

International Journal of Pattern Recognition and Artificial Intelligence  
© World Scientific Publishing Company

## ROBUST ADAPTIVE PRINCIPAL COMPONENT ANALYSIS FOR FACE RECOGNITION

SHAOKANG CHEN,\* TING SHAN,<sup>†</sup> and BRIAN C. LOVELL<sup>‡</sup>

*NICTA*

*Level 20, 300 Adelaide Street, Brisbane, QLD 4000, Australia*  
*School of Information Technology and Electrical Engineering*

*The University of Queensland*  
*Brisbane, QLD 4072, Australia*

*Shaokang.Chen@nicta.com.au\**

*Ting.Shan@nicta.com.au<sup>†</sup>*

*lovell@itee.uq.edu.au<sup>‡</sup>*

*[http://www.itee.uq.edu.au/~shaokang\\*](http://www.itee.uq.edu.au/~shaokang*)*

*<http://www.itee.uq.edu.au/~shanting<sup>†</sup>>*

*<http://www.itee.uq.edu.au/~lovell<sup>‡</sup>>*

Recognizing faces with uncontrolled pose, illumination, and expression is a challenging task due to the fact that features insensitive to one variation may be highly sensitive to the other variations. Existing techniques dealing with just one of these variations are very often unable to cope with the other variations. The problem is even more difficult in applications where only one gallery image per person is available. In this paper, we describe a recognition method, Adaptive Principal Component Analysis (APCA), that can simultaneously deal with large variations in both illumination and facial expression using only a single gallery image per person. We have now extended this method to handle head pose variations in two steps. The first step is to apply an Active Appearance Model (AAM) to the non-frontal face image to construct a synthesized frontal face image. The second is to use APCA for classification robust to lighting and pose. The proposed technique is evaluated on three public face databases — Asian Face, Yale Face, and FERET Database — with images under different lighting conditions, facial expressions, and head poses. Experimental results show that our method performs much better than baseline recognition methods including PCA, FLD and PRM. More specifically, we show that by using AAM for frontal face synthesis from high pose angle faces, the recognition rate of our APCA method increases by up to a factor of 4.

*Keywords:* Face Recognition; Pose, Illumination, and Expression; Face subspace; Space Rotation.

### 1. Introduction

Uncontrolled face recognition is an extremely challenging task because of the overall similarity of all human faces accompanied by large differences between face images of the same person due to image capture variations such as changes in lighting, view point, head pose, and facial expression. An ideal face recognition system should recognize faces from face images acquired under quite natural uncontrolled

photographic conditions. However, the fact that differences between images of the same face due to these nuisance variations are normally much greater than those between faces<sup>1</sup> of different people makes it extremely difficult to compensate for common image capture variations. Most current face recognition systems can only deal with images taken under very constrained image capture conditions, thereby increasing implementation cost and also limiting the usefulness of such systems.

Some researchers refer to pose, illumination, and expression as PIE variations and have directed their recent research to diminishing the impact of PIE. Two main approaches have been proposed for illumination invariant recognition. One is to represent images with features that are less sensitive to illumination changes<sup>16,38,40</sup> such as the edge maps of the image. However, there are also disadvantages with edge representations because a shift in edge locations resulting from a small rotation or location error will degrade recognition performance. Yilmaz and Gokmen proposed to use hills for face representation which overcome the locality problem by spreading the edge profile.<sup>40</sup> Others prefer to use derivatives of gray-level distributions of face images<sup>3,14</sup> because hills may carry an artificial “edginess.” No matter what kind of representation is used, these approaches assume that features do not change dramatically with variable lighting conditions. But this is not always the case. All of the above representations suffer from the fact that features generated from shadows are related to illumination changes and may have an impact on recognition. Experiments done by Adinij et al.<sup>1</sup> show that even with the best image representations using illumination insensitive features and the best distance measurement, the misclassification rate is more than 20%.

Another approach is to construct a low dimensional linear subspace for images of faces taken under different lighting conditions. This approach is based on an assumption that images of a convex Lambertian object under variable illuminations form a convex cone in the space of all possible images.<sup>3</sup> If we ignore shadows, this subspace has three dimensions<sup>41</sup>. To account for attached shadows, higher dimensionality should be involved and around 5 to 9 images are required to construct the convex cone.<sup>2,18,22</sup> All these methods suppose that the surface of human faces is Lambertian reflected and convex. Therefore, it is hard for these systems to deal with cast shadows. Furthermore, these systems need several images of the same face taken under different controlled lighting source directions to construct a model of a given face. However, sometimes it is difficult or impossible to obtain a set of different images of a given face under specific conditions (e.g., this is certainly true in the case of historical photographs of deceased persons).

In the case of expression invariant recognition, this is problematic for machine recognition and is quite a difficult task for humans. One approach<sup>5,8</sup> is to morph images to be the same neutral expression as the gallery image. Yet, it is not guaranteed that all images can be morphed correctly, for example an image with closed eyes cannot be morphed to a neutral image because there is no texture available for the eyes. In another approach, Xiaoming Liu et. al<sup>27</sup> proposed to use optical flow for face recognition with facial expression variations. But it is hard to learn the

local motions within the feature space to determine the expression change of each face, since the way one person expresses a certain emotion is normally somewhat different from another. Martinez<sup>28</sup> proposed a weighting method to deal with facial expressions. A face image is divided into several local areas and those that are less sensitive to expression change are chosen and weighted accordingly. But features that are insensitive to expression changes may be sensitive to illumination change. Adin et al<sup>1</sup> says that “when a given representation is sufficient to overcome a single image variation, it may still be affected by other processing stages that control other imaging parameters.”

Pose invariant face recognition can be classified into two categories: 2D based approaches and 3D based approaches. Though 3D face models can be used to describe the human face appearance under various poses accurately, there are several disadvantages that limit its application.<sup>9</sup> First, to construct 3D face models, expensive 3D scanners have to be installed. Second, to acquire 3D data, the depth of field of scanners has to be well controlled and the range of data acquisition is limited. Third, using 3D scanners to obtain 3D data is time consuming and is not suitable for real-time applications. Experiments on FRGC 2006 show that face recognition on 2D still images achieves comparable performance to that of 3D approaches.<sup>32</sup> Thus, we choose to focus mainly on 2D methods.

Existing 2D pose invariant face recognition techniques can be divided into three classes: Pose-Invariant features, Multi-view spaces, and face synthesis approaches. Wiskott et. al proposed Elastic Bunch Graph Matching, which applied Gabor filter to extract pose invariant features.<sup>39</sup> In Ref. 6 and Ref. 7, Beymer used multiple-view templates to represent faces with different poses. Pentland et. al applied PCA for face recognition with pose variations by constructing view-based eigen spaces, where each view is represented by an individual space.<sup>30</sup> A multi-view approach based on Modular PCA is proposed by Sankaran and Asari to improve the performance.<sup>34</sup> Multiple view approaches requires several images per person under controlled view conditions to identify a face, which restrict its application when only one image available per person. Thus, synthesized methods emerged. Beymer et. al. extended their earlier work of view-based approach to example-based by applying the rotation seen in the prototypes to construct other virtual view templates from a single view.<sup>5</sup> In Ref. 15, Gao et. al constructed Face-Specific Subspace by synthesizing novel views of a single image. A simplified 3D pose recovery model was developed in 2001 to convert a rotated view to a frontal view, which is used for recognition.<sup>17</sup> In 2000, Cootes et. al proposed Active Appearance Models<sup>13</sup> to learn the relationship between model parameters and head orientation and to synthesize face images at different views. Sanderson et. al addressed the pose mismatch problem by extending each frontal face model with artificially synthesized models for non-frontal views.<sup>33</sup>

The above approaches may be able to deal with a certain kind of face variation well but there have constraints restricting their application when multiple variations involved and only one image available per subject. It is a risk for those approaches

that rely heavily on choosing invariant features,<sup>3,14,16,28,38,39,40</sup> because features insensitive to one variation may be highly sensitive to other variations and it is very difficult to abstract features that are completely immune to all kinds of variations. Thus their performance may degrade seriously in the case of multiple variations. Those approaches based on specific face models<sup>2,3,15,18,22,41</sup> will require multiple images per person taken under controlled conditions to construct the subspace for each person for the face representation. Other approaches apply multiple specific interval variation models<sup>6,7,30,34</sup> to divide the range of variation into several intervals and use the corresponding variation model to describe the face appearance change that lies in this interval. These approaches fuse several images representing different variations per person into the corresponding variation models so that matching can be done in each interval individually. However, acquiring the multiple images per person under specific conditions is often impractical and sometimes impossible.

In this paper, we present a method in an attempt to overcome the drawbacks of the above approaches called Adaptive Principal Component Analysis (APCA). We first apply Principle Component Analysis (PCA) for feature abstraction. Then we rotate and warp the space by whitening according to overall covariance and filtering the eigen features according to between-class and within-class covariance. Experiments show that our method performs considerably better than benchmark PCA<sup>37</sup> and Fisher Linear Discriminant(FLD)<sup>4</sup> on face recognition against both illumination and expression variations. As the method inherits benefits from both PCA and FLD methods, we suggest that APCA features could be called “eigen-fisherfaces.”

In the following, we first review PCA and FLD techniques and discuss out the reasons for their limitations in Section 2. We also briefly review the Active Appearance Model in Section 2. Then in Section 3, we describe and analyze our proposed method for variations on lighting conditions. Section 4 modify the method to deal with variations on facial expressions. We then extend our approach for pose invariant face recognition in Section 5. Finally, we will draw conclusions and discuss future extension in Section 6.

## 2. Related Work

We first review three image processing techniques that are highly related to the proposed method.

