

Face recognition invariant to partial occlusions

Azeem Aisha, Sharif Muhammad, Shah Jamal Hussain, Raza Mudassar

Department of Computer Sciences,

COMSATS Institute of Information Technology Wah Cantt., 47040, Pakistan

[e-mail: aishaazeem19@gmail.com, muhammadsharifmalik@yahoo.com, jamalhussainshah@gmail.com,
mudassarkazmi@yahoo.com

*Corresponding author: Raza Mudassar

*Received December 9, 2013; revised February 8, 2014; revised March 14, 2014; accepted March 27, 2014
; published July 29, 2014*

Abstract

Face recognition is considered a complex biometrics in the field of image processing mainly due to the constraints imposed by variation in the appearance of facial images. These variations in appearance are affected by differences in expressions and/or occlusions (sunglasses, scarf etc.). This paper discusses incremental Kernel Fisher Discriminate Analysis on sub-classes for dealing with partial occlusions and variant expressions. This framework focuses on the division of classes into fixed size sub-classes for effective feature extraction. For this purpose, it modifies the traditional Linear Discriminant Analysis into incremental approach in the kernel space. Experiments are performed on AR, ORL, Yale B and MIT-CBCL face databases. The results show a significant improvement in face recognition.

Keywords: Feature extraction, sub-classes, partial occlusion, linear discriminate analysis, incremental approach.

1. Introduction

As it is known, Biometrics is an active field of research in image processing. Among other features in the field, iris, face and fingerprints recognition are specifically used for security purposes. Face recognition is most commonly used in real time applications. The main emphasis of such systems is the accurate identification and recognition of a person. The robustness of such systems depends on the extent to which changes in conditions of light or expression or even in the presence of partial occlusion can be measured. However, there is hardly any study which may have overcome these limitations satisfactorily. Several researchers have proposed various techniques for recognition of faces under conditions in which Eigenfaces [1] are variant to aforementioned factors. Neural networks [2] is also an important classification and feature extraction mechanism but unfortunately as the size of the database increases, its computational cost increases and recognition rate declines accordingly. Since face recognition are broadly classified into linear and non-linear methods. In the linear method, there are two most commonly used techniques viz. Principal Component Analysis (PCA) [1], [3] and Linear Discriminant Analysis (LDA) [4]. The Main advantage of these techniques is that they convert the high dimensional image into low dimensional image space where the features can be linearly separable. The non-linear methods include the kernel version of above-mentioned methods i.e., Kernel Discriminating Analysis (KDA) [5], Kernel Principal Component Analysis (KPCA) [6], Laplacian faces [7] and many others. These methods are further categorized according to the nature of feature extraction processes. Some researchers have used holistic methods involving the whole face whereas others simply involve local features such as eyes, nose and mouth regions. Yet there are others who prefer to use the variants of both approaches. Nevertheless face recognition still remains a complex issue which presents numerous challenges when it comes to real time applications. In view of the above discussion this study specifically examines this problem and attempts to deal with a particular aspect of partial occlusion. As noted above the question of accurate measurement of facial expression is subject to many limitations. This is highlighted in the following Fig. 1.



1. Fig. 1. Presence of natural and synthetic occlusion in upper and lower part of face

Fig. 1 shows the presence of partial occlusion in face where the lower and upper part of face is covered by natural and synthetic occlusions.

The next Section briefly reviews the literature relating to partial occlusions. It proposes a framework for face recognition which is provided in Section 3. Section 4 discusses the various experiments conducted to check the applicability of the proposed framework. The broad conclusions of the study are given in the final section.

2. A Brief Review of the Literature

Several techniques have been proposed by researchers that particularly deal with partial occlusions in face recognition systems which have been used as a basis for classification. Much of the work relates to part based methods, which includes techniques like 2D-PCA [8]. These methods initially identify the occlusions using k-NN and 1-NN classifiers and then check fractional similarities by excluding the occluded portions. In particular, the fractional similarity matrix based Self-Organizing Maps (SOM) is used [9] for partial distance and determines the nearest neighbor and the image having the least distance is considered as recognized image.

Some researchers have used modification of Support Vector Machines (SVM) named Partial SVM [10]. Still others make use of Local Gabor Binary Patterns (LGBP) [11] on the face images and change them into multi-scale images. Local Salient (LS-ICA) [12] method introduced ICA, is used to calculate the local features of face images. However, algorithm in [13] uses the expansions of Posterior Union Model (PUM) and recognition is determined on the basis of single sample per class. Apart from this, Lophoscopic PCA [14] based technique is used for partially occluded images and different expressions. Undoubtedly, this method demonstrates significant improvement in its performance but escalates the computational cost accordingly. A Selective Local Non-Negative Matrix Factorization (S-LNMF) algorithm for occlusion is provided in [15]. Here, the face is partitioned into non-overlapping fragments and on each fragment PCA is implemented and after that, 1-NN is used to detect the differences. In [16] Partitioned Iterated Function System (PIFS) based face recognition is proposed and the algorithms based on it computes self-similarities among the PIFS based local features of the image. Yet another method works with the Support vector machine (SVM) with Gaussian summation kernel [17] to handle local distortion. There are studies which use probabilistic approaches to cater the issue as mentioned in [18]. In all, each method has its own boundaries and none of them appears to work for all type of problems.

