when the image set size is greater than or equal to 10.

If one names the sets "left" and "right," one can then describe the closest pair of linear combinations of images as the left and right principal vectors. Fig. 1 shows principal vectors and angles for four image set comparisons from the YDB. Specifically, the first three principal vectors and angles are shown for four distinct comparisons, two involving different sets of images of the same person (upper row) and two involving different sets of images of different people (bottom row). Vectors and angles for sets containing 6 and 12 images are shown. These examples are drawn from the YDB study presented below.

One thing to note in Fig. 1 is that the minimum/first principal angle between sets of images of the same person are not zero. Given that the theory of illumination subspaces leads us to expect a single low-dimensional subspace for a face, given all the appropriate caveats, one might be excused for at first expecting explicit overlap in the illumination subspaces and, consequently, a zero angle. Such an overlap would also manifest itself in the left and right principal vectors being identical. In practice, this never happens, or at least, we have never in our studies encountered a zero minimum principal angle.

There are myriad reasons why sampled illumination spaces never truly overlap, but one particularly easy one to explain is the simple presence of independent sensor noise at each pixel. There are of course other more important factors at play, including, for example, the fact that faces are not rigid, convex, or perfectly Lambertian.

While sampled illumination spaces do not overlap, as the example in Fig. 1 clearly demonstrates, the minimum principal angle for comparisons between image sets of the same person are notably smaller than for comparisons involving different people. The role played by k, the number of images in the set, is also evident in the minimum/first principal vectors and associated minimum