### 2.1. PCA (*Principal Component Analysis*)

PCA is a second-order method for finding the linear representation of faces using only the covariance of data and determines the set of orthogonal components (feature vectors) which minimize the reconstruction error for a given number of feature vectors. Consider the face image set  $I=[I_1, I_2, \dots, I_n]$ , where  $I_i$  is a  $p \times q$  image,

$i \in [1 \cdots n]$ ,  $p, q, n \in \mathbb{Z}^+$ . The average face of the image set is defined by:

$$\Psi = \frac{1}{n} \sum_{k=1}^n I_k. \quad (1)$$

Normalizing each image by subtracting the average face, we have the normalized difference image:

$$\tilde{D}_i = I_i - \Psi. \quad (2)$$

Unpacking  $\tilde{D}_i$  row-wise, we form the  $N$  ( $N = p \times q$ ) dimensional column vector  $d_i$ . We define the covariance matrix  $C$  of the normalized image set  $D = [d_1, d_2, \dots, d_n]$  by:

$$C = \sum_{i=1}^n d_i d_i^T = DD^T. \quad (3)$$

An eigen-decomposition of  $C$  yields eigenvalues  $\lambda_i$  and eigenvectors  $u_i$  which satisfy:

$$Cu_i = \lambda_i u_i, \quad (4)$$

$$DD^T = C = \sum_{i=1}^N \lambda_i u_i u_i^T, \quad (5)$$

where  $i \in [1 \cdots N]$ .

Normally, the dimensionality of  $C$  is very large and singular decomposition is expensive. In order to reduce computation, we decompose the following expression instead.

$$D^T D = \sum_{i=1}^n \sigma_i^2 v_i v_i^T, \quad (6)$$

$$u_i = \frac{1}{\sigma_i} D v_i, \quad (7)$$

where  $i = [1 \cdots n]$ . It is easily shown that the first  $n$  eigenvalues of  $C$  are the same as those of  $D^T D$  and the corresponding eigenvectors  $u_i$ , often called the eigenfaces, are obtained by (7). Generally, we select a small subset of  $m$  ( $m < n$ ) eigenfaces to define a reduced dimensionality facespace that yields higher recognition performance on unseen examples of faces.

PCA defines a face subspace with reduced dimensionality which contains the greatest covariance and yields good generalization capacity. Nevertheless, since the PCA projection is optimized for minimizing mean square error and does not consider the classification of samples, the features abstracted are not necessarily the best choice for classification. It may retain principal components that represent large within-class variations which are pernicious for recognition. It is been declared that it is only when within- and between-classes variations have the same dominant direction, that principal components are efficient for recognition.<sup>25</sup>

## 2.2. FLD (*Fisher Linear Discriminant*)

FLD finds the optimum projection for classification of the training data by simultaneously diagonalizing the within-class and between-class scatter matrices.<sup>25</sup> The FLD procedure consists of two operations: whitening and diagonalization.<sup>24</sup> Given  $M$  classes  $S_j, j \in [1 \cdots M]$ , we denote the exemplars of each class by  $s_{j,k} = [s_{j,1}, s_{j,2}, \cdots, s_{j,K_j}]$  where  $K_j$  is the number of exemplars in class  $j$ . Let  $\mu_j$  denote the mean of class  $j$  and  $\bar{\mu}$  denote the grand mean for all the exemplars, then the between class scatter matrix is defined by:

$$B = \sum_{j=1}^M K_j (\mu_j - \bar{\mu})(\mu_j - \bar{\mu})^T, \quad (8)$$

and the within-class scatter matrix is defined by:

$$W = \sum_{j=1}^M \sum_{k=1}^{K_j} (s_{j,k} - \mu_j)(s_{j,k} - \mu_j)^T. \quad (9)$$

Therefore the optimised projection  $A$  of size  $M \times M$  is chosen to maximize the ratio of the determinant of the between-class scatter matrix to the determinant of the within-class scatter matrix, i.e.,

$$W_{FLD} = \arg \max_A \frac{|A^T B A|}{|A^T W A|}. \quad (10)$$

The ratio is maximized when  $A$  consists of the leading eigenvectors of  $W^{-1}B$ . But the direct inversion of  $W$ , normally a singular matrix, is numerically unstable. One way to solve this problem is to first diagonalize  $W$  by singular value decomposition, i.e. ,

$$W = \Phi \Lambda \Phi^T, \quad (11)$$

$$(\Phi \Lambda^{-\frac{1}{2}})^T W \Phi \Lambda^{-\frac{1}{2}} = I. \quad (12)$$

Then replace  $W$  with  $B$  to construct a new matrix  $B_{trans}$  for diagonalization,

$$B_{trans} = (\Phi \Lambda^{-\frac{1}{2}})^T B \Phi \Lambda^{-\frac{1}{2}}, \quad (13)$$

$$B_{trans} \Xi = \Xi \Gamma. \quad (14)$$

The final transformation matrix then becomes

$$M_{FLD} = \Phi \Lambda^{-\frac{1}{2}} \Xi. \quad (15)$$

In other words, FLD extracts features that are strong between classes but weak within classes. While FLD often yields higher recognition performance than PCA, it tends to overfit to the training data, since it relies heavily on how the within-class scatter captures reliable variations for a specific class.<sup>24</sup> In addition, it is optimised for specific classes, so the method needs several samples in every class and thus can determine only a maximum of  $M-1$  features, which may degrade the generalization ability.

### 2.3. AAM (Active Appearance Model)

Active Appearance Model (AAM) is based on an assumption that an object can be represented by a linear combination of a series of basic shape and appearance templates. Shape and appearance models are constructed separately from training data and are combined to construct the active appearance model. The statistical shape model is needed to learn the shape variations of an object. For instance, a face image is labeled by  $t$  landmark points  $\{(x_1, y_1), (x_2, y_2), \dots, (x_t, y_t)\}$ , which are located on key feature points, such as profiles of eyes, nose, mouth, chin, etc. The shape of the face  $i$  is then represented by a  $2t$  element vector  $v_i = (x_1, y_1, x_2, y_2, \dots, x_t, y_t)$ . In order to reduce the effect of rotation, translation, and scale on the shape variation, Procrustes Analysis<sup>19</sup> is applied on the vectors of the training images to minimize the distance sum of each shape to the mean shape. After shape normalization, Principal Component Analysis<sup>37</sup> is applied on the sample shape as follows.

We calculate the mean shape vector of all the samples by

$$\bar{v} = \frac{1}{n} \sum_{i=1}^n v_i. \quad (16)$$

Then we calculate the correlation between vector elements by

$$M_{shape} = \frac{1}{n-1} \sum_{i=1}^n (v_i - \bar{v})(v_i - \bar{v})^T. \quad (17)$$

Finally we perform an eigen decomposition of  $M_{shape}$  to get the corresponding eigenvectors  $\phi_i$ . Then each shape can be represented by the following formula:

$$v = \bar{v} + P_s b_s, \quad (18)$$

where  $P_s$  is a shape matrix containing  $m$  eigenvectors with largest eigenvalues. The statistical appearance model can be constructed in a similar way. First each training sample is warped to the mean shape to form a shape invariant texture vector  $g_i$ . Then PCA is applied to the set of training texture vectors to obtain the appearance matrix  $P_a$ . Therefore, each appearance can be described by :

$$g = \bar{g} + P_a b_a, \quad (19)$$

where  $\bar{g}$  is the mean appearance vector.

The shape and appearance parameters  $b_s$  and  $b_a$  can be used to describe the shape and appearance of any example. As there are correlations between the shape and appearance variations of the same person, we can combine the shape and appearance parameters together to form a united vector as follows:

$$b = \begin{pmatrix} W_s b_s \\ b_a \end{pmatrix} = \begin{pmatrix} W_s P_s^T (v - \bar{v}) \\ P_a^T (g - \bar{g}) \end{pmatrix}, \quad (20)$$

8 *S.Chen, T. Shan & B.C. Lovell*

where  $W_s$  is a diagonal matrix which represents the change between shape and texture. We then apply PCA to these vectors and get:

$$b = P_c c = \begin{pmatrix} P_{cs} \\ P_{ca} \end{pmatrix} c, \quad (21)$$

where  $P_c$  is eigenvector matrix and  $c$  is a vector of active appearance parameters controlling both shape and appearance of the model. Finally, the active appearance model which can describe the shape and appearance variations simultaneously is given by:

$$\begin{aligned} v &= \bar{v} + Q_s c \\ g &= \bar{g} + Q_a c, \end{aligned} \quad (22)$$

where  $Q_s = P_s W_s^{-1} P_{cs}$  is the shape variation matrix,  $Q_a = P_a P_{ca}$  is the appearance variation matrix.

### 3. Illumination Variation

Many existing methods have drawbacks when dealing with face images under variations in illumination. Approaches that rely on choosing insensitive features for representation of face images cannot deal with illumination variation well because there are no features sufficiently immune to changes in illumination direction.<sup>1</sup> Other approaches that construct illumination cones for faces need several images for every class and it is hard to extend this method to cope with the other PIE variations. If only one image is available for each class, most existing approaches will not perform well.

We devised a method that inherits merits from both PCA and FLD to overcome many of the above problems.<sup>11,12</sup> We first apply PCA for feature abstraction for the following reasons. First, though it is generally considered that FLD is superior to PCA in pattern classification, sometimes PCA performs better than FLD especially when the number of training samples per class is small.<sup>29,42</sup> For the case of only one gallery image available per class, we would expect that features abstracted by PCA are more suitable than FLD for classification. Second, PCA has good generalization capability while FLD may overfit to the training data which is hazardous when new classes must be handled appropriately. We use the raw data as input for PCA since preprocessing such as edge maps might introduce features that are highly sensitive to certain facial variations. Consequently, every face image  $I_{j,k}$  can be projected into a subspace with reduced dimensionality to form an  $m$ -dimensional feature vector  $s_{j,k}$  with  $k = 1, 2, \dots, K_j$  denoting the  $k^{th}$  sample of the class  $S_j$ .

#### 3.1. Bayes Decision Rule

After constructing the face subspace for image representation, we need to warp this face space to enhance the class separability and improve recognition performance.

