Engineering Science and Technology, an International Journal 19 (2016) 1321–1333

HOSTED BY
ELSEVIER

Contents lists available at ScienceDirect

Engineering Science and Technology, an International Journal

journal homepage: www.elsevier.com/locate/jestch

Full Length Article

Kernel Locality Preserving Symmetrical Weighted Fisher Discriminant Analysis based subspace approach for expression recognition

G.P. Hegde [a,*], M. Seetha [b], Nagaratna Hegde [c]

[a] Dept. CSE, SDMIT, Ujire affiliated to VTU Belgaum, Karnataka State, India
[b] Dept .of CSE, GNITS, Hyderabad, India
[c] Dept. of CSE, VCE, Hyderabad, India

ARTICLE INFO

Article history:
Received 26 August 2015
Revised 5 March 2016
Accepted 6 March 2016
Available online 13 April 2016

Keywords:
Discriminant analysis
Gabor filter
Expression recognition
Feature extraction
Subspace
Symmetrical weight

ABSTRACT

This paper mainly focuses on dimensional reduction of fused dataset of holistic and geometrical face features vectors by solving singularity problem of linear discriminant analysis and maximizing the Fisher ratio in nonlinear subspace region with the preservation of local discriminative features. The combinational feature vector space is projected into low dimensional subspace using proposed Kernel Locality Preserving Symmetrical Weighted Fisher Discriminant Analysis (KLSWFDA) method. Matching score level fusion technique has been applied on projected subspace and combinational entire Gabor subspace is framed. Euclidean distance metric (L2) and support vector machine (SVM) classifier has been implemented to recognize and classify the expressions. Performance of proposed approach is evaluated and compared with state of art approaches. Experimental results on JAFFE, YALE and FD expression database demonstrate the effectiveness of the proposed approach.

1. Introduction

Several face appearance situations exhibit varieties of expressions like happy, sad, angry, fatigue, confusion, surprise, thinking, fear, pain, wink, fun and disgust. The purpose of recognizing the expressions is to understand the feelings of face appearances for several applications in various fields like pattern recognition and computer vision. Sometimes it is needed to know the nonverbal capability of lecture class by the students during smart class teaching, this can be determined through recognizing the expressions of individual students. Driver feelings can be determined during vehicle driving to avoid the accidental hazards. In real time election voting, determination of persons identity at different expressions, is made possible with standard facial expression recognition system.

One and half decade back Ekman et al. [1] carried out study of expressions and observed that at least six expressions like anger, disgust, fear, happiness, sadness and surprise are exhibited by human beings and author noted a neutral state as normal expression. In pattern recognition, computer vision and biometrics, facial expression recognition task is one of the most challenging works due to larger variations of illuminations, noisy environments during recognition [1–4]. There are many applications in psychological studies, medical diagnosis, during different painful situations, and determination of human emotional states for criminal and security issues [4]. In this work singularity problem of scatter matrix has been resolved by introducing symmetrical weighted principal components at kernel region by preserving local discriminant features. To synthesize a complete image face under an appearance based approach, both shape and textures features found a significant domain in several studies as given in [5–7,66]. The rest of the paper is organized as follows: Section 2, presents overview of earlier works. Section 3, illustrates proposed framework by comparing related earlier works. In Section 4, results and discussion are made. In Section 5, conclusions are summarized.

2. Literature survey

Linear discriminant analysis (LDA) method is an important task for recognition of objects used in many fields. When all the scatter matrices are become singular, it causes degradation of efficiency of expression recognition for small data samples, hence it is difficult to maintain larger variability between classes. Most of the LDA extension methods or algorithms proposed in earlier studies were unable to optimize the singularity matrix problems. This problem was resolved by Belhumeur et al. [9] who proposed a Fisherface method (FF) in 1997, which uses a principal component analysis (PCA) [64], based projection and a change of matrix size so that

* Corresponding author.
Peer review under responsibility of Karabuk University.

http://dx.doi.org/10.1016/j.jestch.2016.03.005
2215-0986/© 2016 Karabuk University. Publishing services by Elsevier B.V.

the matrix is nonsingular. But still for Fisher LDA algorithm it is needed to improve the singularity problem scenarios. In this paper singularity problem of within class scatter matrix have been considered as one of the main issues for dimensional reduction purpose. LDA is a linear supervised subspace method using class labels it discriminates the different expression classes and converts the high dimensional space into subspace. There are several LDA based methods improved and proposed by several researchers and authors. Li et al. [10] worked on discriminant analysis with non parametric approach based face recognition. Ming et al. [11] introduced spectral regression kernel discriminate analysis (SRKDA) based on regression and spectral graph analysis. They have suggested that when the sample vectors are non-linear then SRKDA method can efficiently give better solutions than ordinary subspace learning approaches. Wang and Sun [50] proposed semi supervised kernel marginal Fisher analysis (SKMFA) in which authors suggested that singularity problem can be avoided by non linear structure captured by the data dependent kernel based on labeled and unlabeled data. Rahulamathavan et al. [5] developed facial expression recognition system with encrypted domain using linear Local Fisher discriminant analysis (LFDA). Author suggested there was a challenge to work with encrypted domain even if there was a not good recognition rate for unencrypted domain. This method has been applied to JAFFE database and achieved 94.37% recognition rates. Depending on facial land marks position features were extracted from specific facial active patches of different face appearances. Discriminative features were considered by further patches obtained from active patches for expression classification. SVM one against one classification technique was implemented by the authors Happy and Routray [62].

