CORE
Metadata, citation and similar papers at core.ac.uk
Provided by Elsevier - Publisher Connector

Available online at www.sciencedirect.com
ScienceDirect
Procedia Computer Science 54 (2015) 532 – 541
Procedia Computer Science

Eleventh International Multi-Conference on Information Processing-2015 (IMCIP-2015)

Face Recognition using R-KDA with Non-Linear SVM for Multi-View Database

Hidangmayum Saxena Devi[a,*], Romesh Laishram[b] and Dalton Meitei Thounaojam[c]

[a] *National Institute of Technology Manipur, Takyelpat Imphal 795 001, India*
[b] *Manipur Institute of Technology, Takyelpat Imphal 795 001, India*
[c] *National Institute of Technology Silchar, Cachar, Assam 795 001, India*

Abstract

This paper develops a new Face Recognition System which combines R-KDA for selecting optimal discriminant features with non-linear SVM for Recognition. Experiment results have been conducted showing the comparison of enhanced efficiency of our proposed system over R-KDA with k-nn as the similarity distance measure.

© 2015 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
Peer-review under responsibility of organizing committee of the Eleventh International Multi-Conference on Information Processing-2015 (IMCIP-2015)

Keywords: BPN (Back propagation network); Discriminant analysis (DA); Direct exact LDA (DLDA); Fisher LDA (FLDA); Generalised discriminant analysis (GDA).

1. Introduction

Face Recognition is an instance of Biometrics System. Biometrics is necessary in our day to day life, where-in a person is necessary to be recognized by his/her personal physiological characteristics which are related to the shape of the body. From the various available biometric systems available viz. Fingerprint recognition, Face recognition, Iris recognition, Retina recognition, Hand geometry, Voice recognition, Signature Recognition, Face Recognition, in particular has received a considerable attention in recent years both from the industry and the research community and also considered as one of the efficient and easiest biometric systems available due to the fact that because face can be captured or acquired from a distance without giving any information to the person intended and can be captured with ease. Different techniques has been implemented in the area of Face Recognition where focus is on the intelligent approaches like PCA, LDA, DFLD, BPN, etc. In the current trend, combinations of the existing techniques are being taken into consideration either in the feature extraction level or in the classification level for effective recognition. Face Recognition is a kind of a template matching problem in which recognition is to be performed in a high dimensionality space which consists of five main steps (shown in Fig. 1(a)): (a) Face acquisition, (b) Face detection, (c) Feature extraction, (d) Face recognition and finally (e) Output.

(a) Face acquisition:Faces can be captured either using a camera or can be taken from existing standard Database. For example ORL Face Database, AT&T Face Database, UMIST database etc. (b) Face Detection: Face detection is a

*Corresponding author. Tel.: 8794826470.
E-mail address: sanahidangmayum@gmail.com

1877-0509 © 2015 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
Peer-review under responsibility of organizing committee of the Eleventh International Multi-Conference on Information Processing-2015 (IMCIP-2015)
doi:10.1016/j.procs.2015.06.061

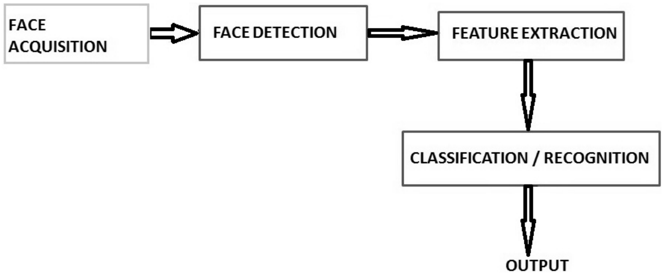

Fig. 1. (a) General model of a face recognition system.