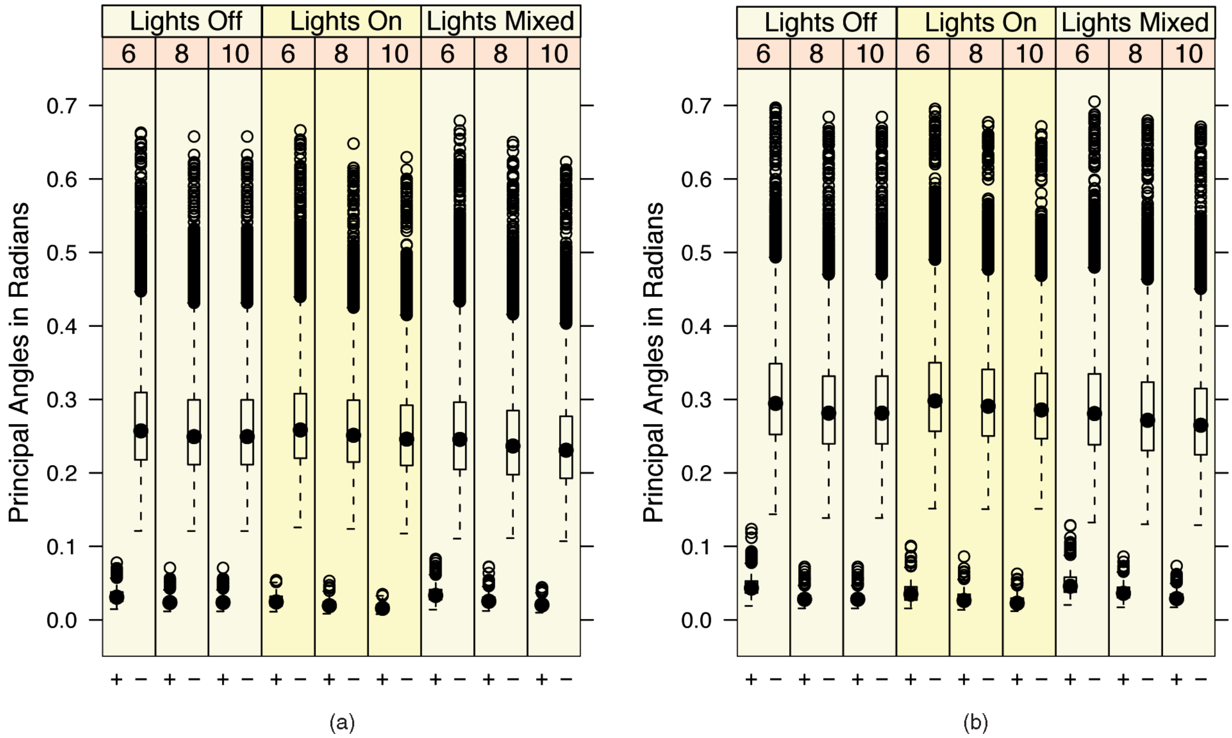

Fig. 3. Box whisker plots of the minimal principal angles of the matching $(+)$ and nonmatching $(-)$ subspaces where image sets are of size 6 to 10 from left to right, accordingly, selected from CMU-PIE Database. (a) Plots when mirrored images are included. (b) The results when mirrored images are not included.

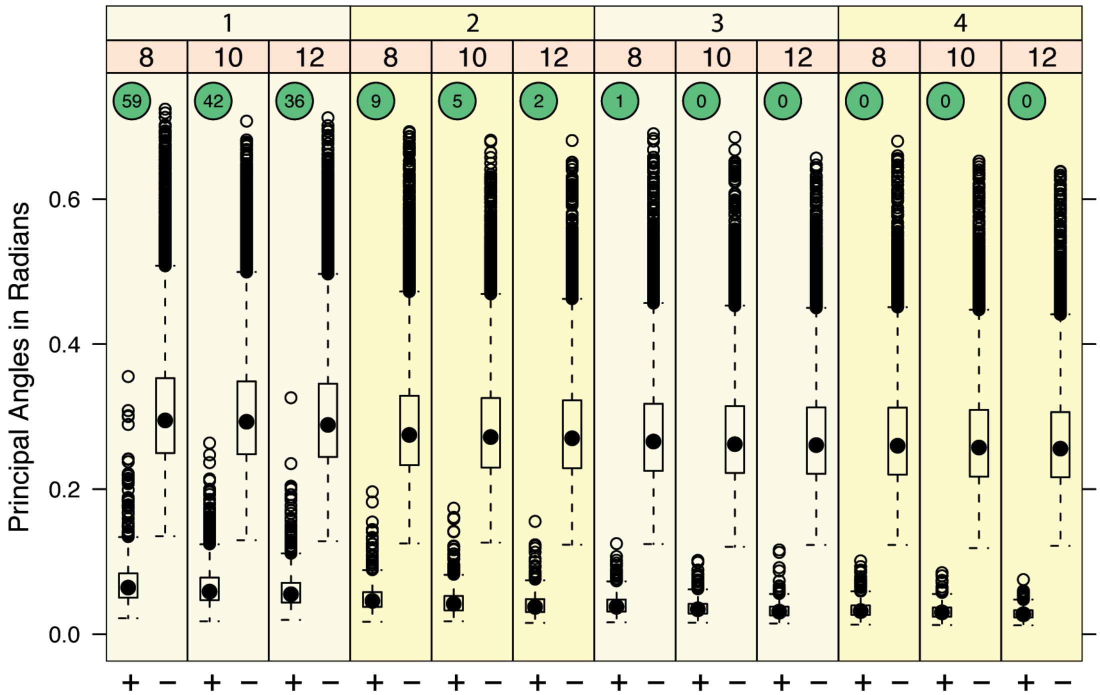

Fig. 4. Box whisker plots of the minimal principal angles of the matching $(+)$ and nonmatching $(-)$ subspaces where the smaller sets contain one through four images and the larger sets contain 8, 10, and 12 images for all i selected from CMU-PIE images with room lights off. As in Fig. 2, the number of match angles greater than or equal to the smallest nonmatch angle is shown along the top of the distributions.

principal angles when the sets are for the same person (top row of Fig. 1). The shift from 6 to 12 images drops the minimum principal angle from 0.0527 to 0.0376, a nearly 30 percent reduction.