The Bayes classifier is the best classifier with minimum error rate for pattern recognition if prior probabilities are known. According to the Bayes decision rule,  $s$  is classified to class  $S_i$  whose posterior probability given  $s$  is the largest among all classes, that is

$$P(S_i|s) = \max_j \frac{p(s|S_j)P(S_j)}{p(s)}, \quad (23)$$

where  $P(S_i|s)$  is the *a posteriori probability* of  $S_i$ ,  $p(s|S_j)$  is the conditional density function of  $S_j$ ,  $P(S_j)$  is the *a priori probability*, and  $p(s)$  is the mixture density function. Since it is difficult to estimate the conditional density functions of certain classes, we generally assume normal Gaussian distribution for simplicity. Consequently, the conditional density function would be:

$$p(s|S_j) = \frac{1}{(2\pi)^{\frac{m}{2}} |cov_j|^{\frac{1}{2}}} \exp[-\frac{1}{2}(s - \mu_j)^T cov_j^{-1}(s - \mu_j)] \quad (24)$$

where  $\mu_j$  is the mean of class  $S_j$  and  $cov_j$  is the covariance matrix of  $S_j$ . Even with this assumption, it is still hard to determine the conditional density function accurately due to the fact that samples are often too few for estimation. Hence, a more strict assumption is needed. PCA treats all features equally so it is implicitly assumed that the within-class covariance is the unit matrix,<sup>26</sup> that is

$$cov = I. \quad (25)$$

But PCA does not take into account the classification of samples and the features extracted represent components of both within-class and between-class covariance. If using the nearest neighbor for classification, the distance between two face vectors is the energy difference between them. In the case of image variations due to illumination, these lighting changes (within-class covariance) become dominant over the characteristic differences between faces (between-class covariance). That is the reason why PCA does not work well in this case.

An improvement of PCA is to apply a standard whitening transformation to compensate for the fact that PCA preferentially weights dominant eigenfaces (low frequencies)<sup>25</sup> according to their eigenvalues. This method estimates the within class covariance as

$$cov = diag\{\lambda_1, \lambda_2, \dots, \lambda_m\}, \quad (26)$$

and the whitening matrix then would be:

$$Z = diag\{\lambda_1^{-\frac{1}{2}}, \lambda_2^{-\frac{1}{2}}, \dots, \lambda_m^{-\frac{1}{2}}\}, \quad (27)$$

where  $m$  is the number of features selected. Though standard whitening improves the performance of PCA slightly,<sup>25</sup> it is not sufficient for illumination invariant face recognition. Indeed, eigenvectors obtained by PCA are the combination of both within-class and between-class covariance. Hence the standard whitening transformation matrix  $cov$  may not weight features optimally and appears to compress the eigenspace so much that class separability is diminished. Chengjun Liu and

Harry Wechsler<sup>26</sup> proposed a method PRM that uses within-class covariance for estimating the conditional density function. They assume that all the within-class covariance matrices are identical and diagonal, that is:

$$cov = diag\{\delta_1^2, \delta_2^2, \dots, \delta_m^2\} \quad (28)$$

where the  $\delta_i^2, i \in m$  are estimated by sample variance in the corresponding eigen direction. However, performance of this method depends on how features capture the within-class covariance. In our trials we have found that this method may also warp the subspace improperly leading to increased misclassifications.

Indeed, in the above methods, we noticed that not all of the features have the same importance for recognition and so we tried whitening adapted to each of the eigenfaces. We finally adopted the technique of a whitening transformation maximizing the ratio of between-class covariance to within-class covariance and proposed our method in two steps: a whitening transformation followed by eigenface filtering.

### 3.2. *Whitening Transformation*

The purpose for whitening is to normalize the scatter matrix for uniform gain control and to compensate for the overweighting of leading eigenfaces. But the above three methods for whitening are not sufficient to compensate for face image variations because the estimation of the conditional density function is not accurate. This is due to the fact that eigen features extracted by PCA represent the overall covariance and the estimation of pdf is affected not only by within-class covariance but also between-class covariance. Moreover, there might exist a bias in the evaluation of the mean for certain classes since the number of samples available may be too few. As illustrated in Figure 1, with the nearest neighbor classifier, if no whitening is applied, we can see from sub-figure A that there is high possibility of misclassifying point A1 in class A and B1 in class B. When we apply standard whitening, class separability might be even worse. As seen from sub-figure B, point A2 and B2 might be misclassified. Only with appropriate whitening can classes A and B be distinguished correctly with the nearest neighbor rule as shown in sub-figure C.

All of the above whitening matrices are influenced by the eigenvectors, so we now assume that our whitening matrix  $Z'$  is a function of eigenvectors,  $Z' = f(\lambda)$ . Though the whitening matrix may improve the recognition performance by taking into account the overall covariance, it could not neutralize the variation of faces resulting from the fact that PCA does not distinguish between the within-class covariance and between-class covariance. Thus, filtering is necessary for feature weighting.

### 3.3. *Filtering the Eigenfaces*

The aim of filtering is to diminish the contribution of eigenfaces that are strongly affected by illumination variation. We want to be able to enhance features that cap-

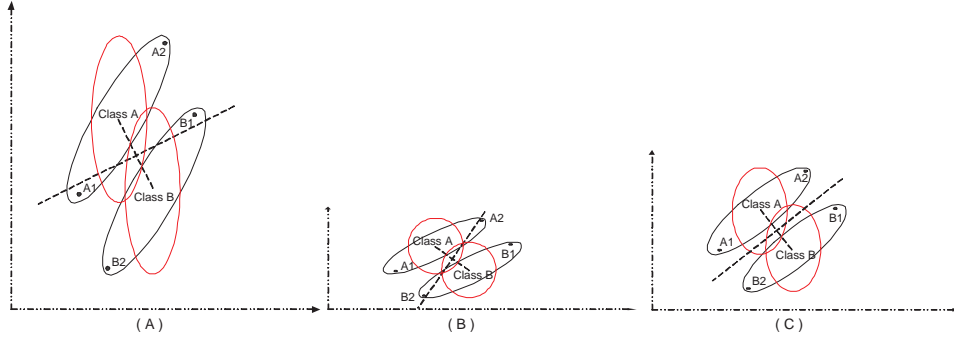


Fig. 1. Whitening of Face Space. Ellipses in black are actual distributions in the space. Ellipses in red are the corresponding estimated distributions.

ture the main differences between classes (faces) while diminishing the contribution of those that are largely due to lighting variation (within-class differences). We thus define a filtering parameter  $\Upsilon$  which is related to identity-to-variation (ITV) ratio. The ITV is a ratio measuring the correlation with a change in person versus a change in luminance for each of the eigenfaces. For an  $M$  class problem, assume that for each of the  $M$  classes (persons) we have examples under  $K$  standardized different lighting conditions — in our case, the lighting source is positioned in front, above, below, left and right as illustrated in Figure 2. Let us denote the  $i^{th}$  element of the face vector of the  $k^{th}$  lighting sample for class (person)  $S_j$  by  $s_{i,j,k}$ . Then

$$\begin{aligned}
 ITL_i &= \frac{\text{Between Class Covariance}}{\text{Within Class Covariance}} \\
 &= \frac{\frac{1}{M} \sum_{j=1}^M \frac{1}{K} \sum_{k=1}^K |s_{i,j,k} - \varpi_{i,k}|}{\frac{1}{M} \sum_{j=1}^M \frac{1}{K} \sum_{k=1}^K |s_{i,j,k} - \mu_{i,j}|}, \\
 \varpi_{i,k} &= \frac{1}{M} \sum_{j=1}^M s_{i,j,k}, \\
 \mu_{i,j} &= \frac{1}{K} \sum_{k=1}^K s_{i,j,k}, \quad i = [1 \dots m].
 \end{aligned} \tag{29}$$

Here  $\varpi_{i,k}$  represents the  $i^{th}$  element of the mean face vector for lighting condition  $k$  for all persons and  $\mu_{i,j}$  represents the  $i^{th}$  element of the mean face vector for person  $j$  under all lighting conditions. We then define the filter matrix  $\Upsilon$  by:

$$\Upsilon = f'(ITL). \tag{30}$$

### 3.4. Cost Function and Optimization

The two weighting matrices  $Z'$  and  $\Upsilon$  are  $m \times m$  matrices. It is hard to search in such a high dimensional space to determine the optimal values for the two matrices. Therefore, we define a simple relationship between  $Z'$  and  $\lambda$ ,  $\Upsilon$  and  $ITV$  to reduce the search space. The whitening matrix is defined by:

$$Z' = \text{diag}\{\lambda_1^p, \lambda_2^p, \dots, \lambda_m^p\}, \tag{31}$$

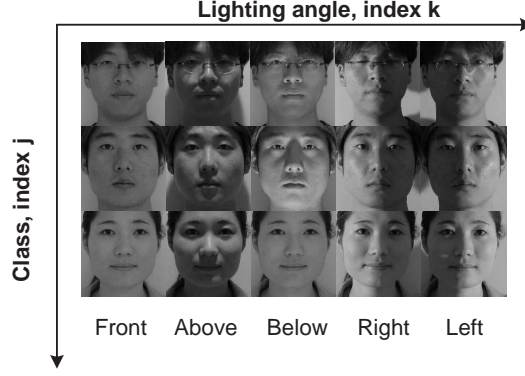


Fig. 2. Examples of images taken under different illuminations in Asian Face Image Database PF01

where the exponent  $p$  is determined empirically. When  $p = -\frac{1}{2}$ , it is a standard whitening matrix. Eigen filtering matrix is defined by:

$$\Upsilon = \text{diag}\{ITL_1^q, ITL_2^q, \dots, ITL_m^q\}, \quad (32)$$

where  $q$  is an exponential scaling factor determined empirically as well.