technique that determines the locations and sizes of human faces in arbitrary images. It detects facial features. It may or may not be included in face Recognition. (c) Feature Extraction: Extraction of relevant information from a face Image is the prime task of Feature Extraction. It involves reduction in dimension, extraction of relevant features and selection of optimal features. The output of feature extraction is mainly dimensionality reduction transforming the data selecting a proper subspace in the original feature space, retaining only the relevant features. Feature Selection selects those subset of extracted features which might result in low classification error. Various existing Feature Extraction Techniques are PCA, Kernel PCA, LDA, ICA, gabor wavelet transforms etc. (d) Face Recognition: The extracted features are used for Recognition. This depends on the user application. For identification purposes, this step will be a comparison between a given picture for the subject and all the biometric templates stored on a database. If it is to be used for verification purpose, the biometric template of the claimed identity will be retrieved (either from a database or a storage medium presented by the subject) and this will be compared to given picture. A classification on various categories of Recognition algorithms are discussed in[1] which are grouped into four categories:

- Geometric/Template based approaches: Geometric based approaches or Feature-based approaches take into account the local facial features and their geometrical relationships. Template based methods perform recognition by comparing the input image with a set of templates which are constructed using any one of the statistical tools viz. PCA, LDA, SVM etc.
- Wholistic Approaches: Here, for recognition, facial features are processed as a whole not extracting relevant information.
- Appearance-based/Model-based Approaches: In appearance based approaches, feature space is derived from the image distribution using statistical techniques and the input image is compared to it. It can be classified as linear or non-linear. On the other hand, in model based approaches, a human face is modeled and the test image is fitted to the model, the face is recognised using the parameters of the fitten model. It can be 2D or 3D.
- Template or Statistical or Neural Network Approaches
  - Template Approaches: Patterns are represented by samples, models, pixels, curves or textures and recognised using a correlation or distance measure.
  - Statistical Approaches: Features represent the patterns which are then recognised using discriminant function.
  - Neural Network Approaches: Here there is various ways of representing features.
- Output: A list or the one with nearest match is given as output after performing the above mentioned steps.

Since different people have unique faces, we can recognize different faces with less difficulty. But automatic recognition by system is a bit difficult task in terms of parameters to be used for the above purpose, computational complexity as well as time complexity.

Face Recognition is one of the demanding area of research as there is still no robust technique against uncontrolled practical situations viz., pose, illumination, expression, rotation and occlusion and also due to the need of automatic recognition, surveillance systems, etc. For a better recognition system, an efficient feature extraction and classification algorithms are required.

PCA[2,3] and LDA[4] are commonly used feature extraction algorithms in any area of image processing. PCA[2] is one of the statistical methods for reducing the dimensionality of a data set while retaining the majority of the variation in the data set thereby producing optimal linear least squares decomposition of a training set. PCA[3], being an unsupervised technique, does not include label information of the data. This problem can be solved using LDA[4] or FLDA[5]. LDA[4] is one of the methods for dimension reduction as well as classification technique by maximizing the ratio of between class variance to within-class variance in any particular data set. FLDA[5] is the application of LDA to PCA face-subspace. LDA optimizes the low dimensional representation of the objects with focus on the most discriminant feature extraction. The subspace LDA method consists of two steps: first, the face image is projected from the original vector space to a face image from the original vector space to a face subspace via PCA where the subspace dimension is carefully chosen, and then LDA is applied to obtain a linear classifier in the space dimension enables us to generate class-separable features via LDA from the full space representation. Hence, we are able to solve full the generalization/over fitting problem when we perform face recognition on a large face dataset but with very few training face images available per testing person.