While dealing particularly with occlusion, several researchers have worked in incremental learning of subspace, i.e. incremental Principal Component Analysis (iPCA) [19] and its kernel versions. However, the evidence on incremental LDA is limited perhaps due to its singularity constraint. In order to overcome this, several studies have used neural networks in combination with LDA [20] and QR decomposition to deal with this singularity constraint [21] or Generalized Singular Value Decomposition (GSVD) [22]. This research work provides the basis for the reformulation in computing the kernel version of incremental LDA named iKFDA which can be used for recognizing partially occluded face images. In this case, each set of class belonging to different subjects is divided into fixed size sub-classes. A detailed illustration is described in Section 3 below.

3. Proposed Work

The facial feature extraction and classification is considered as a complex task in the face recognition systems even if the images are taken under controlled settings. However, this task becomes more complex when the appearance is affected due to expressions or partial occlusion. Hence, much of the work on the subject is directed to tackle feature extraction process in partially occluded or expression variant images. This section attempts to provide a framework for extraction of features in partially occluded (sunglass/scarf) images. For this

purpose, an incremental learning is introduced in the non-linear i.e., kernel version of Fisher's Discriminate Analysis, named as iKFDA. A linear representation of incremental LDA has been proposed in place of incremental PCA [23] nevertheless, the application of LDA is challenging due to singularity constraint imposed by Small Sample Size problem (SSS). This is especially the case where the input samples are small in amount as compared to the sample dimensions [24] which renders the calculation of inverse of within-class scatter S_w^{-1} virtually impossible. A detailed illustration of these issues have been discussed in the following Section.

3.1 Incremental Kernel Fisher Discriminate Analysis (iKFDA)

As its is known, facial feature extraction is a crucial phase in face recognition applications. The extracted features are highly non-linear in nature therefore, it is necessary to device techniques that can extract non-linear facial patterns. On the contrary Kernel methods have the capability to overcome this difficulty by using inner products. It projects the features to high dimensional spaces, making it feasible to draw a line of separation between the overlapping features.

One of the most important objectives of this study is to transform Kernel methods presented so far from Linear Discriminate Analysis (LDA) to kernel LDA in order to discriminate different patterns of classes in high dimensional spaces. A kernel is a function k that satisfies $k(\phi(x), \phi(y))$ where $x, y \in \mathfrak{R}$ and ϕ is a mapping from \mathfrak{R} to an inner product space f i.e. $\phi: \mathfrak{R} \rightarrow f$ by assuming that Gaussian kernel is a kernel function for this study. When a set of features vectors are given, the kernel methods formulates the Gram matrix. This matrix contains the basic relationship between the feature points as well as it satisfies the property of being semi definite.

In view of the above findings, we consider the Eigen-value problem as:

$$\lambda S_w^\phi U^\phi = S_B^\phi U^\phi, \quad (1)$$

It may be noted that S_B and S_w variances in feature space is given by:

$$S_B^\phi = \sum_{i=1}^c n_i (m_i^\phi - M^\phi)(m_i^\phi - M^\phi)^T, \quad (2)$$

$$S_w^\phi = \sum_{i=1}^c \sum_{j=1}^{n_i} (\phi(X_{ij}) - M^\phi)(\phi(X_{ij}) - M^\phi)^T. \quad (3)$$

From the above equations c, n_i, m_i^ϕ, M^ϕ are the number of classes, samples in class i , local class mean, and global class mean respectively. With these formulations Eigen-vectors will lie in the high dimensional feature space as:

$$U^\phi = \sum_{i=1}^c \sum_{j=1}^{n_i} \alpha_{ij} \phi(X_{ij}). \quad (4)$$

The Fisher's Criterion in kernel space $U \in f$ is to find argument that maximizes the following:

$$J(U) = \frac{|U^T S_B^\phi U|}{|U^T S_W^\phi U|}. \quad (5)$$

In the application of above methodology in face recognition scenario containing partially occluded images, we consider n gray-scale frontal images belonging to set $A = \{a_1, \dots, a_n\} \in \mathfrak{R}^{p \times q}$. These images are aligned according to the eye coordinates as given by the relevant databases. Due to the complex nature of patterns in facial images especially when the images are occluded, it gives rise to severe non-linearity. This non-linear nature of occluded images affects adversely the recognition performance. In order to avoid the unwanted recognition results, each facial image of set A is manually partitioned into three horizontal strips of fixed size known as sub-classes. These sub-classes corresponds to local facial features eyes, mouth and nose regions viz one sub-class(horizontal strip) containing eyes, second sub-class for the nose and third sub-class for the mouth; given by:

$$\begin{aligned} L_1 &= [X_1^1 X_1^2 \cdots X_1^n] \\ L_2 &= [X_2^1 X_2^2 \cdots X_2^n], \\ L_3 &= [X_3^1 X_3^2 \cdots X_3^n] \end{aligned} \quad (6)$$