After the affine transformation, the conditional pdf would be:

$$p(s|S_j) = \frac{1}{(2\pi)^{\frac{m}{2}} \prod_{i=1}^m \lambda_i^{-p} ITV_i^{-q}} \exp\left[-\frac{1}{2} \sum_{i=1}^m \frac{(s_i - \mu_{i,j})^2}{\lambda_i^{-2p} ITV_i^{-2q}}\right] \quad (33)$$

and the distance  $d$  between two face images  $I_{j,k}$  and  $I_{j',k'}$  is defined as the Euclidean distance of their transformed face vectors  $\tilde{s}_{j,k}$  and  $\tilde{s}_{j',k'}$ :

$$\begin{aligned} d_{jj',kk'} &= \|\tilde{s}_{j,k} - \tilde{s}_{j',k'}\|_2 \\ &= \|Z'\Upsilon s_{j,k} - Z'\Upsilon s_{j',k'}\|_2 \\ &= \|Z'\Upsilon(s_{j,k} - s_{j',k'})\|_2. \end{aligned} \quad (34)$$

Therefore, our final set of transformed eigenfaces would be:

$$U' = Z'\Upsilon \text{diag}\left\{\frac{1}{\sigma_1}, \frac{1}{\sigma_2}, \dots, \frac{1}{\sigma_m}\right\} DV, \quad (35)$$

where  $V$  is a matrix composed of eigenvectors from Equation 6.

The whitening transformation acts as a generative classifier which is used to control the overall scatter of all samples and tends to isotropize the subspace, while eigenface filtering plays a role as a discriminative classifier which is designed to enhance the separability of classes and may stretch the space. There should be a trade off between these two effects. Therefore, we only need to search in a

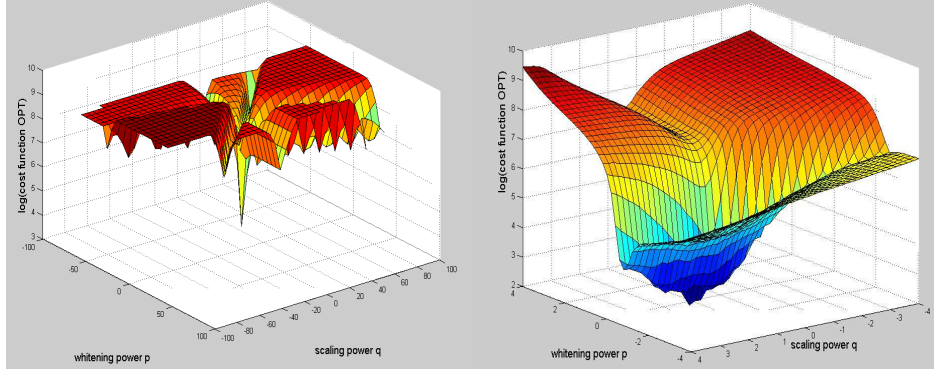


Fig. 3. Relationship between  $OPT$ , whitening power  $p$  and, scaling power  $q$ .

two-dimensional space to determine the two exponents  $p$  and  $q$  for  $Z'$  and  $\Upsilon$ . We introduce a cost function and optimize empirically. It is defined by:

$$OPT = \sum_{j=1}^M \sum_{k=1}^K \sum_m \left( \frac{d_{jj,k0}}{d_{jm,k0}} \right), \forall m \in d_{jm,k0} < d_{jj,k0}, m \in [1 \cdots M]. \quad (36)$$

where  $d_{jj,k0}$  is the distance between the sample  $I_{j,k}$  and  $I_{j,0}$  which is the standard image reference for class  $S_j$  (typically the normally illuminated image). Note that the condition  $d_{jm,k0} < d_{jj,k0}$  is only true when there is a misclassification error. Thus  $OPT$  is a combination of error rate and the ratio of between-class distance to within-class distance. By minimizing  $OPT$ , we can determine the best choices for  $p$  and  $q$  to maximally separate different classes. We find the minimum  $OPT$  by searching in a compact region with  $p$  and  $q$  in the interval  $[-100, 100]$  respectively. Although this is an extremely large range, it illustrates that there is a unique minimum. In order to speed up the search procedure, we apply the pyramid approach. We firstly search in the range  $[-100, 100]$  with a step of 4 to find the possible minimum  $OPT$ . Then we reduce the step size and search in a local area centered at the possible minimum  $OPT$ . This procedure is recursively processed until certain precision is reached. Figure 3 shows the relationship between  $OPT$  and  $p, q$  for one of the training databases in the interval  $[-100, 100]$  and  $[-4, 4]$  respectively. The minimum  $OPT$  at 7.80 is obtained at  $p = -0.3, q = 1.4$  compared to the standard whitening with an  $OPT$  at 22.38. In all of our tests,  $OPT$  is obtained in the interval  $[-2, 2]$ .

### 3.5. APCA Algorithm

From the above, the algorithm works as follows:

- Apply PCA on samples to abstract eigenfaces and eigenvalues as in (6) and (7) and choose the first  $m$  leading eigenfaces. Then project all of the training images into this face space.

Table 1. Comparison of PCA, FLD and APCA for illumination variations with 15 features

Methods	PCA	Fisher	APCA
Training Data	68.33%	91.11%	96.11%
Testing Data	48.73%	69.01%	83.66%

Table 2. Comparison of PCA, FLD and APCA for illumination variations with 25 features

Methods	PCA	Fisher	APCA
Training Data	73.33%	97.22%	99.44%
Testing Data	50.98%	77.74%	90.42%

- Calculate overall covariance  $\lambda$  and ITV as shown in Equation 29.
- Search in the space to minimize the cost function OPT to determine the optimal parameters  $p$  and  $q$ .

### 3.6. Database Evaluation

The method is tested on an Asian Face Image Database PF01,<sup>36</sup> consisting of 535 facial images under 5 different standardized illuminations corresponding to 107 subjects<sup>a</sup>. The size of each image is  $171 \times 171$  pixels with 256 grey levels per pixel. Figure 2 shows some examples from the database. To evaluate the performance of our method, we perform a 3-fold cross validation on the database as follows. We choose one-third of the 107 subjects as training data to construct our APCA model. Then we enrolled the normally lit images of our 71 unseen people as the gallery (pictures in the first column in Figure 2). We then use other unseen images of the above 71 people under different lighting conditions for testing (total 284 images). This process is repeated three-fold using different partitions and the performance is averaged on both training and unseen test sets. Table 1 is the experimental results achieved with only 15 eigen features and Table 2 is the results obtained with 25 features.

It is clear from the results that APCA performs much better than both PCA and FLD in face recognition under variable lighting conditions. Although FLD is fine for training data with 91.11% recognition rate, the performance decreases significantly for the testing data, which demonstrates the lack of generalisation ability. The recognition rate for training data is not 100% because we use the normal lighting image for matching instead of the mean of a certain class. The proposed APCA outperforms PCA and FLD remarkably in recognition rate with 96.11% for training data and 83.66% for testing data with little reduction in performance for normally lit faces.

<sup>a</sup>The current Asian Face Database contains 103 subjects. We use an older version which is a superset of the current one and has 56 male and 51 female.

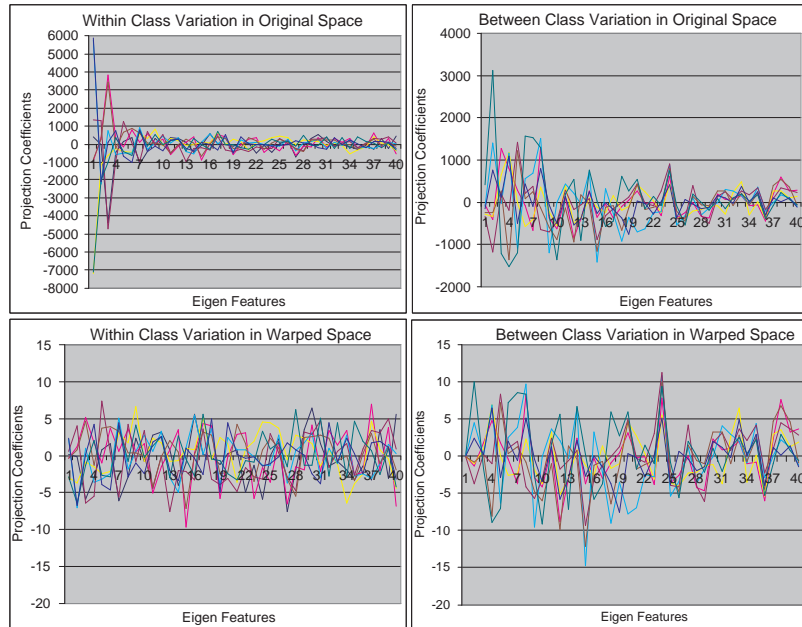


Fig. 4. Within and between class projection coefficients variation in original PCA subspace and the warped APCA subspace

We have compared the variation of projection coefficients of face images in the original PCA constructed subspace and the new warped APCA face space. Figure 4, demonstrates within-class and between-class projection coefficients variation in original and warped face subspace respectively. It can be seen in figure 4 that within class coefficient variation is much greater than between class variation in the original face subspace, which proves that the difference in face images with pure illumination changes (within-class variation) is dominant over the characteristic difference between faces (between-class variation). Comparing the bottom two sub-figures we can see that after our proper whitening and eigen filtering, with-class projection coefficients variation in wrapped subspace is reduced to a range from -5 to 5, which is smaller than that of between-class which ranges from -10 to 10. It confirms that our proposed APCA method weights features appropriately, which enhances those features that are highly correlated to between class differences and weakens features that are strongly related to nuisance within class variations.