In general, questions of closeness of linear subspaces lead into the broader question of how to compute distances between subspaces; metrics for calculating distances between subspaces can be found in [28]. Further, it can be shown that all unitarily invariant metrics for subspaces are functions of principal angles between subspaces [26], and hence, the robustness of the method can be explored through the study of matrix perturbation theory. A detailed discussion of this approach as it relates to face imagery can be found in [29], and as already stated in our introduction, it is not necessary for us to pursue these greater questions in order to convey the essential results of this paper.

3.2 Augmenting the Samples with Mirror-Reflected Images

Before continuing to the empirical results for the YDB and PIE data sets, there is an important additional detail associated with estimating illumination subspaces. Kirby and Sirovich [30] first introduced the use of symmetry-augmented image sets including both an original face image and an additional face image obtained by reflecting the image about the vertical centerline of the face. This operation has consistently shown itself to be useful when estimating linear subspaces from face imagery, and it will be used here. Formally, the image sets defined in (1) are augmented here:

$$X = \left[x_1, x_1^m, x_2, x_2^m, \ldots, x_k, x_k^m\right] \text{ and } \\ Y = \left[y_1, y_1^m, y_2, y_2^m, \ldots, y_k, y_l^m\right], \quad (3)$$

where column vectors x^m and y^m are the mirror-reflected versions of face images x and y. The dimensionality of the subspaces $\mathcal{X}$ and $\mathcal{Y}$ associated with X and Y rises to $2k$ and $2l$, respectively, with the addition of mirror-reflected images. Due to the symmetry of the face, one may think of this step as approximating how a face might appear under a greater variety of lighting conditions. It is clear from the results below that this step improves the separation between subspaces associated with the same person and subspaces associated with different people.

To simplify the visual appearance and interpretation of the principal vectors and principal angles shown in Fig. 1, they were determined not using mirror-reflected images. However, henceforth, unless otherwise stated, it should be assumed that image sets are augmented with mirror-reflected images as defined in (3).

4 The Yale Face Database B (YDB)

Although the YDB contains imagery for only 10 people, it is the oldest and most studied illumination database with images spanning a large range of possible illuminations. The box whisker plots in Fig. 2 show the range of minimum principal angle values for matching and nonmatching pairs of image sets sampled from YDB. These are standard box whisker plots from the R statistics package. The top and bottom lines of the box in the box whisker plot represent the upper and lowest quartile values, whereas the circle in the box represents the median. The whiskers attempt to extend above and below the box to a distance 1.5 times the interquartile range. However, they must terminate on a data point, so the whiskers "shrink back" until they land on a sample.

When the largest angle[1] between any two matching subspaces is less than the smallest angle between any two nonmatching subspaces, the data is perfectly separable by the minimal principal angle. The comparison is made for sets of 6, 8, 10, 12, and 16 Yale images acquired under frontal pose and 64 possible illumination conditions. The set sizes are shown along the top of the plots.

To generate this plot, two disjoint image sets of size k are randomly selected for each of the 10 Yale subjects. For a random selection of sets, there are 10 matches where the angle is measured between sets associated with the same person. There are 90 angles associated with different people. In the plot, a "$+$" is used to indicate the match distribution, and "$-$" indicates the nonmatch distribution. This process of randomly selecting disjoint sets is then repeated 10 times. Thus, there are a total of 100 match angles and 900 nonmatch angles in the distribution shown in Fig. 2. This random sampling approach is used to test Claim 1 because exhaustive sampling is infeasible.

At the top of each pair of distributions, the actual number of match angles greater than or equal to the smallest nonmatch angle is shown. For example, for sets of size 8, there is one match angle that is greater than or equal to the smallest nonmatch angle; the distributions are almost but not completely disjoint. For sets of size 10, there is no match angle larger than the smallest nonmatch angle; the sample distributions are disjoint. Indeed, for sets of size greater than or equal to 10 in YDB, perfect separation between the matching and nonmatching subspaces by the minimal principal angle is observed.