2DLDA[6], DLDA[7], DF-LDA[8], F-LDA[9], GDA[10], KDDA[11] are some variants of LDA which have been used in Face Recognition. In LDA[4], a fisher space is established from the training samples in the training phase and the training faces are projected onto the same subspace. Eigen decomposition on the scatter matrices is applied to compute optimal projection. It fails when all the scatter matrices are singular. In 2DLDA[6], LDA is applied directly on the data matrix without transforming it into vector overcoming the singularity problem. In direct, exact LDA (D-LDA)[7], Fishers criterion is optimized directly without reducing the dimension of the data discarding the null space of between class scatter matrix by first diagonalizing between class scatter matrix and then within class scatter matrix is diagonalized. But the discarded null space may contain significant discriminatory information. DF-LDA[8] is a combination of D-LDA[7] and F-LDA[9] where D-LDA[7] is used to obtain a low dimensional SSS free subspace and subsequent F-LDA[9] is used to re-orient the SSS-free subspace resulting in a set of optimal discriminant features for face representation to enhance the discriminatory power of the obtained D-LDA feature space. In GDA[10] (also known as K-LDA), kernel method is applied to LDA thereby maximizing the inter-classes inertia and minimizing the intra-classes inertia. Here, the standard Fisher's criterion is maximized leading to SSS problem due to the high dimensionality of feature space. The SSS problem is solved removing the null space of within-class scatter matrix, losing some of the significant discriminatory information. To obtain a set of optimal discriminant basis vectors in the feature space and also to solve the above problems mentioned above in GDA[10], a new algorithm called KDDA[11] is proposed, in which a kernel method is applied to a variant of D-LDA (JD-LDA) for dealing with non-linearity of the face pattern's distribution which cannot be dealt by PCA/LDA.

SVM[12] finds a hyperplane that separates the largest possible fraction of points of the same class on the same side maximising the distance from either class to the hyperplane. To deal with the non-linearity problems, data are transformed into high-dimensional feature space using kernel function. The hyperplane is found out using eqn. no (1)

$$y(x) = \sum_{i=1}^{N} \alpha_i y^i K(x^i, x) + b \tag{1}$$

where $K(x^i, x)$ is the kernel function. The idea behind the SVM classifier is to model a given training set with a corresponding group vector and classifies a given test set using an SVM classifier according to a one vs. all relation transforming multi-class problem into two-class problem. For linearly non-separable data, SVM maps the input to a higher dimensional feature space where a linear hyperplane can be found. Although there is no warranty that a linear solution will always exist in the higher dimensional space, it is able to find effective solutions in practice. SVM has been used in classification in many face recognition systems.

In our Face Recognition System, R-KDA[13] is used for feature extraction and non-linear SVM, for classification. Section 2 and Section 3 discuss the motivation behind the proposed system and a brief Literature Survey on various Feature Extraction and Face Recognition Techniques respectively. The Background Work and the Proposed Face recognition System describing the Feature Extraction and Recognition Phase are explained in Section 4 and Section 5 respectively. Section 6 discusses the Implementation Issues of the proposed system.

2. Motivation

As compared with other biometrics system viz.fingerprint, palm print and iris etc., Face Recognition has distinct advantages because of its non-contact process as face images can be captured from a distance without touching the person being identified, and the identification does not require interacting with the person and it serves the crime deterrent purpose because face images that have been recorded and archived can later help identify a person.

The various existing Face Recognition System cannot handle effectively the various uncontrollable conditions viz. illumination, expression, poses, occlusion, rotation, scaling and translation etc, which are discussed in brief below. These factors are considered as hindrances to achieving 100% recognition rate. Maximum of these factors should be taken into consideration for achieving an accurate Face Recognition System.

(a) Illumination: Illumination occurs due to variations caused by various lighting environments which lead to large differences in appearance and hence could affect the recognition system.
(b) Pose: Pose variation can happen due to persons movement while taking the shot or cameras angle. Face Recognition system works well only when the frontal images are involved. So, when such pose variations occur, face recognition system doesn't work correctly.
(c) Expression: Different facial expressions results in not only the spatial location change but also change in the shape of facial feature.
(d) Rotation, Scaling and Translation: These can be caused by variations in image acquisition process leading to difficulty in both face detection and recognition.
(e) Occlusion: Face Recognition process can be hindered if some or whole of the face is covered by another different object, resulting in unable to extract the relevant features which is needed for successful Recognition.

The objective of this Research is to create a Face Recognition System invariant to multi-view images (poses).

So, in order to perform an effective Face Recognition against poses, R-KDA is used with SVM as Regularized KDA (R-KDA)can extract optimal set of discriminant features and SVM can classifier better over k-nn classifier in pose variations images (multi-view database) than previous linear and other non-linear methods. Hence, R-KDA is used with nonlinear-SVM classifier for multi-view Face Recognition.