From the above equation X_j^i is the sub-class of any image a_i of i th class in set A , where $j = 1, 2, 3$. corresponds to three horizontal partitions or sub-classes and $i = 1, \dots, n$ represents the sub-classes of n images in set A . L_1, L_2, L_3 are the matrix notations for the three sub-class divisions with dimension $l \times s$. Placing the extracted sub-classes in a dataset L yields:

$$L = \begin{bmatrix} X_1^1 X_1^2 \cdots X_1^n \\ X_2^1 X_2^2 \cdots X_2^n \\ X_3^1 X_3^2 \cdots X_3^n \end{bmatrix} \Rightarrow L = \begin{bmatrix} L_1 \\ L_2 \\ L_3 \end{bmatrix}. \quad (7)$$

This arrangement is used for calculating local and global sub-class means which is represented by the symbol m^ϕ, M^ϕ respectively. Global sub-class mean is the mean of aggregate sub-classes in dataset L and local sub-class mean is the mean of each sub-class in L_1, L_2, L_3 i.e. global sub-class is given as:

$$M^\phi = \frac{1}{n} \sum_{i=1}^n \phi(X_i^c). \quad (8)$$

In this equation, n is the total number of samples in dataset L , c are the sub-classes i.e. $c = 1, 2, 3$. and local sub-class mean is given as:

$$m_i^\phi = \frac{1}{n_i} \sum_{j=1}^{n_i} X_j^i. \quad (9)$$

Having derived the values of global and local sub-class means we now find the between and within-class scatters matrices as computed earlier. Total scatter matrix S_T^ϕ can also be calculated as:

$$S_T^\phi = \frac{1}{n} \sum_c \sum_{i=1}^n X_i^c X_i^{cT} . \quad (10)$$

Total scatter matrix is a combination of S_B^ϕ and S_W^ϕ :

$$S_T^\phi = S_B^\phi + S_W^\phi . \quad (11)$$

Rearranging the above equation and placing the value of S_W^ϕ in Fisher's criterion equations as:

$$J(U) = \frac{|U^T S_B^\phi U|}{|U^T (S_T^\phi - S_B^\phi) U|} . \quad (12)$$

This formulation is used to project the sub-classes in kernel LDA space and to introduce incremental learning in feature space. In traditional Kernel LDA each time a new image is presented, inner product is computed with training set patterns closer in distance with the test pattern. In order to avoid the inner product computation simply update S_W^ϕ , S_B^ϕ and M^ϕ global sub-class mean as:

$$M^\phi = \frac{nM^\phi + a}{n+1} . \quad (13)$$

where a , is the new image added among the existing images, n is the total number of samples in dataset L i.e. total images are incremented on arrival of new image. If the new image is not contained in the dataset L , in that case S_B^ϕ is given as:

$$S_B^\phi = S_B^\phi + (a - M^\phi)(a - M^\phi)^T , \quad (14)$$

Provided the new image is already a member of the training set then update S_B^ϕ as:

$$S_B^\phi = S_B^\phi + (n_j + 1)(m_j^\phi + M^\phi)(m_j^\phi + M^\phi)^T . \quad (15)$$

In the above formulation increment n_j which belongs to j th sub-class i.e. $\forall a \in j$ th class.

Likewise, S_W^ϕ can be updated as:

$$S_W^\phi = S_W^\phi + \sum_{i=1}^n \sum_{x \in a_i} \frac{n_j}{n_j + 1} (\phi(X_i) - M^\phi)(\phi(X_i) - M^\phi)^T . \quad (16)$$

Once the subspace is updated with new subjects, distance between the subclasses is calculated to provide a compact classification.

The purpose of the following section is to calculate the distance for similarity by using the distance measures.

3.2. Computing the Distance for Similarity

After the image has been localized and projected on the non-linear kernel space, the distance d between the corresponding sub-classes of training face a_i and probe face b is computed. Let the set of test images be represented as:

$$B = \{b_1, \dots, b_m\} \in \mathfrak{R}^{p \times q} . \quad (17)$$

Here distance between each sub-class of the test image X_j^i and the mean of equivalent training image sub-class m_i is computed which is shown by the following matrix representation:

$$d(X_j^i, m_i) = \begin{bmatrix} d(X_1^1, m_1), d(X_1^2, m_2), \dots, d(X_1^n, m_n) \\ d(X_2^1, m_1), d(X_2^2, m_2), \dots, d(X_2^n, m_n) \\ d(X_3^1, m_1), d(X_3^2, m_2), \dots, d(X_3^n, m_n) \end{bmatrix}, \quad (18)$$

The function $d(X_j^i, m_i)$ is basically the gradient distance function [25] between the two data points in the lower dimensional space, which is given by the following formulation:

$$d(X_j^i, m_i) = \frac{\|(X_j^i - m_i)^T \nabla E(i | X_j^i)\|}{\|\nabla E(i | X_j^i)\|}, \quad (19)$$

where $\nabla E(i | X_j^i)$, is the direction of gradient of i th subject with probability E and is given [25] by Eq. (20).

$$\nabla E(i | X_j^i) = \sum_{j=1}^c E(X_j^i | j)(m_j - m_i) . \quad (20)$$

In this Equation, m_i and m_j are mean features of the i th and j th subjects respectively. The entity with least distance among the sub-classes in the f dimension of non-linear subspace is considered to give the final identity using the formula [25] in Eq. (21).

$$\nabla E(X_j^i | j) = \frac{1}{\sqrt{(2\pi)^f}} \exp\left\{-\frac{1}{2}(X_j^i - m_j)^T (X_j^i - m_j)\right\} . \quad (21)$$