The distributions shown in Fig. 2 use the mirror image augmentation defined in Section 3.2. Therefore, for example, set size 10 means that the matrix of image samples X defined in (3) includes 20 columns. For the sake of comparison, were the mirror images not included, then a set size of 10 would imply an X with only 10 columns as defined by (1). When mirror images are not included the number of match angles larger than or equal to the smallest nonmatch angle for the cases shown in Fig. 2 would be 40, 7, 1, 1, and 0, respectively. Hence, pure separability does not arise until $k = 16$ when mirrored images are excluded.

5 The CMU-PIE Database

There are 67 subjects but only 42 illumination variants in the CMU-PIE database. In addition, 21 illumination variants are sampled with the background (ambient) room lights on, and 21 are sampled with the room lights off. Fig. 3 shows box whisker plots of principal angle distribution for matching and nonmatching pairs of image sets sampled from PIE for the cases where the room lights are off, the room lights are on, and there is a pool of both.

The distributions shown are obtained in a manner similar to that used to generate Fig. 2. Since there are only 21 images available for each of the room lights on and room lights off cases, the disjoint sets of equal size cannot be larger than 10. Fig. 3a shows the distributions with mirror images included, and Fig. 3b shows distributions where mirror images are not included.

As before, random disjoint image sets are selected independently 10 times. Thus, the distribution of match angles is based upon 670 samples, and the distribution of nonmatch angles is based upon 44,220 nonmatch angles. Unlike the YDB results in Fig. 2, here, in all cases, there is complete separation of match versus nonmatch angles; thus, counts of overlaps between the distributions are not included. It is true that the case of mixed lighting and $k = 6$ in Fig. 3b comes very close to having an overlap, and generally, the distributions are better separated in Fig. 3a than in Fig. 3b.

6 Asymmetric Comparisons with Few Image Samples

Up to this point, illumination spaces were empirically estimated based upon an equal number of samples in each set. This need not be the case. Fig. 4 shows results for asymmetric set-to-set comparisons where the sizes of sets are unequal. In particular, it shows comparisons between small sets of between one and four images to larger sets of size 8, 10, or 12 images. As before, disjoint random sets are selected independently 10 times. Since our previous results established the value of including mirror images as described in Section 3.2, only results using mirror images are shown in Fig. 4.

The numbers of match angles exceeding the smallest nonmatch angles can be seen to decrease monotonically reading from left to right. Perfect separation is first achieved when a subspace based upon three images is compared to a subspace based upon 10 images. These results are for the case of more extreme lighting changes associated with the room lights off. Separation between distributions was even better for the room-lights-on case, with overlaps of only seven, four, and five for the comparison of one image to 8, 10, and 12, respectively. Room-lights-on comparisons of two or more images yielded zero overlap.

These exact cutoffs are probably artifacts of the PIE data set, but scanning across Fig. 4, the trend is clear: large separations in

1. Where there is little risk of confusion, "minimum principal angle" will be shortened to "angle." Further, "match angle" is short for "the minimum principal angle between two subspaces derived from images of the same person" and "nonmatch angle" is short for "the minimum principal angle between two subspaces derived from images of different people."

principal angle are observed between matching and nonmatching subspaces, even when one of the subspaces is estimated from a relatively small number of samples. This may be important in operational face recognition scenarios where larger numbers of images can be collected during enrollment than operation.

It is useful to relate this result to illumination subspace theory. If the images satisfied the assumptions perfectly and the larger set spanned the face's illumination subspace, then we would expect the minimum principal angle to be zero, even if the smaller set contained only a single image. This is not what we see in Fig. 4. Instead, the minimum principal angles between single PIE images[2] and sets of PIE images of the same people can be quite large, ranging up to 0.35 radians when the larger set contains eight images—the leftmost distribution in Fig. 4. Earlier, we discussed why in practice minimum principal angles are never zero. The observation that minimum principal angles measured to a single image can be large further reinforces our understanding that in real data such as PIE, there are factors not modeled by illumination theory. Fortunately, as the size of the smaller set grows, the match angles drop quickly. The overall point is that while it is not necessary to have as many sample images as the expected dimensionality of the illumination space, there is considerable value in using two or three images rather than just one.