3. Literature Survey on various Statistical Feature Extraction and Recognition Techniques in Face Recognition

Literature Survey on the various Feature Extraction and Face Recognition techniques like PCA, KPCA, DA and LDA involved in the area of Face Recognition is discussed in this section.

3.1 Principal component analysis

PCA is one of the statistical method for reducing the dimensionality of a data set while retaining the majority of the variation in the data set thereby producing optimal linear least squares decomposition of a training set. Mathematically, it can be explained as below[2]:

Let there be N images with n pixels, then the entire data set is of Nn order, where each row represents an image of the data set.

The mean of the data is found out by averaging the columns of the data matrix and then the mean image is subtracted from each image of the dataset to create the mean centered data vector say U.

The covariance matrix is calculated as:

$$\sum = \frac{U^T U}{N-1} \tag{2}$$

Eigen values and Eigen vectors are calculated from the covariance matrix. Eigen vectors are arranged in such a manner that the Eigen values are in the decreasing order denoted by E. So for any given image, only a few of the highest Eigen vectors are considered thus reducing the dimension. PCA is given by

$$PCA = \{\mathrm{Int}_i - m\} x E \tag{3}$$

where, Int_i is the intensity values of each pixel of the original image (I) and m is the mean of the corresponding image.

Turk *et al.*,[2] presented a face detection and identification method using eigenfaces in which face images are projected into a feature space (face space), characterizing an individual face by the weighted sum of the eigenface features. Eigenfaces are nothing but the principal components of the initial training set of face images.

Kim[3] discussed PCA for constructing 1-D vector of pixels from 2-D facial image for face recognition. After the eigenfaces are computed, face identification is done computing the distance between the stored faces and the eigenface.

Chung *et al.*,[14] proposed a new PCA based face recognition method in which robust facial features are represented using Gabor features, which are again transformed into Eigenspace using PCA for classification. Gabor filters handle illumination and pose variations in original face images are handled during classification.

Moon *et al.*,[15] proposed a generic modular PCA which consists of normalization and PCA projection. Recognition is obtained using nearest neighbor classifier. The performance is improved by normalization.

Khan *et al.*,[16] applied PCA to transform directional images into Eigenspace for increasing recognition accuracy. Using directional filter bank, directional images are created from original images.

Scholkopf *et al.*,[17] used KPCA in which dot product matrix $K_{ij} = (k(x_i, x_j))_{ij}$ is computed and diagonalised to calculate the eigenvalues and eigenvectors of the data matrix. Projections onto the eigenvectors provide the principal components corresponding to the kernel k. A linear SVM is trained for the object classification purpose.

Xie *et al.*,[18] used doubly non-linear mapping KPCA to reduce the effect of feature variations due to illuminations, expression and perspective disturbance. Facial features are extracted using gabor wavelets.

Welling[5] discussed that in KPCA, if the centered kernel can be computed in terms of the non-centered kernel in the high-dimensional space, no other unnecessary features need to be accessed.

3.2 DA and LDA

Balakrishnama *et al.*,[19] discussed mathematical steps of LDA and implemented as a classification technique mentioned below. They also discuss different approaches to LDA viz., class dependent and class independent.

Mathematical steps of LDA: The data set set_j and the test vectors are formulated. Mean of each data set u_j where $j = 1, 2, \ldots, n$ and the mean μ_{mean} of the entire data set is computed as:

$$\mu_{\text{mean}} = \sum (p_j x u_j) \tag{4}$$

where, p_js are the a priori probabilities of the classes. Within class scatter (S_w) is the expected covariance of each of the classes, given as

$$S_w = \sum p_j x \text{cov } j \tag{5}$$

where cov $j = (x_j - \mu_j)(x_j - \mu_j)^T$.