We have also tested our method on the Yale Face Database B.<sup>18</sup> This requires rescaling the images to 171 by 171 pixels to match the Asian Face Database<sup>36</sup> and image registration according to the position of eyes. Figure 5 shows the recognition rate for the APCA and PCA methods applied to Yale Face Database B<sup>18</sup> when trained on the Asian Face Database.<sup>36</sup> APCA is once again shown to be more reli-

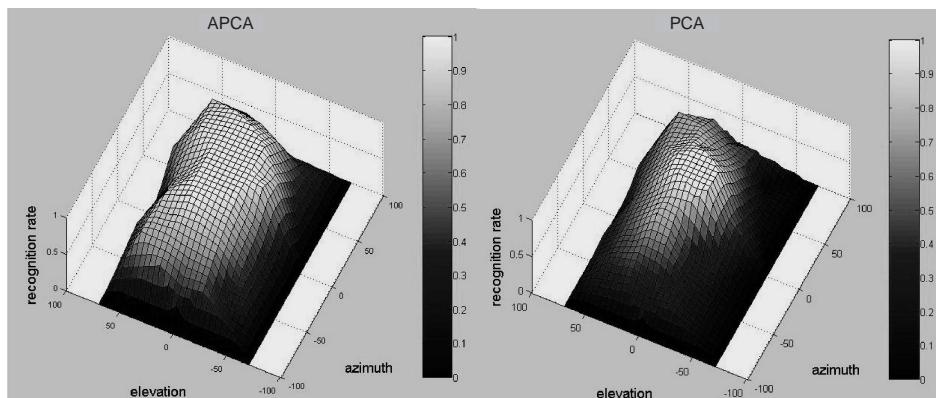


Fig. 5. Recognition rate of APCA and PCA on Yale Face Database B under varying lighting source angles when trained on the Asian Face Database PF01.

able than PCA on face recognition under variant illuminations. The recognition rate of APCA on face images with given lighting source angles is always higher than that of PCA on face images under the same lighting conditions. Moreover, the recognition rate of APCA does not change dramatically with the azimuth and elevation, while it drops greatly for PCA with the azimuth and elevation away from origin. Hence, peak of the “mountain” of APCA looks quite flat while the “mountain” of PCA is sharper. Figures 6 demonstrates the 95% correct recognition contour for APCA and PCA respectively. This result shows that APCA can recognize pictures taken with lighting source angles with azimuth angle range from  $-50$  to  $+50$  degrees and elevation angle range from  $-30$  to  $+30$  degrees compared to PCA with an angle range of only about  $-20$  to  $+20$  degrees for both azimuth and elevation. This quite good result was achieved even though we used the Asian Face Image Database<sup>36</sup> for training instead of the Yale Database.<sup>18</sup>

#### 4. Expression Variation

We applied similar techniques to face images with variations in expression, but could not attain levels of performance comparable to those obtained on illumination variant faces<sup>11</sup> and suggest the following explanation. Eigenfeatures extracted by PCA on face images with illumination variation naturally cluster into two groups: features strongly related to within-class variance, and features strongly related to between-class variance. Usually the first three eigenfaces are strongly related to illumination (within-class) variation. Therefore, it is easy to find the eigenfeatures that represent within-class variations and suppress these with eigenfiltering. However, for expression change, since different people display the same expression in different ways, PCA does not successfully separate between-class and within-class features. This means that the estimation of conditional pdfs may be significantly affected by between-class covariance, which may result in low recognition rates. This is corrob-

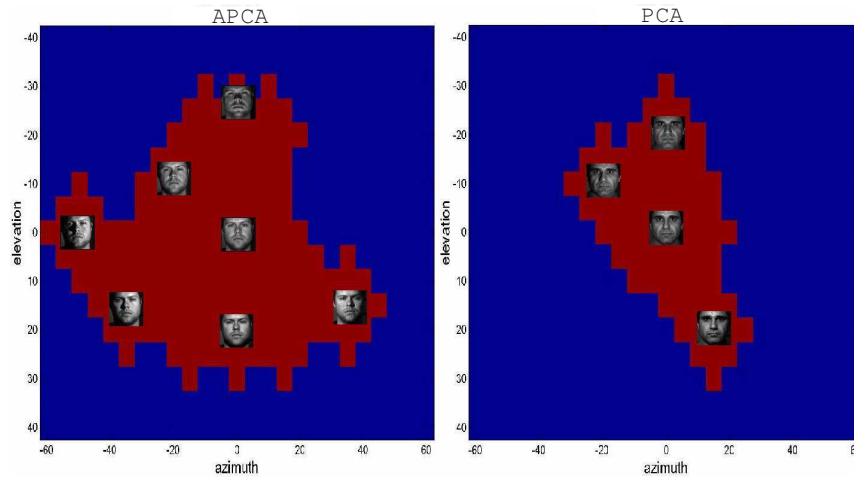


Fig. 6. 95% Recognition rate contour for APCA and PCA on Yale Face Database B under varying lighting source angles when trained on Asian Face Database PF01.

orated by the fact that the identity-to-variation ratio  $ITV$  is roughly close to 1 for most eigenfeatures. This means that when we compress the space in one direction, the within-class covariance and between class-covariance are both affected leading to poor separability.

#### 4.1. Space Rotation

It is difficult to completely separate within- and between-class features because they might be highly correlated and are not orthogonal. We therefore rotate the feature space according to within-class covariance to enhance the representativeness of the features to improve our estimation of the conditional pdfs. After appropriate rotation, most features now represent predominantly between- or within- class variation and by enhancing these between-class features and weakening those within-class features via eigenfiltering the influence of the within-class variation on recognition is diminished.

Moreover, after rotation, since within-class covariance dominates over between-class covariance in the best within-class features, whitening compression in this general direction affects within-class covariance more than between-class covariance and hence improves separability. This effect is illustrated in Figure 7. Sub-figure A shows the original distribution of class A and class B in black ellipses and their corresponding estimated distributions are shown in red ellipses. Sub-figure B illustrates the effect of standard whitening, where points A2 and B2 may be misclassified. The effect of the original space rotation is depicted in sub-figure C, in which the red lines indicate new axes after rotation. As can be seen from sub-figure D that with proper whitening after space rotation, classes are completely separated from each

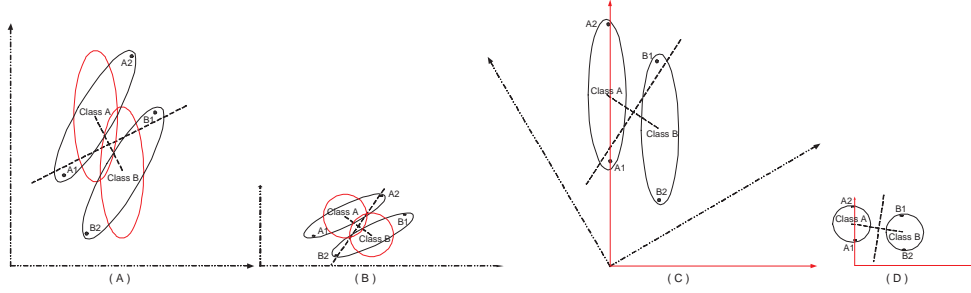
18 *S.Chen, T. Shan & B.C. Lovell*


Fig. 7. Effect of Face Space Rotation. (A) Original distribution. (B) Distribution after standard whitening. (C) Distribution after rotation. (D) Distribution after rotation and whitening.

other.

Figure 8 demonstrate the ITV distribution of features in the original and rotated space for illumination, expression and combined variation faces respectively. We can see that after rotation there are much fewer features with an ITV value at around 1.0 compared to the original set of features. Moreover, features become highly discriminative since most features have an ITV bigger than 2 or less than 0.5, which demonstrates an improvement of the representativeness of the features.

The rotation matrix  $R$  is obtained by applying singular value decomposition on within class covariance  $W$  as in Equation 11,  $R = \Phi$ . Then every face vector  $s$  is transformed into the new space by  $R$ ,

$$r = R^T s. \quad (37)$$

After rotation, the new whitening matrix  $Z'_{rotate}$  would be:

$$Z'_{rotate} = \text{diag}\{\lambda'_1, \lambda'_2, \dots, \lambda'_m\}, \quad (38)$$

$$\lambda' = R^T \lambda \quad (39)$$

Then we apply our previous technique of whitening and filtering to the new space. Figure 9 illustrates the procedure of the recognition process.

#### 4.2. Final Algorithm

The final algorithm works as follows:

- Apply PCA on samples to compute eigenfaces and eigenvalues as in Equation 6 and 7 and choose the first leading  $m$  eigenfaces. Then project all the raw images into these face space.
- Apply PCA on within class covariance to calculate the rotation matrix  $R$ .
- Calculate overall covariance  $\lambda$  and ITV in rotated space.
- Search in the space to minimize the cost function  $OPT$  to determine the optimal parameters  $p$  and  $q$ .

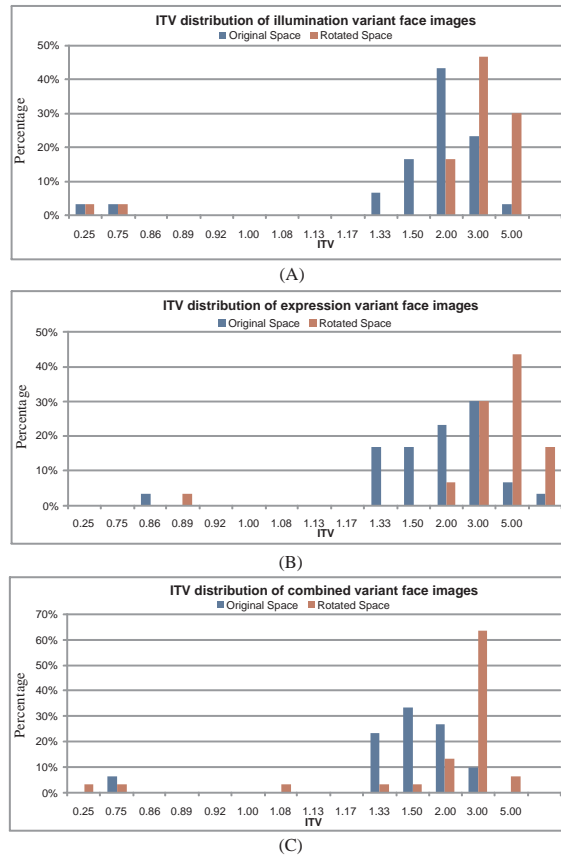


Fig. 8. ITV distribution in original and rotated space. (A) Illumination variant face images. (B) Expression variant face images. (C) Combining illumination and expression variant face images.