Overall flow of proposed framework is summarized in Algorithm 1 as in the following:

Algorithm 1:iKFDA

Input: Generate a set of training images $\{X_j^i\}$ for each subject using Eq. (7)

Output: A correctly classified occluded image in kernel space f

Step 1: Compute the kernel matrix via kernel trick in [26]

Step 2. Calculate M and m_i using Eq. (8) and Eq. (9)

Step 3. Calculate S_B^Φ using Eq. (2) and S_W^Φ via Eq. (3)

If $|S_W^\Phi| = 0$ use Eq. (12) to compute the inverse of S_T^Φ otherwise use conventional LDA methodology

Step 4. Perform Eigenvalue decomposition via Eq. (1)

Step 5. Update the global mean using Eq. (13)

Step 6. Update S_B incrementally on arrival of new subject using Eq. (14)

If subject is already member of set L then use Eq. (15)

Step 7. Increment S_W via Eq. (16)

Step 8. Compute gradient distance of project features using Eq. (21)

4. Results and Experiments

To determine the robustness of the proposed work, experiments were performed on four renowned databases. The datasets involved in the evaluation process included AR database [27] which is considered to be the benchmark for evaluating the effectiveness of algorithms. It comprises of 4000 colored images of 126 individuals. Each subject has a total of 26 images taken in two different sessions which comprises of 13 images in each session. Each set of

images contains variant illumination, expression and above all natural occlusion due to scarf and sunglasses. ORL [28] dataset contains 40 subjects, each having 10 images. The images vary in slight head orientations, expressions and presence and absence of glasses. Yale B database contains 15 subjects, each having 11 images. These images vary in lighting conditions and expressions. This database is selected to check the performance of the system without occlusion information and images belonging to MIT-CBCL database contain faces and non-face images. It is provided with a training set of 2429 faces and 4548 non-face images while the probe set comprises of 472 face images and 23573 non-face images. Images of faces include frontal as well as different profile views. All the experiments thus performed are based on the assumption that images are aligned according to eye position and histogram equalized to obtain stable results. The experiments proceeded in three steps. Firstly, tests are performed on whole images using different training samples, presence of random pixel corruption, natural and synthetic occlusion. Secondly, experiments are conducted on partitioned images containing synthetic, natural occlusion and availability of partial features. Finally, average recognition rates are calculated, i.e. after performing each experiment several times using the following mechanism:

$$R_a = \frac{\sum_{t=1}^t e_c^t}{e_T} \quad (22)$$

Where t is the number of experiments, e_c^t is the number of correct classifications for t^{th} run, e_T is the total experiments and R_a is the average correct matching rate. Criteria for matching the two face images are based on the distance function of similarity computed in the earlier sections. After the images are projected on the kernel feature space, the subclasses with least distance are taken as correct identity. The following subsections explain the process of experiments for evaluating the effectiveness of the technique.

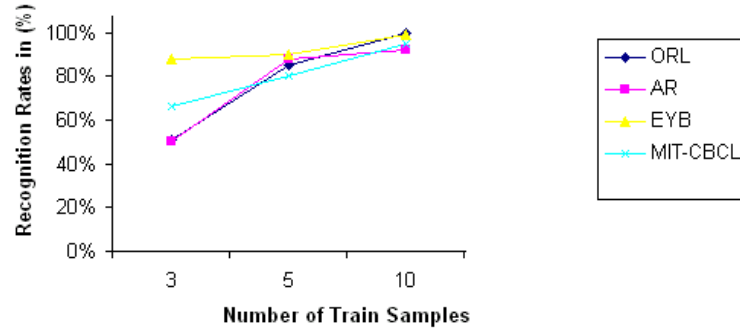
4.1 Without occlusion

This set of experiment deals with face recognition scenario in which the images are taken from above mentioned databases. Experiments are performed with varying training samples of each subject i.e., 3, 5 and 10 and corresponding recognition rates are obtained. Size of images from ORL database is kept at its original i.e., 112×92 , for AR database images size is 120×165 , in Yale B database images are cropped to 133×122 whereas for MIT-CBCL face images are cropped to 128×128 pixels. From each dataset a total of 10 sample images of each subject are considered in which 3 images are trained at the initial stage, and then 5 and lastly 10 images are used. Later experiment is performed with randomly selecting three training images per person from each dataset while rest of eight samples constitutes of testing set. The next step involves a series of experiments which requires selecting five random images per person for training and remaining five images from the probe set and lastly 10 training samples are considered for gallery and 10 images forms the probe set. In these experiments whole face is taken into consideration and the recognition rates obtained on the above databases are represented in [Table 1](#) and graphical illustration in [Fig. 2](#).

