Illumination and Expression Invariant Face Recognition with One Sample Image

Shaokang Chen Brian C. Lovell
Intelligent Real-Time Imaging and Sensing(IRIS) Group
The School of Information Technology and Electrical Engineering
The University of Queensland, Australia QLD 4072
shaokang@itee.uq.edu.au lovell@itee.uq.edu.au

Abstract

Most face recognition approaches either assume constant lighting condition or standard facial expressions, thus cannot deal with both kinds of variations simultaneously. This problem becomes more serious in applications when only one sample images per class is available. In this paper, we present a linear pattern classification algorithm, Adaptive Principal Component Analysis (APCA), which first applies PCA to construct a subspace for image representation; then warps the subspace according to the within-class covariance and between-class covariance of samples to improve class separability. This technique performed well under variations in lighting conditions. To produce insensitivity to expressions, we rotate the subspace before warping in order to enhance the representativeness of features. This method is evaluated on the Asian Face Image Database. Experiments show that APCA outperforms PCA and other methods in terms of accuracy, robustness and generalization ability.

1. Introduction

Within the last several years, research on face recognition has been focused on diminishing the impact of changes in lighting conditions, facial expression and poses. Two main approaches have been proposed for illumination invariant face recognition. One is to represent images with features that are less sensitive to illumination change such as the edge maps of an image. But edge features generated from shadows are related to illumination changes and may have a significant impact on recognition. The other main approach supposes that the surface of human faces is Lambertian reflected and convex and tries to construct a low dimensional linear subspace for face images taken under different lighting conditions [3]. But it is hard for these systems to deal with cast shadows. Furthermore, these systems need several images of the same face taken under specific lighting source directions to construct a model of a given face. In many cases, it is hard to meet this requirement, such as recognizing face images from historic photographs.

As for expression invariant face recognition, one approach is to morph images to be the same shape as the one used for training. But 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 of the lack of texture inside the eyes. Another approach is to use optical flow. However, it is difficult to learn the local motions within feature space to determine the expression changes of each face, since different persons express a certain expression with different ways. Martinez [6] proposed a weighting method that weights independently those local areas which are less sensitive to expressional changes. But features that are insensitive to expression changes may be sensitive to illumination changes as noted in [5].

Previous methods dealing with illumination or facial expression variations cannot compensate for both variations simultaneously. We present a new method, Adaptive Principal Component Analysis (APCA) to warp the face space by whitening and filtering eigen features according to the second order statistics of the samples. We further improve APCA by space rotation to enhance the representativeness of features. Experiments show that our method outperforms PCA [1] and Fisher Linear Discriminant (FLD) [2] on face recognition with both illumination and expression changes.

2. Adaptive Principal Component Analysis

We first apply Principal Component Analysis (PCA) [1] for feature abstraction because of its good generalization capacity. We choose to use raw data as samples for PCA since preprocessing such as edge maps might introduce features that are highly sensitive to certain facial variations. Consequently, every face image can be projected into a subspace with reduced dimensionality to form an m-dimensional feature vector $s_{j,k}$ with $k = 1, 2, \cdots K_j$ denoting the k^{th} sample of the class S_j.

Proceedings of the 17th International Conference on Pattern Recognition (ICPR'04)
1051-4651/04 $ 20.00 IEEE

2.1 Bayes Decision Rule

After constructing the face subspace for image representation, we need to warp this face space to enhance the class separability. The Bayes classifier is the best classifier with minimum error rate for pattern recognition if prior probabilities are known. Because it is difficult to obtain conditional density function of certain classes, we generally assume normal distribution for simplification. Consequently, the conditional density function would be:

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

where μ_j is the mean of class S_j and cov_j is the covariance matrix of S_j. A more strict assumption might be needed due to the fact that samples are often too few for estimation. PCA [1] treats all features equally so it is assumed that the within-class covariance is the unit matrix [4], that is

$$cov = I. \quad (2)$$

But PCA does not take into account the classification of samples. 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.

Chengjun Liu and Harry Wechsler [4] proposed a method PRM which assumes all the within-class covariance matrices are identical and diagonal, that is:

$$cov = diag\{\delta_1^2, \delta_2^2, ..., \delta_m^2\} \quad (3)$$

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.

2.2 Whitening and Eigenface Filtering

The above 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 overall covariance and the estimation of the pdfs is affected not only by within-class covariance but also between-class covariance. In order to compensate for the influence of between-class covariance on the estimation of pdf, we introduce a whitening power p to control the distribution, that is

$$cov = diag\{\lambda_1^{-2p}, \lambda_2^{-2p}, ...\lambda_m^{-2p}\}, \quad (4)$$

where $\lambda_i (i = [1...m])$ are the eigenvalues extracted by PCA. Consequently, the whitening matrix Z is:

$$Z = diag\{\lambda_1^p, \lambda_2^p, ...\lambda_m^p\}, \quad (5)$$

where the exponent p is determined empirically.