7 Conclusion

There is, we think, a growing consensus that frontal imagery from the YDB and PIE data sets is in some sense easy. Certainly, a variety of good algorithms achieve very high levels of performance on this data. The contribution of this work is that it helps explain why this is true. Illumination spaces associated with individual people in these data sets are very distinctive, and by Claim 1 above, 10 images are sufficient to perfectly disambiguate individuals. Further, as indicated by Claim 2, similar separability is observed with fewer samples when those fewer samples are compared to a larger set of samples. By implication, any algorithm that taps into this information, either through training or directly through sampling, ought to do very well on these data sets.

Stepping beyond just the YDB and PIE data sets, our results strongly support the view that changes in illumination are informative and may provide a basis for highly accurate recognition systems. The orientation of an illumination cone, after all, is a function of both the 3D shape and the reflectance properties of a face. It makes sense that it should be discriminating, and this intuition is supported by the data presented here. More definitive support for this claim, of course, will require tests on larger data sets, and we are in the early stages of collecting such a set.

Acknowledgments

This study was partially supported by the US National Science Foundation under Award DMS-0434351 and the US DoD-USAF-Office of Scientific Research under Contract FA9550-04-1-0094. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the US National Science Foundation or the US DoD-USAF-Office of Scientific Research.

References

[1] P. Belhumeur and D. Kriegman, "What Is the Set of Images of an Object under All Possible Illumination Conditions," *Int'l J. Computer Vision,* vol. 28, no. 3, pp. 245-260, July 1998.
[2] K.-C. Lee, J. Ho, and D. Kriegman, "Acquiring Linear Subspaces for Face Recognition under Variable Lighting," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 27, no. 5, pp. 684-698, May 2005.
[3] A. Georghiades, D. Kriegman, and P. Belhumeur, "Illumination Cones for Recognition under Variable Lighting: Faces," *Proc. IEEE Conf. Computer Vision and Pattern Recognition,* pp. 52-58, 1998.
[4] R. Basri and D. Jacobs, "Lambertian Reflectance and Linear Subspaces," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 25, no. 2, pp. 218-233, Feb. 2003.
[5] A. Georghiades, P. Belhumeur, and D. Kriegman, "From Few to Many: Illumination Cone Models for Face Recognition under Variable Lighting and Pose," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 23, no. 6, pp. 643-660, June 2001.
[6] T. Sim, S. Baker, and M. Bsat, "The CMU Pose, Illumination, and Expression Database," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 25, no. 12, pp. 1615-1618, Dec. 2003.
[7] T. Riklin-Raviv and A. Shashua, "The Quotient Image: Class Based Re-Rendering and Recognition with Varying Illuminations," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 23, no. 2, pp. 129-139, Feb. 2001.
[8] R. Gross, I. Matthews, and S. Baker, "Appearance-Based Face Recognition and Light-Fields," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 26, no. 4, pp. 449-465, Apr. 2004.
[9] J.-M. Chang, J. Beveridge, B. Draper, M. Kirby, H. Kley, and C. Peterson, "Illumination Face Spaces are Idiosyncratic," *Proc. Int'l Conf. Image Processing and Computer Vision,* vol. 2, pp. 390-396, June 2006.
[10] Y. Weiss, "Deriving Intrinsic Images from Image Sequences," *Proc. Eighth Int'l Conf. Computer Vision,* vol. 2, pp. 68-75, July 2001.
[11] G. Finlayson, M. Drew, and C. Lu, "Intrinsic Images by Entropy Minimization," *Proc. Eighth European Conf. Computer Vision,* pp. 582-595, May 2004.