#### 4.3. Database Evaluation

We tested our method on the Asian Face Image Database PF01<sup>36</sup> again. A set of 144 images with various facial expressions from 36 people were used to construct the face subspace, the remaining 71 people with 284 images were used for testing except in the case of APCA with rotation. In this case, because rotation according to within-class covariance of sample images may result in an overfitting to the training set training, we used 36 people for training and another 36 for testing. All the experiment is done using three fold cross validation and the performance was averaged. Table 3 shows the experimental results for face recognition on facial expression variations. The same technique has also been tested on both lighting changes and expression changes by combining all images from the above experiments and the result are shown in Table 4.

It can be seen that FLD performs better than PCA on both training and testing,

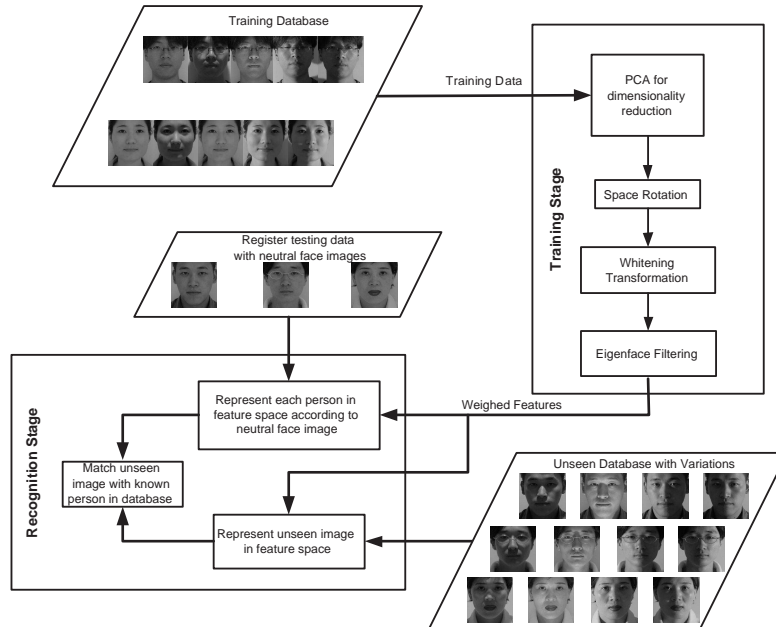


Fig. 9. Procedure for the Recognition Process.

but its generalization ability is worse since the performance drops more than PCA for testing data. APCA has similar performance compared to FLD with a little lower recognition rate in training and higher rate in testing. However, the performance of APCA on training and testing data does not change significantly, which shows that APCA does not overfit to training and generalises well to unseen data. If we apply the rotation transformation, APCA can further improve generalization capacity and outperforms both PCA and FLD remarkably.

Figure 10, 11 and 12 show the test performance for the following techniques

Table 3. Face recognition on facial expression variations with 15 features

Methods	PCA	FLD	APCA	APCA with rotation
Training Data	88.89%	91.67%	88.89%	93.06%
Testing Data	83.45%	86.97%	83.45%	90.0%

Table 4. Face recognition on both expression and illumination variations with 15 features

Methods	PCA	FLD	APCA	APCA with rotation
Training Data	73.96%	90.28%	86.81%	91.43%
Testing Data	59.68%	75.53%	81.87%	90.97%

PCA, FLD, APCA, PRM, Rotated PRM(RPRM) and Rotated APCA(RAPCA) applied on three different kinds of facial variations respectively. We can see again that APCA after rotation achieves the best performance among all techniques. Though FLD can achieve around 90% accuracy for expression variant faces, its performance drops significantly for illumination variant and combination variant faces. Moreover, the recognition rate of FLD changes considerably with the number of eigen features. All the test have shown that rotation can improve the performance since recognition rate for APCA and PRM increase in 12 percent and 5 percent respectively after rotation. Among all the techniques, APCA with rotation is the most robust one since the performance is insensitive to the number of features used for representation and it is also immune to all three kinds of variations.

We also tested the top- $N$  recognition performance of PCA, FLD and RAPCA methods. The top- $N$  classification accuracy is the percentage of times that one of the top- $N$  candidates is the correct match when choosing  $N$  closest candidates from the database. If  $N = 1$ , the accuracy is the normal recognition rate. Figures 13, 14 and 15 show the cumulative accuracy of the top- $N$  classification of face images with combined variations using three different methods with different number of

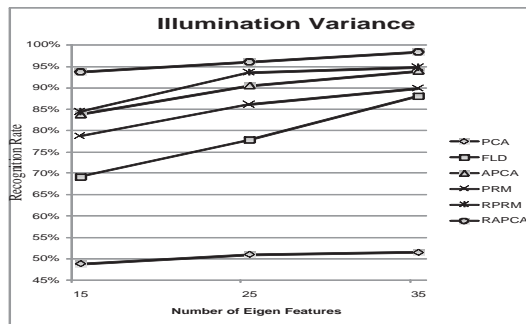


Fig. 10. Test performance for different techniques on illumination variant faces

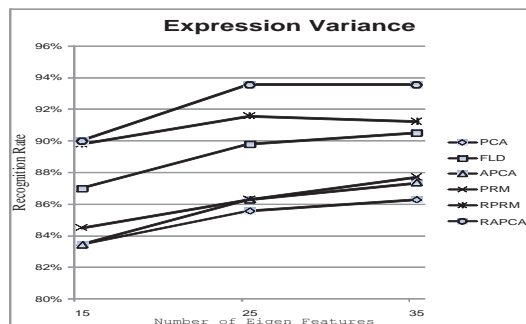


Fig. 11. Test performance for different techniques on expression variant faces

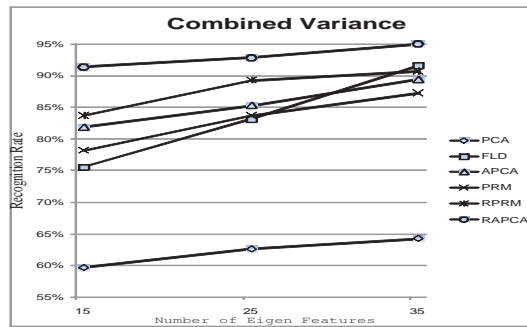


Fig. 12. Test performance for different techniques on combined variant faces

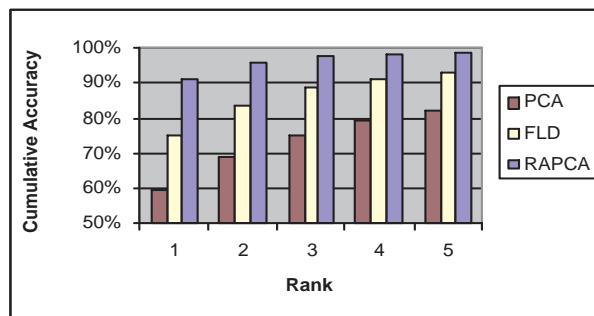


Fig. 13. Top- $N$  accuracy of face images with combined variations using 15 features.

features respectively. It can be seen that RAPCA always outperforms PCA and FLD regardless of the value  $N$  and the number of features used. With an increase in the value  $N$ , the cumulative accuracy of PCA and FLD raises significantly. The recognition rate of PCA and FLD also increases when more features are used. While for RAPCA, the recognition rate does not change dramatically with variations in  $N$  or the number of features. Because the performance of RAPCA is quite good and stable, the recognition rate is possibly too high for us to see further improvement. However, if we look at the percentage of the relative top- $N$  error rate, we can see the improvement of accuracy with an increase in the value  $N$ . Figures 16, 17 and 18 plot the relative error rate of the top- $N$  classification of face images with combined variations using three different methods with different number of features respectively. It can be seen that when  $N = 1$  the relative error rates of three methods in three figures are all 100% because we consider the number of error of the top-1 classification for the corresponding method as a baseline. It is apparent that with the increase in value  $N$ , the relative error rate drops significantly for all of the three methods. RAPCA performs the best with the lowest relative error rate when  $N$  ranges from 2 to 5 and PCA is worse than FLD with a higher relative error

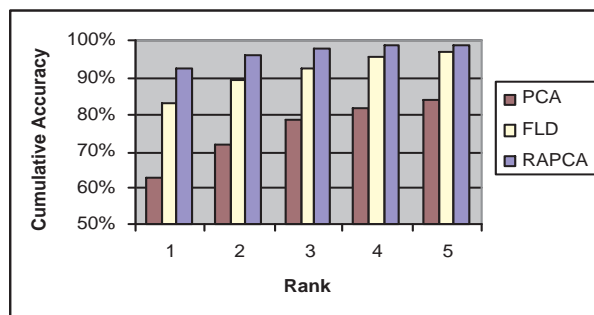


Fig. 14. Top-N accuracy of face images with combined variations using 25 features.

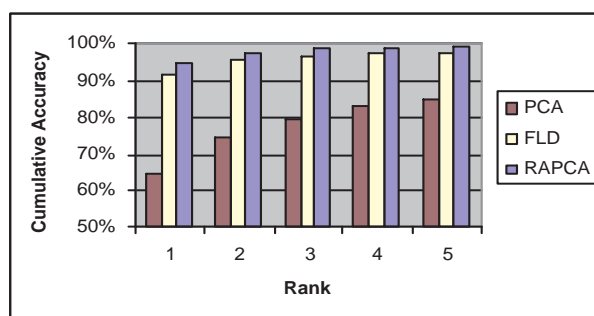


Fig. 15. Top-N accuracy of face images with combined variations using 35 features.

rate. In addition, the relative error rate decreases dramatically for RAPCA when  $N \leq 3$  from 100% to around 25% whilst the relative error rate for PCA and FLD are about 60% and 45% respectively with an  $N$  at 3. Moreover, results show that the top-2 recognition rate of RAPCA is always greater than 95.7% even with only 15 features. However, the top-2 recognition rate for PCA and FLD with 15 features are 69.1% and 83.8% respectively. Hence, the distinguishability of RAPCA is much better than that of PCA and FLD not only in the cumulative recognition rate but also in the top-N relative error rate.