Table 1. Recognition rates on different training samples

Databases	Recognition Rates on Different Training Samples		
	3	5	10

ORL	51%	85%	100%
AR	50%	88%	92%
EYB	88%	90%	99%
MIT-CBCL	66%	80%	95%



2. Fig. 2. Recognition rates on different training samples

4.2 Random Corruption of Pixels

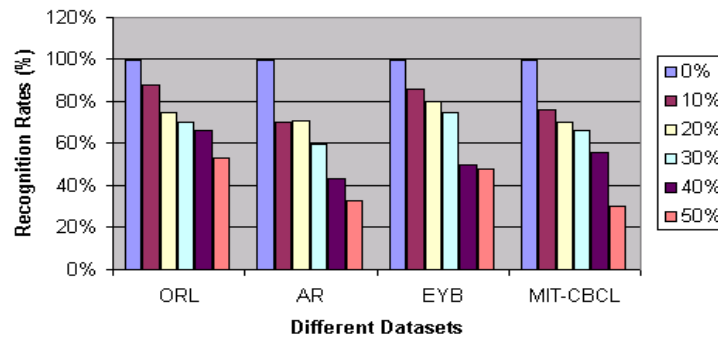
For this experiment, images from above databases are chosen and random noise is introduced on the images. Images used for training and testing are same which are performed on a total of 100 training and testing samples for explaining the robustness of the proposed framework under random pixel corruption. In this case, salt and pepper noise is incorporated in the testing images, no noise is introduced in the training set. Salt and pepper noise is incorporated due to the fact that it lays the appearance of image having black and white dots which give the effect of pixels values as being occluded. The size of the images is kept the same as in previous experiments. The amount of noise intensity introduced is varied from 0%, 10%, 20%, 30% up till 50% and the effect of each is recorded and displayed as an average of recognition rates obtained after performing each test several times. Fig. 3 shows the presence of noise on sample AR database image. Table 2 and Fig. 4 demonstrate the recognition rates on the datasets.



3. Fig. 3. Sample image with random noise

4. Table 2. Recognition rates with random corruption of pixels

Percentage Corruption	0%	10%	20%	30%	40%	50%
ORL	100%	88%	75%	70%	66%	53%
AR	100%	70%	71%	60%	43%	33%
EYB	100%	86%	80%	75%	50%	48%
MIT-CBCL	100%	76%	70%	66%	56%	30%



5. Fig. 4. Recognition rates with random corruption of pixels

4.3 Natural occlusion

In this experiment, AR database is considered for experimentation since it is the only available dataset containing natural occlusion. The training set includes 5 images i.e. neutral, smile, angry and partially occluded images (sunglass and scarf), of 50 male and 50 females. While its alternative images from second session are used for the probe set. Neutral face is taken so as to check the accuracy of the system while partially occluded images are used for determining the robustness of recognition system. Table 3 shows the recognition rate in percentage of the proposed method while considering natural occlusion and expression images. This test did not produce good results on occluded images while better recognition rates were obtained on expression images.

Table 3. Recognition rates (%) with partial occlusion and expression on AR database

	Partial Occlusions		Varying Expressions		
	Sunglasse	Scarf	Smile	Angry	Screa
ikFDA	60	65	97	96	90

4.4 Synthetic occlusion

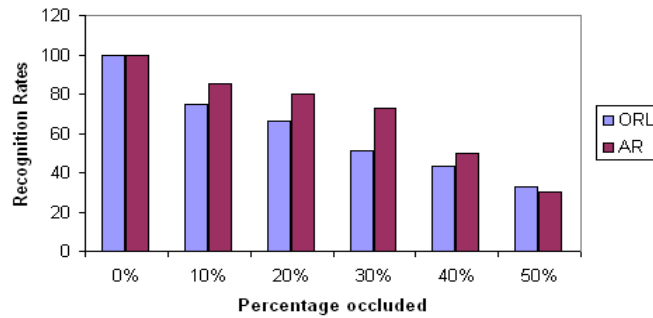
In order to further evaluate the performance of the proposed work, block occlusion is added on the frontal faces of AR and ORL databases. In this experiment, training and testing set is the same, images from both the dataset are firstly trained without occlusion and then tests are performed with block occlusion. Percentage of block occlusion is estimated using the block size of $r \times r$ pixels where $r = \{10, 20, \dots, 50\}$ placed at arbitrary location on the images. If block size is 10×10 then 10% occluded block is incorporated on the image. From ORL database neutral, smile and slightly tilted head of each subject is used for training while the occluded version of these images forms the testing set. These images are used for checking the performance of the system under synthetic occlusion. From AR database neutral, smile and angry images of each subject is used. Synthetic occlusion added by a black square block on the face images is shown in Fig. 5. Table 4 and Fig. 6 represent the recognition rates in % on images from ORL and AR datasets containing synthetic occlusion. They also depict the percentage of occluded block.



6. **Fig. 5.** Sample images from ORL dataset with synthetic occlusion

7. **Table 4.** Recognition rates on databases under synthetic occlusion

Databases	Synthetic Occlusions					
	0%	10%	20%	30%	40%	50%
ORL	100	75	66	51	43	33
AR	100	85	80	73	50	30



8. **Fig. 6.** Recognition rates (%) on whole images with synthetic occlusion

4.5 Synthetic occlusion on partitioned images

After performing a series of experiments on whole image, now we partition the image into three subclasses. All the images used in this experimentation were aligned according to the eye coordinate; frontal faces were selected with slight head rotation as in ORL dataset. A horizontal strip of size $m \times n$ pixels is manually extracted for each local feature hence making the sub-class which is used in the recognition procedure. The size of m and n is kept at a level such that each subclass contains the local facial features i.e. first sub-class contains both the eyes; second subclass contains nose region and lastly lips and mouth region. Partitioned images of face are used for training and testing, difference lies in the fact that training is done without occluded images and testing is done with occluded blocks randomly placed on the strips. These subclasses correspond to local facial features; the size of each partition is kept at the same level so that it does not affect the recognition accuracy. In this set of experiments, AR and ORL databases are used. Tests were conducted on 100 subjects, each containing 5 images, so a total of 500 training as well as testing images were used. **Table 5** represents the recognition rates with synthetic occlusion. **Fig. 7** represents a sample image from AR dataset partitioned into its subsequent sub-classes.