Between class scatter (R_b) is the covariance of data set whose members are the mean vectors of each class and is given by

$$R_b = \sum (\mu_j - \mu_{\text{mean}})(\mu_j - \mu_{\text{mean}})^T \tag{6}$$

The optimizing criterion is the ratio of Rb to Sw and the solution obtained by maximising this criterion defines the axes of the transformed space. The optimizing criteria for class dependent type is given by

$$\text{criterion } j = \text{inv}(\text{cov } j) x Rb \tag{7}$$

and, class independent type is

$$\text{criterion} = \text{inv}(Sw) Rb \tag{8}$$

The transformations are found as the eigenvector matrix transformed$_j$ defined by eqn. no. (7) and (8).

The class dependent LDA transform is given by

$$\text{transformed}_{\text{set}_j} = (\text{transformed}_j)^T \text{set}_j \tag{9}$$

and the class independent LDA is given by

$$\text{transformed}_{\text{set}} = (\text{transformed}_{\text{spec}})^T \, \text{data}_{\text{set}}^T \tag{10}$$

Similarly, test vectors are transformed. Once the transformations are completed using the LDA transforms, Euclidean distance is calculated from each class mean as:

$$\text{dist}_n = (\text{transformnspec})^T x - \mu_n \text{trans} \tag{11}$$

where, $\mu_n$trans is the mean of the transformed data set and x is the test vector, which is used for classification.

Chelali *et al.*,[4] presented LDA based face recognition wherein a fisher space is established from the training samples in the training phase and the training faces are projected onto the same subspace. Eigen decomposition on the scatter matrices is applied to compute optimal projection. In the classification phase, Euclidean distance is used to find the similarity measure. It is also reported that recognition rate is poor under varying head tilts and illumination. Many LDA-based algorithms suffer from the so-called small sample size (SSS) problem which exists in high-dimensional pattern recognition tasks.

Ye *et al.*,[6] proposed 2DLDA for face recognition, in which each datum is represented as a matrix and the collection of data is represented as a collection of matrices and a combination of 2DLDA and LDA is also studied. It is also concluded in this paper that 2DLDA implicitly avoids the singularity problem and both the above methods have distinctly lower time and space complexity as well as higher classification accuracy than a combination of PCA and LDA.

Yu *et al.*,[7] proposed a direct, exact LDA (D-LDA), which optimizes Fishers criterion directly without reducing the dimension of the data. It discards the null space of between class scatter matrix by first diagonalizing between class scatter matrix and then within class scatter matrix is diagonalized.

Lu *et al.*,[8] proposed DF-LDA, which is a combination of D-LDA and F-LDA for feature representation for face recognition and classification is done using nearest neighbor classifier. A new variant of D-LDA is introduced which utilizes modified Fishers criterion and a weighing function so that a low dimensional SSS free subspace is obtained and subsequent F-LDA step can be applied to re-orient the SSS-free subspace resulting in a set of optimal discriminant features for face representation to enhance the discriminatory power of the obtained D-LDA feature space.

Baudat *et al.*,[10] proposed GDA, which deals with non-linear discriminant analysis using kernel function operator. The input space is mapped into another feature space, which are non-linearly related to the input space. GDA tries to maximize the inter-classes inertia and minimize the intra-classes inertia and can be used for supervised and nonlinear problem for feature extraction and for classification. It is discussed that using GDA, a reduced number of discriminant coordinate that are optimal for separating the groups can be found out. With two such coordinates a classification map that partitions the reduced space into regions is constructed.

Lu *et al.*,[11] introduced a kernel machine-based DA method, which is a combination of GDA and D-LDA, an improvement over DLDA by introducing kernel features in which the optimal discriminant features are exactly extracted from both inside and outside of the within class scatter matrix in the feature spaces null space.

These are a brief survey on some of the existing Feature Extraction and Face Recognition Techniques.