[12] J. Ho, M.-H. Yang, J. Lim, K.-C. Lee, and D. Kriegman, "Clustering Appearances of Objects under Varying Illumination Conditions," *Proc. IEEE Conf. Computer Vision and Pattern Recognition,* vol. 1, pp. 11-18, June 2003.
[13] X. Lui, A. Srivastava, and K. Gallivan, "Optimal Linear Representations of Images for Object Recognition," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 26, pp. 662-666, 2004.
[14] O. Yamaguchi, K. Fukui, and K. Maeda, "Face Recognition Using Temporal Image Sequence," *Proc. Third Int'l Conf. Automatic Face and Gesture Recognition,* pp. 318-323, 1998.
[15] K. Fukui and O. Yamaguchi, "Face Recognition Using Multi-Viewpoint Patterns for Robot Vision," *Proc. 11th Int'l Symp. Robotics Research,* pp. 192-201, 2003.
[16] M. Nishiyama, O. Yamaguchi, and K. Fukui, "Face Recognition with the Multiple Constrained Mutual Subspace Method," *Proc. Fifth Int'l Conf. Audio- and Video-Based Biometric Person Authentication,* pp. 71-80, 2005.
[17] Q.-S. Sun, Z.-D. Liu, P.-A. Heng, and D.-S. Xia, "A Theorem on the Generalized Canonical Projective Vectors," *Pattern Recognition,* vol. 38, no. 3, pp. 449-452, Mar. 2005.
[18] Q.-S. Sun, P.-A. Heng, Z. Jin, and D.-S. Xia, "Face Recognition Based on Generalized Canonical Correlation Analysis," *Proc. Int'l Conf. Intelligent Computing,* pp. 958-967, 2005.
[19] T.-K. Kim, J. Kittler, and R. Cipolla, "Discriminative Learning and Recognition of Image Set Classes Using Canonical Correlations," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 29, no. 6, pp. 1005-1018, June 2007.
[20] L. Wolf and A. Shashua, "Learning over Sets Using Kernel Principal Angles," *J. Machine Learning Research,* vol. 4, no. 10, pp. 913-931, 2003.
[21] T.-J. Chin, K. Schindler, and D. Suter, "Incremental Kernel SVD for Face Recognition with Image Sets," *Proc. Seventh Int'l Conf. Automatic Face and Gesture Recognition,* pp. 461-466, Apr. 2006.
[22] T.-K. Kim, O. Arandjelović, and R. Cipolla, "Boosted Manifold Principal Angles for Image Set-Based Recognition," *Pattern Recognition,* vol. 40, pp. 2475-2484, 2007.
[23] Y. Freund and R.E. Schapire, "A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting," *Proc. Second European Conf. Computational Learning Theory,* pp. 23-37, 1995.
[24] M. Nishiyama, M. Yuasa, T. Shibata, T. Wakasugi, T. Kawahara, and O. Yamaguchi, "Recognizing Faces of Moving People by Hierarchical Image-Set Matching," *Proc. IEEE Conf. Computer Vision and Pattern Recognition,* pp. 1-8, 2007.
[25] G. Golub and C.V. Loan, *Matrix Computations,* third ed. Johns Hopkins Univ. Press, 1996.
[26] G. Stewart and J.-G. Sun, *Matrix Perturbation Theory.* Academic Press, 1990.
[27] A. Björck and G. Golub, "Numerical Methods for Computing Angles between Linear Subspaces," *Math. Computation,* vol. 27, no. 123, pp. 579-594, 1973.
[28] A. Edelman, T. Arias, and S. Smith, "The Geometry of Algorithms with Orthogonality Constraints," *SIAM J. Matrix Analysis and Applications,* vol. 20, no. 2, pp. 303-353, 1999.
[29] J.-M. Chang, "Classification on the Grassmannians: Theory and Applications," PhD dissertation, Dept. of Math., Colorado State Univ., Apr. 2008.
[30] M. Kirby and L. Sirovich, "Application of the Karhunen-Loève Procedure for the Characterization of Human Faces," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 12, no. 1, pp. 103-108, Jan. 1990.

2. Recall that these results use mirror-reflected images, so results for sets containing a single original PIE image have two columns in X and the subspace $\mathcal{X}$ spans two dimensions. This fact does not alter the underlying thread of our argument since the small angle predicted by the theory for a single image without the additional mirror images certainly holds true for the case including mirror images.