9. **Fig. 7.** Showing three subclasses of a sample face

10.

Table 5. Recognition rates (%) with synthetic occlusion on partitioned images

Database	Percentage of Synthetic Block Occlusions					
	0%	10%	20%	30%	40%	50%
ORL	100	100	89	80	75	50
AR	100	95	92	85	80	77

4.6 Natural occlusion on partitioned images

A similar experiment is performed with partitioned images containing natural occlusions on face. This test is performed to minimize the overall effect of deformed pixels. Prior to this, tests were conducted on whole face image with no partitioning scheme. AR database is used because of natural occlusion since none of the other datasets considered experimentation containing natural occlusion. The partitioned images of classes are trained without occlusion and then probe images with occluded faces are used in the testing phase. Tests were conducted on 100 subjects, each containing 5 images, so a total of 500 training as well as testing images were used. Recognition rates obtained on natural occlusion with partitioned images are provided in [Table 6](#).

Table 6. Recognition rates on partitioned images containing natural occlusion

Partial Occlusions on Partitioned Images					
Method	Sunglasses	Scarf	Smile	Angry	Scream
ikFDA	99	95	97	96	90

4.7 Partial face recognition on AR database

Finally, experiment is performed to check if the system is able to recognize the correct identity despite giving partial features. The test is conducted first by eyes portion only i.e., training and testing sets consisting of eyes of the individuals. Secondly, nose portion is tested to recognize it and lastly, mouth region is given as input and decision is made based on the available contents. Further experiment is conducted on AR database with 100 subjects, each containing 5 images, so a total of 500 training as well as testing images were used. It may be pointed out that no occluded image is used in this experiment. [Table 7](#) represents the obtained recognition rates.

Table 7. Recognition rates (%) with partial facial features

Method	Partial Features		
	Eyes	Nose	Mouth
ikFDA	96	75	91

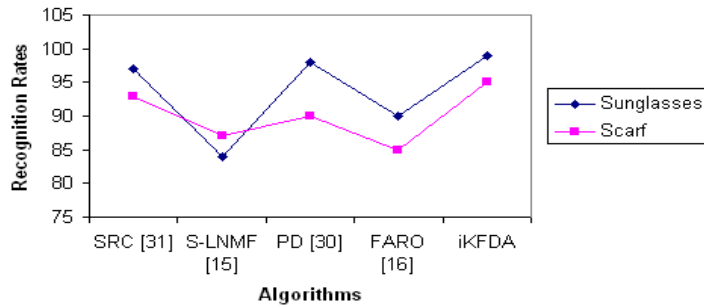
4.8 Comparisons and Discussion

Linear Discriminate Analysis (LDA) is generally considered as one of the most renowned classification techniques in general pattern classification tasks. However, linear methods show reduced classification capability when appearance is affected. In such instances non-linear versions are mostly used but their increased computational cost becomes a major issue. In view of this, the present study attempts to design a general framework for recognition of

partially occluded faces in kernel space in combination with incremental learning. **Table 8** and the corresponding **Fig. 8** provide an overview of the most commonly used methods which particularly deal with occluded faces.

Table 8. Recognition rates (%) of different methods

Algorithms	Sunglasses	Scarf
SRC [29]	97	93
S-LNMF [15]	84	87
PD [30]	98	90
FARO [16]	90	85
iKFDA	99	95



11. **Fig. 8.** Recognition rates of different methods

12.

To further evaluate the functionality of the proposed method a comparison is made with different methods dealing with synthetic occlusions. Comparison is made on the basis of block occlusion placed at arbitrary locations on the neutral images of AR database. **Fig 9(a)-(d)** shows some of the example occluded images used in the experiment. **Table 9** shows the comparative evaluation of different methods whose average recognition rates are shown below.



13. **Fig. 9(a)-(d).** Example images from AR database with synthetic block occlusions.

14. **Table 9.** Recognition rates (%) of different methods dealing in synthetic block occlusion

15. **Table 9.** Recognition rates (%) of different methods dealing in synthetic block occlusion

Algorithms	Synthetic occlusion at random locations as shown in Fig 9			
	a	b	c	d
R-PCA[31]	100%	100%	46%	24%
LNMF[32]	100%	96%	28%	10%
KLD-LGBP [11]	100%	100%	100%	100%
S-LNMF[15]	100%	100%	100%	100%

iKFDA	100%	100%	100%	100%
-------	------	------	------	------

The robustness of our work is based on the presence of natural occlusion on the face images due to sunglasses and scarf. The system is checked with whole images as well as partitioned images. These results represent the average recognition rates obtained after performing the tests several times. These experiments confirm that better recognition rates are obtained when the image is divided into subclasses as compared to whole image. This is due to the fact that partitioning the images into subclasses reduces the overall effect of any deformity present in the images.