4. Background Work

LDA is one of the methods for dimension reduction as well as classification technique. From a training set Z, each containing C classes with

$$\{Z_i\}_{i=1}^{c} \tag{12}$$

a number of localized face images, $\{Z_{ij}\}$ (each of size J), the total number of face images, LDA finds a set of $M \ll J$ feature basis vectors, $\{\psi_m\}_{m=1}^{M}$ maximizing the ratio of between-class scatter matrix to within-class scatter matrix as

$$\psi = \frac{S_b}{S_w} \tag{13}$$

where 'Sb' and 'Sw' are the between class scatter matrix and within class scatter matrix respectively. Modified Fisher criterion is introduced in D-LDA to solve SSS problem given in eqn. no. (17).

$$\psi = \text{argmax}\frac{\psi^T S_b \psi}{\psi^T S_w \psi} \tag{14}$$

Face data, being high dimensional and number of samples per class is less, SSS problem arises leading high variance in the sample-based estimation of Sb and Sw making the system ill-posed or poorly posed. Other variants of LDA do exist such as DF-LDA which uses weighting function into the Sb but ultimately leads to misclassification in the output space. The parameter is again modified in KDDA[11] approach using eqn. no. (18).

$$\psi = \text{argmax}\frac{|\psi^T S_b \psi|}{|\psi^T S_b \psi + \psi^T S_w \psi|} \tag{15}$$

This method even cannot solve the SSS problem effectively. To effectively solve the SSS problem and to reduce high computational complexity, a new Face Recognition approach is developed introducing a regularisation parameter and subspace decomposition techniques, known as R-KDA[13] using eqn. no. (19)

$$\psi = \text{argmax}\frac{|\psi^T S_b \psi|}{|\eta\{\psi^T S_b \psi\} + \{\psi^T S_w \psi\}|} \tag{16}$$

This regularization strategy provides a balance between the variance and the bias in the sample based estimates solving the SSS problem. Non-linear models can handle the various variations in face patterns, so R-KDA[13] is implemented using a kernel function, where F is the high-dimensional feature space in which the distribution of the mapped data is linearized and simplified.

5. Proposed Face Recognition System

The proposed Face Recognition System consists of two phases: Feature Extraction Phase and Face Recognition Phase in which R-KDA[13] is used for feature extraction and non-linear SVM[12] for Recognition.

5.1 Feature extraction phase

5.1.1 Steps involved in feature extraction using R-KDA (For Training Set)

Input: Image Training set, number of training samples per class, RBF kernel parameter
Output: R-KDA subspace

- The kernel matrix using RBF kernel for the training set is calculated.
- Eigen-analysis of between-class scatter matrix and within-class scatter matrix is performed.

Regularised eigenvalues of within class scatter matrix to find the kernel discriminant subspace. (R-KDA subspace).

5.1.2 Steps involved in Feature Extraction using R-KDA (For Testing Set)

It projects the test samples to the kernel discriminant subspace, R-KDA subspace obtained from the training set.
Input: Testing data, training examples, kernel discriminant space, Gaussian variance for RBF kernel.
Output: R-KDA Subspace for Testing Set

- Calculate the size of the test set (m) and size of the learning set (n).
- Create an array of size $m * n$.
- Calculate the kernel matrix using RBF kernel for the test set.
- Calculate the projection using R-KDA subspace with the Kernel matrix.

5.2 Face Recognition Phase

Face Recognition is performed using SVM classifier in which SVM finds an optimal hyperplane that separates the largest possible fraction of points of the same class on the same side maximizing the distance from either class to the hyperplane. Determination of the optimal hyperplane is a constrained optimization problem and can be solved using quadratic programming techniques. The discriminant hyperplane is defined using eqn. no. (1).

One Against-All strategy is employed in Face Recognition using SVM for transforming multiclass problem to two-class problem. SVM finds the optimal linear hyperplane such that the expected classification error for future test samples is minimized, i.e., good generalization performance is achieved.

Training set

$$T = \{(x^i, y^i), x^i \epsilon R^n; y^i \epsilon \{-1, +1\}; i = 1, 2, \ldots, N\} \quad (17)$$

is transformed to a series of

$$D_k = \{(x^t, y_k^t \epsilon \{-1, +1\}\} \quad (18)$$

where,

$$y_k^i = +1 if y^i = k \quad (19)$$

$$y_k^i \neq -1 if y^i = k \quad (20)$$

Discriminant functions are computed corresponding to D_k using eqn. no. (18). Recognition of class for the particular input x is done using eqn. no. (21)

$$k = argmax(f_x(k)) \quad (21)$$

6. Implementation Issues

6.1 Database used

The Sheffield Face Database (previously UMIST Face Database) consists of 564 images of 20 people. Each covers a range of poses from profile to frontal views. Subjects cover a range of race/sex/appearance. Each subject exists in their own directory. The files are all in PGM format, approximately 220 × 220 pixels in 256 shades of grey. Out of these total images, 10 different poses each of 18 different subjects are taken for our experimental purpose. Samples of two different persons are given in Fig. 2(a) and (b). Six images of different poses are used for training and the rest for testing.