Figure 19 shows the first 9 eigen-faces extracted by PCA and space rotation of RAPCA from face images with combined variations according to the corresponding largest 9 eigenvalues individually. It can be seen from the bottom row that the first three eigen-features of RAPCA are highly correlated to illumination variation and the fourth and last three features are relevant to expression variations. While those features extracted by PCA in the top row do not have significant representation of within-class or between-class variations. It shows that space rotation can extract more representative features. In addition, the corresponding eigenvalues of the first 9 features are the top 9 among all features, which implies that variation of face

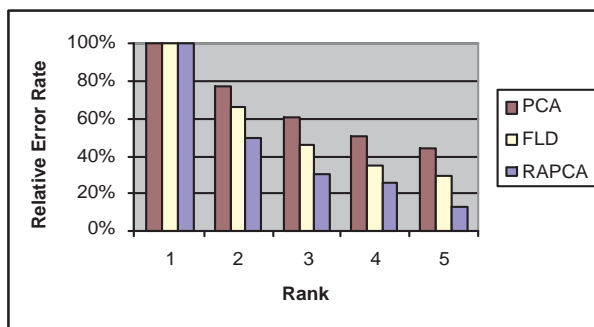


Fig. 16. Top-N relative error rate of face images with combined variations using 15 features.

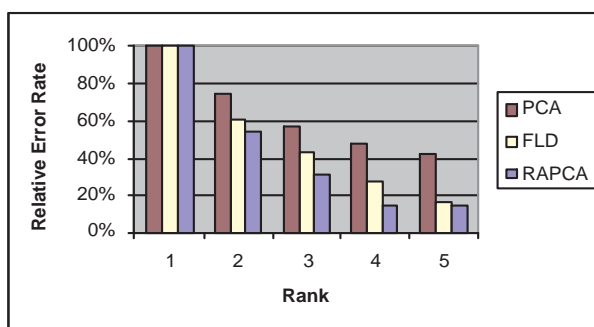


Fig. 17. Top-N relative error rate of face images with combined variations using 25 features.

images on these features are greater than that of other features. Because these features are highly correlated to lighting changes or expression changes, it confirms that differences between face images of the same person due to these nuisance variations is greater than those between different faces. Figure 20 illustrates the last 9 eigen-faces that are assigned minimal weight after whitening and eigen-filtering of RAPCA. It is apparent that the first three images are highly correlated to lighting changes. One can even suggest that these images illustrate the effect of light coming from left, top and bottom respectively. The fourth, sixth and eighth images appear to expressions of smiling and surprise respectively. Because these features are assigned smaller weights compared to other features, the RAPCA method diminishes the contribution of these features which weakens the effect of corresponding variations on face recognition. Moreover, comparing eigen-faces in figure 20 and those in the bottom row of figure 19, we can see that 8 out of 9 faces are the same in both figures, but their order are different. That is because we change the importance of features by two operations: whitening and eigen-filtering. Hence, not only the eigenvalues but also the *ITV* value affect the contribution of features.

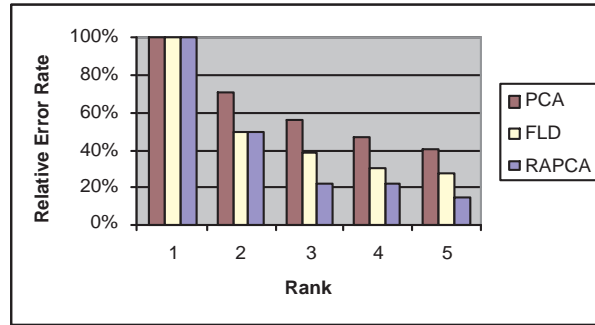


Fig. 18. Top-N relative error rate of face images with combined variations using 35 features.

### 5. Pose Variant Faces

The proposed APCA and RAPCA method is based on PCA extracted features, which are global features and are sensitive to pose variations<sup>23,33,35</sup>. Normalization is therefore needed to align facial feature<sup>42</sup> and it can improve the performance of face recognition as well<sup>10,20,21,35</sup>. We apply Active Appearance Model(AAM) to align the key points on face images and construct a model to synthesize a frontal view image for recognition as the following.

#### 5.1. Pose Estimation

Cootes et al<sup>13</sup> assume that the model parameters  $c$  from (22) trace out an approximately elliptical path when the orientation changes and is affected by the

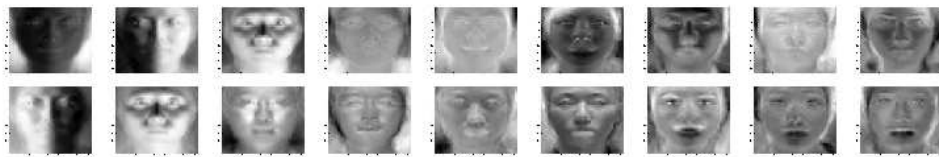


Fig. 19. The first 9 eigen-faces extracted by PCA and RAPCA. The first row shows the first 9 eigen-faces extracted by PCA according to their corresponding eigen-values. The second row illustrates the eigen-faces extracted after space rotation of RAPCA.



Fig. 20. The last 9 eigen-faces after whitening and eigen-filtering of RAPCA. The weights assigned to these eigen-features are the minimal 9 among all the features.

26 *S.Chen, T. Shan & B.C. Lovell*

orientation angle  $\theta$  as the following:

$$c = c_0 + c_c \cos \theta + c_s \sin \theta, \quad (40)$$

where  $c_0, c_c$  and  $c_s$  are the vectors learned from the training data. For each training sample  $k$  with orientation angle  $\theta_k$ , we perform AAM search to find the best parameter  $c_{\theta_k}$  that describe the image appearance in the model. We can apply a regression matching between vectors  $c_{\theta_k}$  and vectors  $(1, \cos \theta_k, \sin \theta_k)$  to obtain  $c_0, c_c$  and  $c_s$ . After building up the correlation model between  $c$  and  $\theta$ , we can estimate the orientation angle of a new face image. We first transform (40) to the following:

$$c - c_0 = (c_c c_s) \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix}. \quad (41)$$

Let  $R_c^{-1}$  be the left pseudo-inverse of the matrix  $(c_c | c_s)$ , then we have:

$$R_c^{-1}(c - c_0) = \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix}. \quad (42)$$

The estimation of  $\theta$  is not entirely accurate which may be caused by land mark annotation errors or regression learning errors. But this model is good enough to be used to synthesize the face image.

## 5.2. Frontal View Synthesis

Once we get the orientation angle  $\theta$ , we can synthesize a frontal view face image and send it to face recognition. Assume we want to synthesize the appearance of the same face at angle  $\alpha$ , following (40) we get:

$$c_\alpha = c_0 + c_c \cos \alpha + c_s \sin \alpha. \quad (43)$$

However, there always have error between the orientation angle and the estimation, it will bring about differences between two sides of (40). This may generate artificial appearance on the reconstructed image, which could significantly degrade the recognition performance. We therefore introduce the residual vector in an attempt to compensate for this unexpected effect. The residual vector is defined as:

$$c_{res} = c - (c_0 + c_c \cos \theta + c_s \sin \theta). \quad (44)$$

Then the new correlation equation becomes:

$$c_\alpha = c_0 + c_c \cos \alpha + c_s \sin \alpha + c_{res}. \quad (45)$$

Because we need to reconstruct the frontal view face image for recognition, angle  $\alpha$  should be 0. Thus, the frontal view AAM parameter is calculated as:

$$c_f = c_0 + c_c + c_{res}. \quad (46)$$

Then we follow (22) to reconstruct the frontal view appearance as:

$$\begin{aligned} v_f &= \bar{v} + Q_s c_f \\ g_f &= \bar{g} + Q_a c_f. \end{aligned} \quad (47)$$

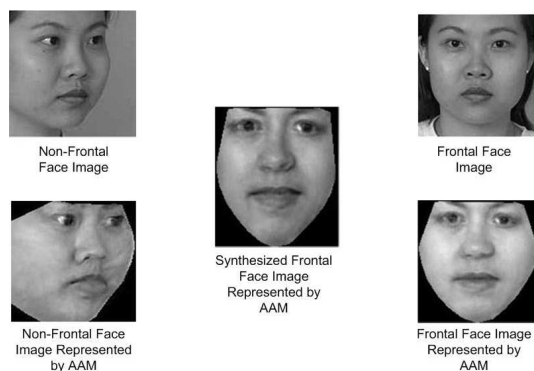


Fig. 21. Reconstruction of frontal view face image from non-frontal face images.

Figure 21 shows an example of frontal view face image reconstruction from non-frontal face images from the FERET database.<sup>31</sup>

### 5.3. Database Evaluation

After frontal view face image reconstruction, we can then apply APCA for face recognition. Our APCA model is trained on frontal face images in the Asian Face Database<sup>36</sup> as done in Section 3 and 4. As the Asian and Yale databases do not include head pose variations, we use the FERET database<sup>31</sup> to evaluate performance. We choose face images from 46 persons with good AAM search results from the FERET database.<sup>31</sup> Each person has 5 face images with the corresponding pose angles ranging from left  $25^\circ$ ,  $15^\circ$ ,  $0^\circ$  to right  $15^\circ$ ,  $25^\circ$ . To evaluate the effect of frontal view synthesis on face recognition, we perform two tests: one is to evaluate on original image dataset, the other is to evaluate on synthesized image dataset. Each of them contains 230 ( $46 \times 5$ ) images and we only register the 46 frontal view images into the gallery and use the rest images with different pose angles for testing. For the original image dataset, face images are aligned according to eye positions and normalized to 171 by 171 pixels. For synthesized face image dataset, all the images are reconstructed to synthesize their corresponding frontal view face images. The test results are illustrated in Figure 22. As can be seen from the figure, the recognition rates of PCA and APCA on synthesized images is much higher than that of the original images. The recognition rate increases by a factor of 4 from 12% to 57% for images with view angle of 25 degrees. For smaller rotation angles less than  $15^\circ$ , the accuracy increases by up to 50 percent. Note that recognition performance of APCA is significantly higher than PCA in disregard of the angle or image dataset, which is in consistent with the test results for illumination and expression variations. There is a trend that the recognition rates of both original and synthesized face images reduce with an increase in rotation angle. However, for synthesized face images this reduction is much less than that of original images.