The underlined merit of the present work is also due to the division of classes into fixed size sub-classes which greatly reduces the effect of spatially deformed images [29]. Moreover, the method showed great performance even in synthetic occlusion and expressions. Gradient distance function [25] used provides an added benefit due to its design principally for feature extraction process involving Fisher's criterion. The incremental learning in kernel space significantly improved the performance of new entries in the system with less consumption of resources involved in retraining new subjects.

5. Conclusion

This study has proposed an incremental Kernel Fisher Discriminate Analysis (iKFDA) to deal with partial occlusion caused by accessories like sunglasses or scarf. It establishes the fact that the division of facial images into subsequent sub-classes has come to play a pivotal role in dealing with likely occlusions. The study has also demonstrated that the incremental learning of non-linear space greatly reduces the resource consumption as well as computational cost involved in retraining the subspace for new subjects. Our results confirm the robustness of this novel technique in dealing with natural and synthetic occlusions, which commonly occur in facial image. The algorithm has achieved a recognition rate of 95-99% with natural occlusion giving a significantly reduced error.

References

- [1] Turk M and Pentland A., "Eigenfaces for recognition," *Journal of cognitive neuroscience*, vol. 3, no. 1, pp. 71-86, 1991. [Article \(CrossRef Link\)](#)
- [2] Lawrence S, Giles CL, Tsoi AC and Back AD., "Face recognition: A convolutional neural-network approach," *IEEE Transaction on Neural Networks*, vol. 8, no. 1, pp. 98-113, 1997. [Article \(CrossRef Link\)](#)
- [3] Kirby M and Sirovich L., "Application of the Karhunen-Loeve procedure for the characterization of human faces," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 12, no. 1, pp. 103-108, 1990. [Article \(CrossRef Link\)](#)
- [4] Belhumeur PN, Hespanha JP and Kriegman DJ., "Eigenfaces vs. fisherfaces: Recognition using class specific linear projection," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 711-720, 1997. [Article \(CrossRef Link\)](#)
- [5] Baudat G and Anouar F., "Generalized discriminant analysis using a kernel approach," *Neural computations*, vol. 12, no. 10, pp. 2385-2404, 2000. [Article \(CrossRef Link\)](#)
- [6] Schölkopf B, Smola A and Müller K-R., "Nonlinear component analysis as a kernel eigenvalue problem," *Neural computations*, vol. 10, no. 5, pp. 1299-1319, 1998. [Article \(CrossRef Link\)](#)
- [7] He X, Yan S, Hu Y, Niyogi P and Zhang H-J., "Face recognition using laplacianfaces," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 27, no. 3, pp. 328-340, 2005. [Article \(CrossRef Link\)](#)