6.2 Performance analysis

In this section, the Performance of R-KDA is compared with the Performance of KPCA with MAHCOS distance as the similarity measure, in which R-KDA performs better than KPCA as shown in Fig. 3 and the Performance Analysis of R-KDA with non-linear SVM is presented as in Fig. 4. The performance of the algorithm is given in terms of correct recognition rate (crr) against the change in RBF kernel. The kernel parameter for RBF is the scale value. The Correct Recognition Rate (crr) is plotted against RBF kernel within the range 1×10^7 to 15×10^7, equally subdivided with a space of 5 intervals each. To enhance the accuracy of the assessment, ten runs of such a partition are executed and

Fig. 2. (a) Sample for different poses of one subject contained in the Sheffield face multi-view database used in our experiment; (b) Sample for different poses of another subject contained in the Sheffield face multi-view database used in our experiment.

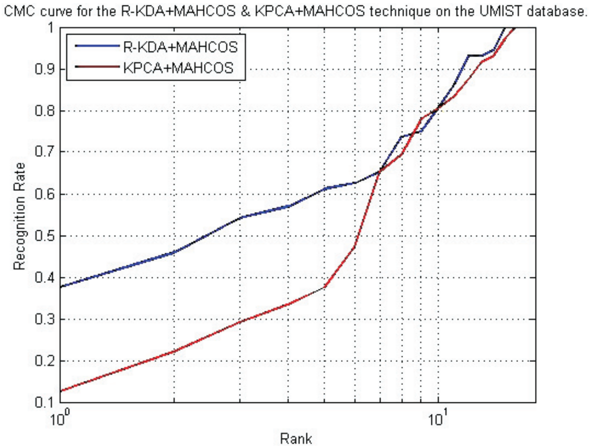

Fig. 3. Performance comparison of R-KDA over KPCA with MAHCOS as similarity measure.

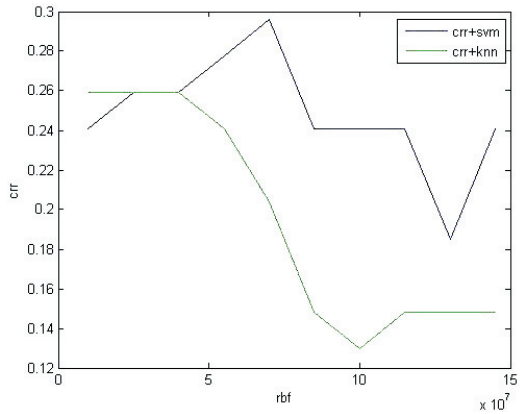

Fig. 4. Performance Comparison of our System over R-KDA with k-nn.

the results have been averaged over these ten runs. To show the robustness of the proposed algorithm, a comparative study is carried out with the more commonly used k-nn classifier. It can be clearly observed from the analysis result (Fig. 4) that the combination of R-KDA with non linear SVM outperforms the traditional k-nn classifier. The best correct recognition rate achieved with proposed algorithm is approximately 30% and the RBF function for the highest crr is found at 7×10^7. The reason for low crr is mainly due to the fact that we are considering multi-view faces in our study. It is well known that the correct recognition rate for any algorithm is quite low in the case of multi-view faces. But the benchmark for face recognition algorithm should be based on multi-view faces.

