Performance analysis of PCA-based and LDA-based algorithms for Face Recognition

Steven Fernandes and Josemin Bala

Steven Fernandes and Josemin Bala are with Department of Electronics and Communication Engineering, Karunya University, Coimbatore, India.

Abstract—Analysing the face recognition rate of various current face recognition algorithms is absolutely critical in developing new robust algorithms. In this paper we report performance analysis of Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) for face recognition. This analysis was carried out on various current PCA and LDA based face recognition algorithms using standard public databases. Among various PCA algorithms analyzed, Manual face localization used on ORL and SHEFFIELD database consisting of 100 components gives the best face recognition rate of 100%, the next best was 99.70% face recognition rate using PCA based Immune Networks (PCA-IN) on ORL database. Among various LDA algorithms analyzed, Illumination Adaptive Linear Discriminant Analysis (IALDA) gives the best face recognition rate of 98.9% on CMU PIE database, the next best was 98.125% using Fuzzy Fisherface through genetic algorithm on ORL database.

Index Terms—face recognition; Principal Component Analysis; Linear Discriminant Analysis; PCA-IN; Illumination Adaptive LDA; Fisher Discriminant.

I. INTRODUCTION

Facial recognition methods can be divided into appearance-based or model-based algorithms. Appearance-based methods represent a face in terms of several raw intensity images. An image is considered as a high-dimensional vector. Statistical techniques are usually used to derive a feature space from the image distribution. The sample image is compared to the training set.

Appearance methods can be classified as linear or non-linear. Linear appearance-based methods perform a linear dimension reduction. The face vectors are projected to the basis vectors, the projection coefficients are used as the feature representation of each face image, and approaches are PCA, LDA, and Independent Component Analysis (ICA) [1], [2]. Non-linear appearance methods are more complicated. Linear subspace analysis is an approximation of a non-linear manifold. Kernel PCA (KPCA) [3] is a method widely used.

Model-based approaches can be 2-Dimensional or 3-Dimensional. These algorithms try to build a model of a human face. These models are often morphable. A morphable model allows classifying faces even when pose changes are present, and approaches are Elastic Bunch Graph Matching [4] or 3D Morphable Models [5]. In this paper we report performance analysis of various current PCA and LDA based algorithms for face recognition. The evaluation parameter for the study is face recognition rate on various standard public databases. The remaining of the paper is organized as follows: Section II provides a brief overview of PCA, Section III presents PCA algorithms analyzed, Section IV provides a brief overview of LDA, Section V presents LDA algorithms analyzed. Section VI presents performance analysis of various PCA and LDA based algorithms finally Section VII draws the conclusion.

II. PRINCIPAL COMPONENT ANALYSIS (PCA)

PCA, also known as the Karhunen-Loeve transform, is a linear dimension-reduction technique. It aims to find the project directions along which the reconstructing error to the original data is minimum, and projects the original data into a lower dimensional space spanned by those directions corresponding to the top eigenvalues. In face recognition, those directions which are the eigenvectors of the covariance matrix of face images are orthogonal basis vectors.

Consider the training sample set of face image \( F = \{ x_1, x_2, \ldots, x_M \} \), where \( x_i \in \mathbb{R}^n \), \( i = 1, \ldots, M \) corresponds to the lexicographically ordered pixels of the \( i \)th face image, and where there are \( M \) face images. PCA tries to mapping the original \( n \)-dimensional image space into an \( m \)-dimensional feature space, where \( m < n \). The new feature vectors \( y_i \in \mathbb{R}^m \) are defined by the following linear transform:

\[
y_i = W^T x_i \quad k = 1, \ldots, M
\]

where \( W = [ w_1, w_2, \ldots, w_m ] \in \mathbb{R}^n \) is the orthogonal with each other is the eigenvector of total scatter matrix \( S_f \) corresponding to the \( m \)th largest eigenvalue. The total scatter matrix is defined as

\[
S_f = \sum_{i=1}^{M} (x_i - \mu)(x_i - \mu)^T
\]

where \( \mu \) is the mean value of all training samples.

III. PCA ALGORITHMS ANALYZED

A. PCA and Support Vector Machine (SVM)

PCA is used to extract the essential characteristics of face images, SVM as classifier. One against one
I. PCA and Minimum Distance Classifier

Different facial images of a single human face are taken together as a cluster. PCA is applied for feature extraction. Minimum distance classifier is used for the recognition that avoids the exploit of threshold value which is changeable under different distance classifiers [14].

IV. LINEAR DISCRIMANT ANALYSIS

Let us consider a set of N sample images \( \{ x_1, x_2, ..., x_n \} \) taking in an n-dimensional image space, and assume that each image belongs to one of c classes \( \{ c_1, c_2, ..., c_c \} \).