- [8] Kim TY, Lee KM, Lee SU and Yim C-H., "Occlusion invariant face recognition using two-dimensional PCA," *Advances in Computer Graphics and Computer Vision: Springer*, 305-315, 2007. [Article \(CrossRef Link\)](#)
- [9] Tan X, Chen S, Li J and Zhou Z-H, "Learning non-metric partial similarity based on maximal margin criterion," in *Proc. of IEEE 2006 Computer Vision and Pattern Recognition Conference*, 168-145, 2006. [Article \(CrossRef Link\)](#)
- [10] Jia H and Martinez AM, "Support vector machines in face recognition with occlusions," in *Proc. of IEEE 2009 Computer Vision and Pattern Recognition Conference*, pp. 136-141, 2009. [Article \(CrossRef Link\)](#)
- [11] Zhang W, Shan S, Chen X and Gao W., "Local gabor binary patterns based on kullback-leibler divergence for partially occluded face recognition," *IEEE Signal Processing Letters*, vol. 14, no. 11, pp. 875-878, 2007. [Article \(CrossRef Link\)](#)
- [12] Kim J, Choi J, Yi J and Turk M., "Effective representation using ICA for face recognition robust to local distortion and partial occlusion," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 27, no. 12, pp. 1977-1981, 2005. [Article \(CrossRef Link\)](#)
- [13] Lin J, Ming J and Crookes D., "Robust face recognition with partially occluded images based on a single or a small number of training samples," in *Proc. of IEEE 2009 Acoustics, Speech and Signal Processing Conference, ICASSP 2009*, pp. 881-884, 2009. [Article \(CrossRef Link\)](#)
- [14] Tarrés F, Rama A and Torres L., "A novel method for face recognition under partial occlusion or facial expression variations," in *Proc. of IEEE 2005 47th International Symposium*, pp. 163-166, 2005. [Article \(CrossRef Link\)](#)
- [15] Oh HJ, Lee KM and Lee SU., "Occlusion invariant face recognition using selective local non-negative matrix factorization basis images," *Image and Vision Computing*, vol. 26, no. , pp. 1515-1523, 2008. [Article \(CrossRef Link\)](#)
- [16] De Marsico M, Nappi M and Riccio D., "FARO: Face recognition against occlusions and expression variations," *IEEE Transactions on Systems Man and Cybernetics—part a: systems and humans*, vol. 40, no. 1, pp. 121-132, 2010. [Article \(CrossRef Link\)](#)
- [17] Hotta K., "Robust face recognition under partial occlusion based on support vector machine with local Gaussian summation kernel," *Image and Vision Computation*, vol. 26, no. 11, pp. 1490-1498, 2008. [Article \(CrossRef Link\)](#)
- [18] Martínez AM., "Recognition of partially occluded and/or imprecisely localized faces using a probabilistic approach," in *Proc. of IEEE 2000 Computer Vision and Pattern Recognition Conference*, pp. 712-717, 2000. [Article \(CrossRef Link\)](#)
- [19] Weng J, Zhang Y and Hwang W-S., "Candid covariance-free incremental principal component analysis," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 25, no. 8, pp. 1034-1040, 2003. [Article \(CrossRef Link\)](#)
- [20] Chatterjee C and Roychowdhury VP., "On self-organizing algorithms and networks for class-separability features," *IEEE Transaction on Neural Networks*, vol. 8, no. 3, pp. 663-678, 1997. [Article \(CrossRef Link\)](#)
- [21] Ye J, Li Q, Xiong H, Park H, Janardan R and Kumar V., "IDR/QR: an incremental dimension reduction algorithm via QR decomposition," *IEEE Transactions on Knowledge and Data Engineering*, vol. 17, no. 9, pp. 1208-1222, 2005. [Article \(CrossRef Link\)](#)
- [22] Zhao H and Yuen PC., "Incremental linear discriminant analysis for face recognition," *IEEE Trans on Syst Man and Cybernetics —Part b: Cybernetics*, vol. 38, no. 1, pp. 210-221, 2008. [Article \(CrossRef Link\)](#)
- [23] Chin T-J and Suter D., "Incremental kernel principal component analysis," *IEEE Transaction on Image Processing*, vol. 16, no. , pp. 1662-1674, 2007. [Article \(CrossRef Link\)](#)
- [24] Chen L-F, Liao H-YM, Ko M-T, Lin J-C and Yu G-J., "A new LDA-based face recognition system which can solve the small sample size problem," *Journal of Pattern Recognition*, vol. 33, no. 10, pp. 1713-1726, 2000. [Article \(CrossRef Link\)](#)
- [25] Kittler J, Li Y and Matas J., "On Matching Scores for LDA-based Face Verification," in *Proc. of 47th Proceedings of British Machine Vision Conference*, pp. 1-10, 2000. [Article \(CrossRef Link\)](#)
- [26] Zhang B., "Gabor-kernel fisher analysis for face recognition," *Advances in Multimedia*

- Information Processing-PCM 2004: Springer*, pp. 802-809, 2005. [Article \(CrossRef Link\)](#)
- [27] Martinez AM., "The AR face database," *CVC Tech Rep*, vol. 24, 1998. [Article \(CrossRef Link\)](#)
- [28] Samaria FS and Harter AC., "Parameterisation of a stochastic model for human face identification," in *Proc. of IEEE 2nd Workshop on Applications of Computer Vision*, pp. 138-142, 1994. [Article \(CrossRef Link\)](#)
- [29] Wright J, Yang AY, Ganesh A, Sastry SS and Ma Y., "Robust face recognition via sparse representation," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 31, no. 2, pp. 210-227, 2009. [Article \(CrossRef Link\)](#)
- [30] Tan X, Chen S, Zhou Z-H and Liu J., "Face recognition under occlusions and variant expressions with partial similarity," *IEEE Transactions on Information Forensics and Security*, vol. 4, no. 2, pp. 217-230, 2009. [Article \(CrossRef Link\)](#)
- [31] Leonardis A and Bischof H., "Robust recognition using eigenimages," *Computer Vision and Image Understanding*, vol. 78, no. 1, pp. 99-118, 2000. [Article \(Cross Ref Link\)](#)
- [32] Li S. Z, Hou X.W, Zhang H.J and Cheng Q.S., "Learning spatially localized, part-based representation," in *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, vol. 1, pp. 1-207, 2001. [Article \(Cross Ref Link\)](#)



Aisha Azeem is a student of COMSATS Institute of Information Technology (CIIT), Wah Campus, Pakistan. Her research interests include Digital Image Processing.



Dr. Muhammad Sharif is working as Assistant Professor in Department of Computer Science, COMSATS Institute of Information Technology, Wah Campus, Pakistan. He has more than 17 years of teaching experience at graduate/ undergraduate level as well as research in different areas. He has more than 90 publications in journals and conferences of national and international repute. His areas of research include Digital Image Processing, Algorithms analysis, Intelligent systems and Parallel & Distributed Computing.



Jamal Hussain Shah is Lecturer in Department of Computer Science, COMSATS Institute of Information Technology, Wah Campus, Pakistan. His research areas are Digital Image Processing and Networking. Mr. Jamal has more than 3 years' experience in developing IT-related projects. He designed and developed a number of ERP systems for different reputable organizations of Pakistan.



Mudassar Raza is working as Assistant Professor in Department of Computer science, COMSATS Institute of Information Technology, Wah Campus, Pakistan. He holds more than eight years of experience of teaching undergraduate classes. Mudassar Raza has also been supervising final year projects of undergraduate students. His areas of interest include Digital Image Processing, Algorithms Analysis and Parallel & Distributed Computing.