The aim of filtering is to enhance features that capture 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 Υ which is related to identity-to-variation (ITV) ratio. The ITV is a ratio measuring the correlation of a change in person versus a change in variation 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. 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

$$\begin{aligned} ITV_i &= \frac{BetweenClassCovariance}{WithinClassCovariance} \\ &= \frac{\frac{1}{M}\sum_{j=1}^{M}\frac{1}{K}\sum_{k=1}^{K}|s_{i,j,k}-\varpi_{i,k}|}{\frac{1}{M}\sum_{j=1}^{M}\frac{1}{K}\sum_{k=1}^{K}|s_{i,j,k}-\mu_{i,j}|}, \\ \varpi_{i,k} &= \frac{1}{M}\sum_{j=1}^{M} s_{i,j,k}, \\ \mu_{i,j} &= \frac{1}{K}\sum_{k=1}^{K} s_{i,j,k}, \quad i=[1,\cdots m]. \end{aligned} \quad (6)$$

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 filtering matrix Υ by:

$$\Upsilon = diag\{ITV_1^q, ITV_2^q, ...ITV_m^q\}, \quad (7)$$

where q is an exponential scaling factor determined empirically as well. After the affine transformation, the conditional pdf would be:

$$p(s|S_j) = \frac{\exp[-\frac{1}{2}\sum_{i=1}^{m}\frac{(s_i-\mu_{i,j})^2}{\lambda_i^{-2p}ITV_i^{-2q}}]}{(2\pi)^{\frac{m}{2}}\prod_{i=1}^{m}\lambda_i^{-p}ITV_i^{-q}} \quad (8)$$

and the distance d between two face vectors $s_{j,k}$ and $s_{j',k'}$ is define by the Euclidean distance of their transformed vectors:

$$d_{jj',kk'} = \|Z\Upsilon(s_{j,k} - s_{j',k'})\|_2, \quad (9)$$

Therefore, our final transformation matrix is:

$$U' = Z\Upsilon V, \quad (10)$$

where V is the set of eigenvectors extracted by PCA.

2.3 Cost Function and Experimental Results

The whitening matrix Z controls the overall scatter of all samples and tends to isotropize the subspace while the filtering parameter Υ 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 need to search in a two-dimension space to determine the two exponents p and q for Z and Υ. We introduce the following cost function which is a combination of error rate and the ratio of within-class distance to between-class distance and optimize it empirically. It is defined by:

$$OPT = \sum_{j=1}^{M} \sum_{k=1}^{K} \sum_{m} (\frac{d_{jj,k0}}{d_{jm,k0}}), \quad \forall m \in d_{jm,k0} < d_{jj,k0}, m \in [1 \cdots m]. \tag{11}$$

where $d_{jj,k0}$ is the distance between the sample $s_{j,k}$ and $s_{j,0}$ 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. By minimizing OPT, we can determine the best choices for p and q to maximally separate different classes. We find the minimum OPT by search in the region with p and q in the interval $[-100, 100]$. Although this is an extremely large range, it illustrates that there is a unique minimum. Figure 1 shows the relationship between OPT and p, q for one of the training databases in interval [-100, 100] and [-4, 4]. The minimum OPT at 7.80 is obtained at $p = -0.3, q = 1.4$. In all of our trainings, the OPT is always obtained in interval [-2 2]. The method is tested on an Asian Face Image Database PF01 [7], consisting of 535 facial images under 5 different standardized illuminations corresponding to 107 subjects. The size of each image is 171×171 pixels with 256 grey levels per pixel. We choose one-third of the 107 subjects to construct our APCA model. Then we just register the normally lit faces (images taken with light source positioned in the front) of the remaining two-thirds of the data into our recognition database. We use other unseen images of the above 71 people under different lighting conditions for testing (a total of 284 images). This process is repeated three-fold using different partitions and the performance is averaged. Table I is the experimental results achieved with 20 eigen 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 the 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 class mean. 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.

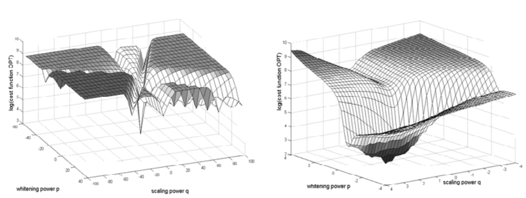

Figure 1. Relationship between OPT, whitening power p and scaling power q.

Table 1. Performance of PCA, FLD and APCA for illumination variations with 20 features

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

3 Rotated APCA

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. That is because Eigenfeatures extracted by PCA on face images with illumination variation naturally cluster into two groups: features strongly related to within-class covariance, and features strongly related to between-class covariance. 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 corroborated by the fact that the identity-to-variation ratio ITV is roughly the same value for all eigenfeatures as shown in Figure 2. When we compress the space in one direction, the within-class covariance and between class-covariance are both affected leading to poor separability.