Let \( N_i \) be the number of the samples in class \( c_i \) (\( i = 1, 2, ..., c \)). Let \( \mu_i \) be the mean of the samples in class \( c_i \). Then the within-class scatter matrix \( S_b \) is defined as

\[
S_b = \frac{1}{N} \sum_{i=1}^{c} N_i (\mu_i - \mu)(\mu_i - \mu)^T
\]

The within-class matrix \( S_w \) is defined as

\[
S_w = \frac{1}{N} \sum_{i=1}^{c} \sum_{x_i \in c_i} (x_i - \mu_i)(x_i - \mu_i)^T
\]

In LDA, the projection \( W_{opt} \) is chosen to maximize the ratio of the determinant of the between-class scatter matrix of the projected samples to the determinant of the within-class scatter matrix of projected samples:

\[
W_{opt} = \arg \max_i \frac{|W^T S_b W|}{|W^T S_w W|} = [w_1, w_2, ..., w_d]
\]

where \( \{ w_i \} = 1, 2, ..., d \) is the set of generalized eigenvectors of \( S_b \) and \( S_w \) corresponding to the m largest generalized eigenvalues \( \{ \lambda_i = 1, 2, ..., d \} \), i.e.,

\[
S_b w_i = \lambda_i S_w w_i, i = 1, 2, ..., d.
\]

V. LDA ALGORITHMS ANALYZED

A. Regularized-LDA (R-LDA)

R-LDA is used for extracting low-dimensional discriminant features from high dimensional training images and then these features are used by Probabilistic Reasoning Model (PRM) for classification [15].

B. Multi-Feature Discriminant Analysis (MFDA)

Feature extraction method that combines advantages of Two-Dimensional Principal Component Analysis (2D)PCA and Incremental PCA (IPCA). It(2D)PCA consumes less computational load than IPCA as well as smaller memory waste than (2D)PCA [16].

C. Rearranged Modular 2D LDA (Rm2D LDA)

Two-dimensional linear discriminant analysis has lower time complexity but it implicitly avoids the small sample problem encountered in classical LDA. Rm2D LDA was developed. It was based on the idea of dividing an image into sub-images and then concatenating them to form a wide image matrix [17].
D. Illumination Adaptive Linear Discriminant Analysis (IALDAl)

The images of many subjects under the different lighting conditions are used to train illumination direction classifier and varieties of LDA projection matrices. Then the illumination direction of a test sample is estimated by illumination direction classifier, the corresponding LDA feature which is robust to the illumination variation between images under the standard lighting conditions and the estimated lighting conditions is extracted [18].

E. Fuzzy Fisherface (FLDA) through genetic algorithm

Searches for optimal parameters of membership function. The optimal number of nearest neighbors to be considered during the training is also found through the use of genetic algorithms [19].

F. Semi-supervised face recognition algorithm based on LDA self-training)

Augments a manually labeled training set with new data from an unlabeled auxiliary set to improve recognition performance [20]. Without the cost of manual labeling such auxiliary data is often easily acquired but is not normally useful for learning.

G. Random sampling LDA

To reduce the influence of unimportant or redundant features on the variables generated by PCA, random sampling LDA was introduced. By incorporating Feature Selection for face recognition (FS_RSLDA) was introduced, in this algorithm unimportant or redundant features are removed at first, this way the obtained weak classifier is made better [21].

H. Revised Non-negative Matrix Factorization (NMF) with LDA based color face recognition

Block diagonal constraint is imposed on the base image matrix and coefficient matrix on the basis of the constraints of traditional NMF. And LDA is then implemented on factorization coefficients to fuse class information [22].

I. Layered Linear Discriminant Analysis (L-LDA)

Decrease False Acceptance Rate (FAR) by reducing the face dataset to very small size through L-LDA. It is intensive to both small subspace (SSS) and large face variations due to light or facial expressions by optimizing the separability criteria. Hence it provides significant performance gain, especially on similar face database and Small Subspace (SSS) problems [23].

VI. PERFORMANCE ANALYSIS

A. Performance Analysis of Various PCA based Algorithms

Illumination invariant face recognition based on DCT and PCA on YALE Database B gives accuracy of 94.2% [28].

As Table I shows, face recognition rate of PCA+SVM method, under small samples circumstance, is better than PCA+NN and SVM [6].