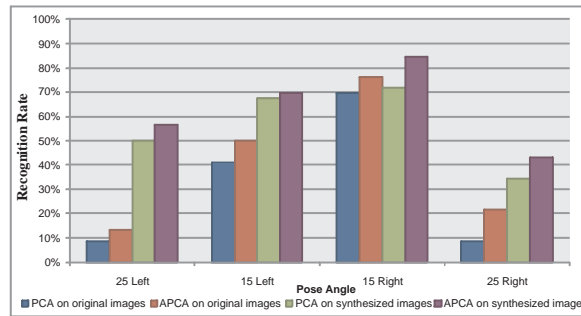


Fig. 22. Recognition rate for PCA and APCA on original and synthesized images.

The decrease of the recognition rate is because we only register frontal images into database and large rotation angles will produce significant distortion of face features compared to frontal face images. The test results show that our AAM model can compensate for this distortion to a certain degree, hence recognition on synthesized images is much less sensitive to pose variations.

## 6. CONCLUSION

In this paper we describe an Adaptive Principal Components Analysis (APCA) method for illumination and expression invariant face recognition. The APCA features are extracted from standard PCA features in three steps: first, a space rotation to improve the representativeness of features; second, a whitening transformation to normalize the scatter matrix; finally, eigenface filtering to enhance the separability of classes. Three-fold cross-validated studies on the Asian Face Database<sup>36</sup> and Yale Face Database<sup>18</sup> show that APCA performs significantly better than both PCA, PRM and FLD methods in terms of accuracy, robustness and generalization ability. We also find out that rotation can enhance the representativeness of features and improve the performance of recognition. We then extend APCA by applying the AAM method to reconstruct frontal face images from non-frontal face images to deal with significant pose variations. Experiments on the FERET Database<sup>31</sup> show that frontal face reconstruction can compensate for pose change and recognition on the synthesized images is much less sensitive to head pose. We are currently investigating the application of APCA on face recognition in real-life environment and will extend it for real-time Intelligent CCTV systems.

## Acknowledgements

Some work of this paper first appear in Ref 12. We would like to thank Sai Sun for preliminary discussions of APCA and the reviewers of Ref 11, 12, 35 and this paper for their feedback. The research described in this paper is supported by a grant from the Australian Government Department of the Prime Minister and Cabinet

and by the Australian Research Council through the Research Network for Securing Australia. NICTA is funded by the Australian Government's *Backing Australia's Ability* initiative, in part through the Australian Research Council.

## References

1. Y. Adinj, Y. Moses, and S. Ullman. Face recognition: The problem of compensation for changes in illumination direction. *IEEE Transactions on PAMI*, 19(7):721–732, 1997.
2. R. Basri and D. W. Jacobs. Lambertian reflectance and linear subspaces. *IEEE Transactions on PAMI*, 25(2):218–233, 2003.
3. P. Belhumeur and D. Kriegman. What is the set of images of an object under all possible illumination conditions. *Intl. J. of Computer Vision*, 28(3):245–260, 1998.
4. P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman. Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. *IEEE Transactions on PAMI*, 19(7):711 – 720, 1997.
5. D. Beymer and T. Poggio. Face recognition from one example view. In *Intl. Conf. on Computer Vision*, 1995.
6. D. J. Beymer. Face recognition under varying pose. In *Intl. Conf. on Computer Vision and Pattern Recognition*, 1994.
7. D. J. Beymer. Feature correspondence by interleaving shape and texture computations. In *Intl. Conf. on Computer Vision and Pattern Recognition*, 1996.
8. M. J. Black, D. J. Fleet, and Y. Yacoob. Robustly estimating changes in image appearance. *Computer Vision and Image Understanding*, 78(1):8–31, 2000.
9. K. W. Bowyer, K. Chang, and P. Flynn. A survey of approaches to three-dimensional face recognition. In *Intl. Conf. on Pattern Recognition*, 2004.
10. X. Chai, S. Shan, and W. Gao. Pose normalization for robust face recognition based on statistical affine transformation. 2003.
11. S. Chen and B. C. Lovell. Illumination and expression invariant face recognition with one sample image. In *Proc. of Intl. Conf. on Pattern Recognition*, 2004.
12. S. Chen, B. C. Lovell, and S. Sun. Face recognition with APCA in variant illuminations. In *Proc. of Workshop on Signal Processing and its Applications*, 2002.
13. T. Cootes, K. Walker, and C. Taylor. View-based active appearance models. In *IEEE Intl. Conf. on Automatic Face and Gesture Recognition*, 2000.
14. S. Edelman, D. Reissfeld, and Y. Yeshurun. A system for face recognition that learns from examples. In *European Conf. on Computer Vision*, 1994.
15. W. Gao, S. Shan, X. Chai, and X. Fu. Virtual face image generation for illumination and pose insensitive face recognition. In *Intl. Conf. on Multimedia and Expo*, 2003.
16. Y. Gao and M. K. Leung. Face recognition using line edge map. *IEEE Transactions on PAMI*, 24(6):764 – 779, 2002.
17. Y. Gao, M. K. H. Leung, W. Wang, and S. C. Hui. Fast face identification under varying pose from a single 2-D model view. *Vision, Image and Signal Processing*, 148(4):248–253, 2001.
18. A. S. Georghiades, P. N. Belhumeur, and D. J. Kriegman. From few to many: Illumination cone models for face recognition under variable lighting and pose. *IEEE Transactions on PAMI*, 23(6):643–660, 2001.
19. C. Goodall. Procrustes methods in the statistical analysis of shape. *Journal of the Royal Statistical Society. Series B (Methodological)*, 53(2):285–339, 1991.
20. R. Gross, I. Matthews, and S. Baker. Appearance-based face recognition and light-fields. *IEEE Transactions on PAMI*, 26(4):449– 465, 2004.
21. S. Gundimada and V. Asari. Face alignment and adaptive weight assignment for

- robust face recognition. *Advances in Visual Computing*, 3804:191–198, 2005.
22. P. W. Hallinan. A low-dimensional representation of human faces for arbitrary lighting conditions. In *IEEE Conf. Computer Vision and Pattern recognition*, 1994.
  23. B. Heisele, P. Ho, J. Wu, and T. Poggio. Face recognition: component-based versus global approaches. *Computer Vision and Image Understanding*, pages 6–21, 2003.
  24. C. Liu and H. Wechsler. Enhanced fisher linear discriminant models for face recognition. In *Intl. Conf. on Pattern Recognition*, 1998.
  25. C. Liu and H. Wechsler. Evolution of optimal projection axes (OPA) for face recognition. In *IEEE Intl. Conf on Automatic face and Gesture Recognition*, 1998.
  26. C. Liu and H. Wechsler. Probabilistic reasoning models for face recognition. In *IEEE Conf. on Computer Vision and Pattern Recognition*, 1998.
  27. X. Liu, T. Chen, and B. V. Kumar. Face authentication for multiple subjects using eigenflow. *Pattern Recognition*, 36:313–328, 2003.
  28. A. M. Martinez. Recognizing imprecisely localized, partially occluded and expression variant faces from a single sample per class. *IEEE Transactions on PAMI*, 24(6):748–763, 2002.
  29. A. M. Martinez and A. C. Kak. PCA versus LDA. *IEEE Transactions on PAMI*, 23(2):228–233, 2001.
  30. A. Pentland and B. M. adn Thad Starner. View-based and modular eigenspaces for face recognition. In *Intl. Conf. on Computer Vision and Pattern Recognition*, 1994.
  31. P. J. Phillips, H. Moon, P. J. Rauss, and S. Rizvi. The FERET evaluation methodology for face recognition algorithms. *IEEE Transactions on PAMI*, 20(10):1090–1104, 2000.
  32. P. J. Phillips, W. T. Scruggs, A. J. OToole, P. J. Flynn, K. W. Bowyer, C. L. Schott, and M. Sharpe. Frvt 2006 and ice 2006 large-scale results. In *Face Recognition Vendor Test 2006*, <http://www.frvt.org/FRVT2006/docs/FRVT2006andICE2006LargeScaleReport.pdf>, 2006.
  33. C. Sanderson, S. Bengio, and Y. Gao. On transforming statistical models for non-frontal face verification. *Pattern Recognition*, 39(2):288–302, 2006.
  34. P. Sankaran and V. Asari. A multi-view approach on modular PCA for illumination and pose invariant face recognition. In *Proceedings of Applied Imagery Pattern Recognition Workshop*, 2004.
  35. T. Shan, B. C. Lovell, and S. Chen. Face recognition robust to head pose from one sample image. In *Proc. of Intl. Conf. on Pattern Recognition*, 2006.
  36. Intelligent Multimedia Lab, Department of Computer Science and Engineering, Pohang University of Science and Technology. Asian face image database PF01. <http://nova.postech.ac.kr/>.
  37. M. A. Turk and A. P. Pentland. Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, 3(1):71–86, 1991.
  38. B. S. Venkatesh, S. Panalivel, and B. Yegnanarayana. Face detection and recognition in an image sequence using eigenedgeiness. In *India Conf. on Vision Graphics and Image Processing*, 2002.
  39. L. Wiskott, J. Fellous, N. Kruger, and C. von der Malsburg. Face recognition by elastic bunch graph matching. *IEEE Transactions on PAMI*, 7(19):775–779, 1997.
  40. A. Yilmaz and M. Gokmen. Eigenhill vs. eigenface and eigenedge. In *Intl. Conf. on Pattern Recognition*, 2000.
  41. L. Zhao and Y. Yang. Theoretical analysis of illumination in PCA-based vision systems. *Pattern recognition*, 32(4):547–564, 1999.
  42. W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld. Face recognition: A literature survey. *ACM Computing Surveys*, 35(4):399 – 458, 2003.