3.1 Space Rotation

We therefore rotate the feature space according to within-class covariance to enhance representativeness of the features and to improve estimation of the conditional pdfs. After rotation, some features represent predominantly within-class variation and by selecting these via eigenfiltering the influence of between-class variation on estimation is diminished. Moreover, after rotation, features are highly distinguished by their ITV and compression in within-class features will affect within-class covariance more than between-class covariance and hence improves separability. The rotation matrix R is a set of eigen vectors obtained by apply singular value decomposition to the

Proceedings of the 17th International Conference on Pattern Recognition (ICPR'04)
1051-4651/04 $ 20.00 IEEE

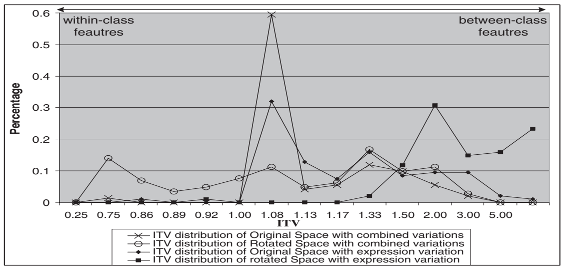

Figure 2. ITV distribution

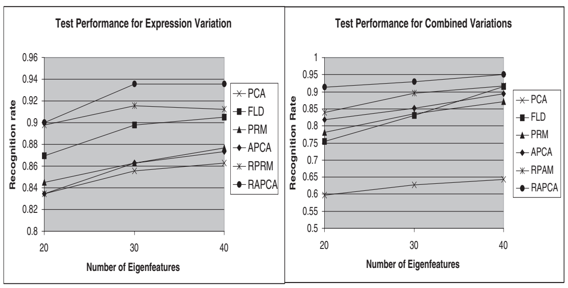

Figure 3. Test performance for different techniques applied on face recognition

within-class covariance matrix. Every face vector s is transformed into the new space by R as follows,

$$r = R^T s. \tag{12}$$

We can see from Figure 2 that after rotation, features are highly distinguished by their ITV thus representativeness of features is improved. After rotation, we need to recalculate the overall covariance λ and ITV, than apply our previous recognition procedure in the new feature space.

3.2 Experimental Results

We tested our methods on the Asian Face Image Database PF01 [7]. We used 284 images with different facial expressions from 71 people to construct the face subspace, the other 36 people with 144 images were reserved for testing. All the experiments are done with three fold cross validation and the performance is averaged. We also apply the same technique on face images with both lighting changes and expression changes by combining all images from the above experiments.

Figure 3 shows the test performance for the following techniques PCA, FLD, PRM, APCA, RPRM and RAPCA applied on two different kinds of facial variations respectively. We can see again that RAPCA achieves the best performance among all techniques. The performance of FLD drops significantly for combined variations and recognition rate changes apparently with the number of eigen features. All the tests have shown that rotation can improve the performance since recognition rates for APCA and PRM increase 10 percent and 5 percent respectively after rotation. Among all the techniques, RAPCA is the most robust one since the performance is insensitive to the number of features used for representation and it is also immune to both variations.

4 CONCLUSION

In this paper we proposed an Adaptive Principle Components Analysis (APCA) method for illumination and expression invariant face recognition. The APCA features are extracted from standard PCA features using two steps: first a whitening transformation to normalize the scatter matrix, then a filtering of the eigenfaces to enhance the separability of classes. Three-fold cross-validated studies show that APCA performs significantly better than both PCA and FLD in terms of accuracy, robustness and generalization ability. We also find out that feature space rotation can enhance the representativeness of features hence improving the performance for recognition. If a rotation transformation is introduced, RAPCA can further enhance the generalization ability and robustness.

However, our rotation is only concentrated on within-class covariance which is likely to be suboptimal. Future work will focus on suitable rotation transformations that achieve better class separability.

References

[1] M. Turk and A. Pentland. " Eigenfaces for recognition", *Journal of Cognitive Neuroscience,* 3(1):71–86, 1991.

[2] P.Belhumeur, J. Hespanha, and D. Kriegman, "Eigenfaces vs. £sherfaces: Recognition using class speci£c linear projection", *IEEE TPAMI,* Vol.19, No.7, 1997, pp.711-720.

[3] Georghiades, A.S. and Belhumeur, P.N. and Kriegman, D.J., "From Few to Many: Illumination Cone Models for Face Recognition under Variable Lighting and Pose", *IEEE TPAMI,* vol.23, No. 6, 2001, pp. 643-660.

[4] Chengjun Liu and Harry Wechsler " Probabilistic Reasoning Models for Face Recognition", *IEEE CVPR,* Santa Barbara, California, USA, 1998.

[5] Y. Adinj, Y. Moses, and S. Ullman, "Face recognition: The problem of compensation for changes in illumination direction", *IEEE TPAMI,* vol. 19, No. 7, 1997, pp721-732.

[6] Aleix M. Martinez, "Recognizing Impercisely Localized, Partially Occluded and Expression Variant Faces from a Single Sample per Class", *IEEE TPAMI,* Vol. 24, No. 6, 2002.

[7] Intelligent Multimedia Lab., "Asian Face Image Database PF01", http://nova.postech.ac.kr/.