<table>
<thead>
<tr>
<th>Class</th>
<th>Training samples</th>
<th>Test samples</th>
<th>Method</th>
<th>Recognition rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>PCA+NN</td>
<td>90</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SVM</td>
<td>85.71</td>
</tr>
<tr>
<td>200 C</td>
<td>60</td>
<td>140</td>
<td>PCA+SVM</td>
<td>94.29</td>
</tr>
<tr>
<td>40 C</td>
<td>120</td>
<td>280</td>
<td>PCA+NN</td>
<td>80.36</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SVM</td>
<td>78.93</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>PCA+SVM</td>
<td>81.10</td>
</tr>
</tbody>
</table>

Table II and Table III shows, face recognition rate of I(2D)PCA is better when compared to PCA, 2DPCA, I(2D)PCA, IPCA, I(2D)PCA on YALE and ORL databases [7].

Table IV shows, Eigen-GEFeW is the best performing instance when compared with Eigenface and Eigen-GEFeS on Face Recognition Grand Challenge (FRGC) dataset.

<table>
<thead>
<tr>
<th>Methods</th>
<th>Recognition rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eigenface</td>
<td>87.14</td>
</tr>
<tr>
<td>Eigen-GEFeS</td>
<td>86.67</td>
</tr>
<tr>
<td>Eigen-GEFeW</td>
<td>91.42</td>
</tr>
</tbody>
</table>
symmetrical property of human face to enlarge the number of training samples, but also employs the weighted PCA space to improve the robustness against variance of illumination and expression [9].

### Table V. Performance Comparison Between PCA, SPCA, WPCA and SWPCA on ORL Database

<table>
<thead>
<tr>
<th>Method</th>
<th>Training samples/class</th>
<th>Recognition rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCA</td>
<td>6</td>
<td>92.50</td>
</tr>
<tr>
<td>SPCA</td>
<td>6</td>
<td>94.37</td>
</tr>
<tr>
<td>WPCA</td>
<td>6</td>
<td>94.37</td>
</tr>
<tr>
<td>SWPCA</td>
<td>6</td>
<td>96.00</td>
</tr>
</tbody>
</table>

### Table VI. Performance Comparison Between PCA, SPCA, WPCA and SWPCA on Yale Database

<table>
<thead>
<tr>
<th>Method</th>
<th>Training samples/class</th>
<th>Recognition rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCA</td>
<td>4</td>
<td>85.71</td>
</tr>
<tr>
<td>SPCA</td>
<td>4</td>
<td>88.57</td>
</tr>
<tr>
<td>WPCA</td>
<td>4</td>
<td>89.52</td>
</tr>
<tr>
<td>SWPCA</td>
<td>4</td>
<td>93.33</td>
</tr>
</tbody>
</table>

Table VII shows, best performance (99.70%) of PCA-IN classifiers. PCA-IN method outperformed all other methods [10]. Face recognition rate of Manual face localization on ORL and SHEFFIELD database consisting of 100 components is 100% [11].

In FRFT face images are transformed into FRFT domain, it uses several angles characters for classifying. Experiments on FERET database shows that FRFT provides new insights into the role that pre-processing methods play in dealing with images [12].

GNP-FDM successfully prevents the accuracy loss caused by a large number of classes in the Multiple Training Images per Person – Complicated Illumination Database (MTIP-CID). GCA reduces the overlaps in the PCA domain [13].

PCA and minimum distance classifier gives a recognition rate of 96.7% on ORL database [14].

### A. Performance Analysis of Various LDA based Algorithms

Table VIII and Table IX shows, R-LDA using PRM gives better recognition when compared to R-LDA on UMIST database, further it is observed that by taking more number of features (32), the recognition rate is maximum (97.5%) for ORL database and by considering 12 number of features in case of UMIST database the recognition rate is 98.48% [15].

### Table VIII. Performance Comparison Between R-LDA and R-LDA using PRM on Yale Database

<table>
<thead>
<tr>
<th>Number of features</th>
<th>Methods</th>
<th>Recognition rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>32</td>
<td>R-LDA</td>
<td>95</td>
</tr>
<tr>
<td>32</td>
<td>R-LDA Using PRM</td>
<td>97.5</td>
</tr>
</tbody>
</table>

### Table IX. Performance Comparison Between R-LDA and R-LDA using PRM on UMIST Database

<table>
<thead>
<tr>
<th>Number of features</th>
<th>Methods</th>
<th>Recognition rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>R-LDA</td>
<td>88.50</td>
</tr>
<tr>
<td>12</td>
<td>R-LDA Using PRM</td>
<td>98.48</td>
</tr>
</tbody>
</table>

Compared to LDA, MFDA significantly boosts the recognition performance. The accuracy for LDA is 60% compared to the 83.9% accuracy of MFDA [16].