7. Conclusion

In this paper we have proposed the combination of R-KDA with non-linear SVM for face recognition in a multi-view face database. R-KDA is used for selecting optimal discriminant features and non-linear SVM is used for classification. The experiment carried out using UMIST multi-view face data set shows that our proposed system outperforms R-KDA with k-nn as the similarity distance measure, in computing the correct recognition rate calculated against RBF kernel. The best correct recognition rate is low but for multi-view faces. The algorithm may further be tested in other standard data set to test its robustness which is considered for our future studies.

References

[1] De Carrera and Proyecto Fin, Face Recognition Algorithms, *Master's Thesis in Computer Science*, pp. 26–28, (2010).

[2] M. Turk and A. Pentland, Eigenfaces for Recognition, *Journal of Cognitive Neuroscience*, vol. 3(1), pp. 71–86, (1991).

[3] K. Kim, Face Recognition using Principle Component Analysis, In *International Conference on Computer Vision and Pattern Recognition*, pp. 586-591, (1999).

[4] F. Z. Chelali, A. Djeradi and R. Djeradi, Linear Discriminant Analysis for Face Recognition, In *IEEE International Conference on Multimedia Computing and Systems ICMCS*, pp. 1–10, (2009).

[5] M. Welling, Fisher Linear Discriminant Analysis, Department of Computer Science, University of Toronto, (2005).

[6] J. Ye, R. Janardan and Q. Li, Two-dimensional Linear Discriminant Analysis, *Advances in Neural Information Processing Systems*, pp. 1569–1576, (2004).

[7] H. Yu and J. Yang, A Direct LDA Algorithm for High-Dimensional Datawith Application to Face Recognition, *Pattern Recognition*, vol. 34(10), pp. 2067–2070, (2001).

[8] J. Lu, K. N. Plataniotis and A. N. Venetsanopoulos, Face Recognition using LDA-Based Algorithms, *IEEE Transactions on Neural Networks*, vol. 14(1), pp. 195–200, (2003).

[9] R. Lotlikar and R. Kothari, Fractional-Step Dimensionality Reduction, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22(6), pp. 623–627, June (2000).

[10] G. Baudat and F. Anouar, Generalized Discriminant Analysis using a Kernel Approach, *Neural Computation*, vol. 12(10), pp. 2385–2404, (2000).

[11] Lu, Juwei, Konstantinos N. Plataniotis and Anastasios N. Venetsanopoulos, Face Recognition using Kernel Direct Discriminant Analysis Algorithms, *IEEE Transactions on Neural Networks*, vol. 14(1), pp. 117–126, (2003).

[12] C. J. Burges, A Tutorial on Support Vector Machines for Pattern Recognition, *Data Mining and Knowledge Discovery*, vol. 2(2), pp. 121–167, (1998).

[13] J. Lu, K. N. Plataniotis and A. N. Venetsanopoulos, Kernel Discriminant Learning with Application to Face Recognition, In *Support Vector Machines: Theory and Applications*, pp. 275–296, (2005).

[14] Chung, Ki-Chung, Seok Cheol Kee and Sang Ryong Kim, Face Recognition using Principal Component Analysis of Gabor Filter Responses, *Recognition, Analysis, and Tracking of Faces and Gestures in Real-Time Systems*, (1999).

[15] Moon, Hyeonjoon and P. Jonathon Phillips, Computational and Performance Aspects of PCA-based Face-Recognition Algorithms, Perception-London, vol. 30(3), pp. 303–322, (2001).

[16] Khan, Mohammad AU, *et al.*, Improved PCA based Face Recognition using Directional Filter Bank, In *Proceedings of IEEE Multitopic Conference*, (2004).

[17] B. Schlkopf, A. Smola and K. R. Mller, Nonlinear Component Analysis as a Kernel Eigenvalue Problem, *Technical Report on Neural Computation*, 44, (1996).

[18] Xie, Xudong and Kin-Man Lam, Kernel PCA with Doubly Nonlinear Mapping for Face Recognition, In *IEEE International Symposium on Intelligent Signal Processing and Communication Systems*, (2005).

[19] Balakrishnama, Suresh and Aravind Ganapathiraju, Linear Discriminant Analysis-A Brief Tutorial, *Institute for Signal and Information Processing*, (1998).