### Table X. Performance Comparison Between 2DLDA, RM2DLDA (2x2) and RM2DLDA (4x4) on ORL Database

<table>
<thead>
<tr>
<th>Methods</th>
<th>Recognition rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2DLDA</td>
<td>95.65</td>
</tr>
<tr>
<td>Rm2DLDA(2 x 2)</td>
<td>96.65</td>
</tr>
<tr>
<td>Rm2DLDA(4 x 4)</td>
<td>97.1</td>
</tr>
</tbody>
</table>

### Table XI. Performance Comparison Between 2DLDA, RM2DLDA (2x2) and RM2DLDA (4x4) on Yale Database

<table>
<thead>
<tr>
<th>Methods</th>
<th>Recognition rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2DLDA</td>
<td>88.68</td>
</tr>
<tr>
<td>Rm2DLDA(2 x 2)</td>
<td>90.75</td>
</tr>
<tr>
<td>Rm2DLDA(4 x 4)</td>
<td>91.55</td>
</tr>
</tbody>
</table>

Table X, Table XI and Table XIII shows, Rm2DLDA gives better recognition when compared to 2DLDA on ORL, YALE and PIE databases [17].

### Table XII. Performance Comparison Between 2DLDA, RM2DLDA (2x2) and RM2DLDA (4x4) on PIE Database

<table>
<thead>
<tr>
<th>Methods</th>
<th>Recognition rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2DLDA</td>
<td>90.58</td>
</tr>
<tr>
<td>Rm2DLDA(2 x 2)</td>
<td>93.0</td>
</tr>
<tr>
<td>Rm2DLDA(4 x 4)</td>
<td>95.04</td>
</tr>
</tbody>
</table>

Table XIII, Table XIV shows, IALDA gives better recognition when compared to LDA on CMU PIE.
The recognition rate is increased from 94.12% using Fuzzy Fisherface (FLDA) to 98.125% using Fuzzy Fisherface through genetic algorithm on ORL database [19].

Experiments on ORL database, AR database and CMU PIE database show that Semi-supervised face recognition algorithm based on LDA is robust to variations in illumination, pose and expression and that it outperforms related approaches in both transductive and semi-supervised configurations [20].

RSLDA is an effective random sampling LDA method, the 1-NN classifier in the feature subspace obtained by RSLDA has better classification performance as compared to that induced by Base-LDA on AR, ORL, YALE, YALEB face datasets [21]. For ORL, the classification accuracy has an increase of 15.1% around.

Experimental results on CVL and CMU PIE databases prove the algorithm improves recognition rate effectively [22]. L-LDA is insensitive to large dataset and also small sample size and it provided 93% accuracy and reduced False Acceptance Rate (FAR) to 0.42 on BANCA face database [23].

B. Performance Comparison between PCA and LDA based Algorithms

Table XV shows, LDA gives better recognition when compared to PCA while 32 features are considered on YALE Database [15].

Table XVI shows, LDA gives better recognition when compared to PCA while 12 features are considered on UMIST Database [15].

Table XVII shows, 2DPCA gives better recognition when compared to RLDA on ORL Database [17].

Table XVIII shows, 2DPCA gives better recognition when compared to RLDA on YALEB Database [17].

Table XIX shows, RLDA gives better recognition when compared to 2DPCA on PIE Database [17].

Table XX shows, LDA gives better recognition when compared to PCA on B1 and CMU PIE Databases respectively [20].

Table XXI shows, PCA and LDA on CMU PIE Databases give the best result.

Table XXII shows, PCA and LDA on ATT, CROPPED YALE, faces94, faces95, faces96, JAFE Databases.

From Table XXII, it is evident that the best algorithm to recognize image without disturbance is PCA, because in the same recognition rate, PCA takes shorter time than LDA. But to recognize image with disturbances, LDA is better to use because it has better recognition rate [24]. In term of time taken, PCA tends to be much better than LDA, especially to recognize images with background disturbance [24].

VII. CONCLUSION

In this paper, we have analysed various current PCA based and LDA based algorithms for face recognition. This analysis is vital in developing new robust algorithms for face recognition. Among various PCA algorithms analysed, the best result was found when Manual face localization was used on ORL and SHEFFIELD database consisting of 100 components. The face recognition rate in this case was 100%. The next best was 99.70% face recognition rate using PCA- IN on ORL database. Among various LDA algorithms analysed, it was found that IALDA gives the best face recognition rate of 98.9 % when 20 test samples and 1
training sample were considered on CMU PIE Database. The next best was 98.125 % using Fuzzy Fisherface through genetic algorithm on ORL database.

REFERENCES