The effect of expression analysis using subspace methods have been made in [35,36].

There are several researchers who implemented subspace projection methods directly on input images to achieve feature extraction and dimension reduction. In [45] different earlier subspace methods were implemented on feature dataset for dimensional reduction and the strength and weakness of subspace methods were compared. From literature survey it has been noted that many linear and nonlinear subspace methods were found to be more robust for expression recognition. Subspace methods like principal component analysis (PCA) [5–15,31], linear discriminant analysis based [52] and Fisher LDA based approaches [16–22,63], and locality preserving projection (LPP) [23–27,30] are linear approaches. Nonlinear approaches include mapping subspace (Isomaps) [28,29], KPCA [34], Locality Preserving Fisher Discriminant Analysis (LFDA) [40], KFDA [45], and KLFDA [44,48,49]. The common drawback of these nonlinear embedding methods are consumes more time while computing high dimensional feature dataset. Yu et al. [32] proposed a direct LDA algorithm for face recognition which incorporates the concept of null space for high dimensional data. A complete kernel Fisher discriminant framework for feature extraction and recognition using KPCA and LDA is proposed in [34]. State of art approaches based on LDA algorithms are listed in Table 1.

Table 1
Comparison of overall recognition accuracy for JAFFE face dataset for proposed and state of art approaches.

Literature	Approaches	OFERR (%)
Zhang et al. [3]	LBP based LDA	73.4 ± 5.6
Zhang et al. [3]	Boosted LBP based LDA	77.67 ± 5.7
Wang et al. [51]	Orthogonal LDA	86.33
Cohen et al. [4]	LFDA	90.70
Shih et al. [53]	2DLDA + SVM	94.13
Dongcheng et al. [54]	Gabor + PCA, Gabor + 2DPCA	91 and 94
Bai et al. [55]	Gabor + LBP + LDA	92–97
Zhi and Ruan [56]	2D discriminant LPP	95.91
Zhang et al. [57]	Multilayer perceptron	90.34
Liejun et al. [58]	SVM based	95.7
Zhao et al. [59]	PCA and NMF	93.72
Lee [60]	RDAB	96.67
Ours	CEGKLSWFDA	97.14

The conventional LDA problem attempts to find an optimal linear transformation by minimizing the total intra class distance and maximizing the between class distance simultaneously. It is well known that this optimization problem can be solved by applying eigenvalue decomposition to the scatter matrices. However, this requires the total scatter matrix to be nonsingular.

Singularity problem of scatter matrices of discriminative analysis subspace methods degrades the Fisher ratio and efficiency of recognition during intra and inter class local feature separation and dimensional reduction.

Less preservation of local and global features of face images causes less recognition rate in nonlinear region is counted as another problem during projection of subspace.

To optimize the above problems in this work kernel based locality preserving symmetrical weighted Fisher discriminant analysis subspace approach is proposed for dimension reduction of higher dimensional feature dataset defined by different scales and orientations of Gabor filter by introducing symmetrical weighted principal components while projecting PCA space into subspace for creating FLDA space. Larger values of eigen components yield good feature for efficient recognition of expressions.

3. Proposed frame work

This paper mainly focuses on and illustrates the projection of high dimensional image space into low dimensional subspace by solving the singularity problems of linear discriminant analysis. This task has been carried out by proposing Kernel Locality Preserving Symmetrical Weighted Fisher Discriminant Analysis algorithm (KLSWFDA). In the beginning of this expression recognition system work, face detection [65] was carried out and a trained database has been created from input raw images by resizing the images as per the previous work as given in [42]. Texture feature extraction has been carried out by implementing Gabor filter as given in [39]. Both Gabor magnitude and phase parts are isolated and features are extracted separately. Then combinational entire Gabor feature dataset has been formed by fusing the Gabor magnitude feature vector and Gabor phase vector with geometrical feature vector (from 18 fiducially created points) as presented in Fig. 1 for equal distribution of feature dimension. These two vectors are named as combinational Gabor magnitude vector (CGMV) and combinational Gabor phase vector (CGPV). These vector sizes are found to be large in dimension and projected into subspace by applying KLSWFDA algorithm. Both these projected subspaces with similar matrix are fused by matching score level fusion as introduced in [33] and combinational entire Gabor subspace has been framed. Using Euclidean distance metric (L2) and SVM [41] classifier technique all the expressions are recognized and classified. Gabor filter is constructed with different scales and orientation parameters for different dimensions are listed in Tables 3 and 4 for JAFFE and YALE database respectively. Geometrical feature are extracted as per the procedure mentioned in [42].

Table 2
Comparison of overall recognition accuracy for YALE dataset for proposed and state of art approaches.

Literature	Approaches	OFERR (%)
Wang and Sun [50]	SKMFA	73.6
Wang and Gong [47]	Gabor + PCA + NN	86.64
Ours	CEGSWKLFDA	83.33

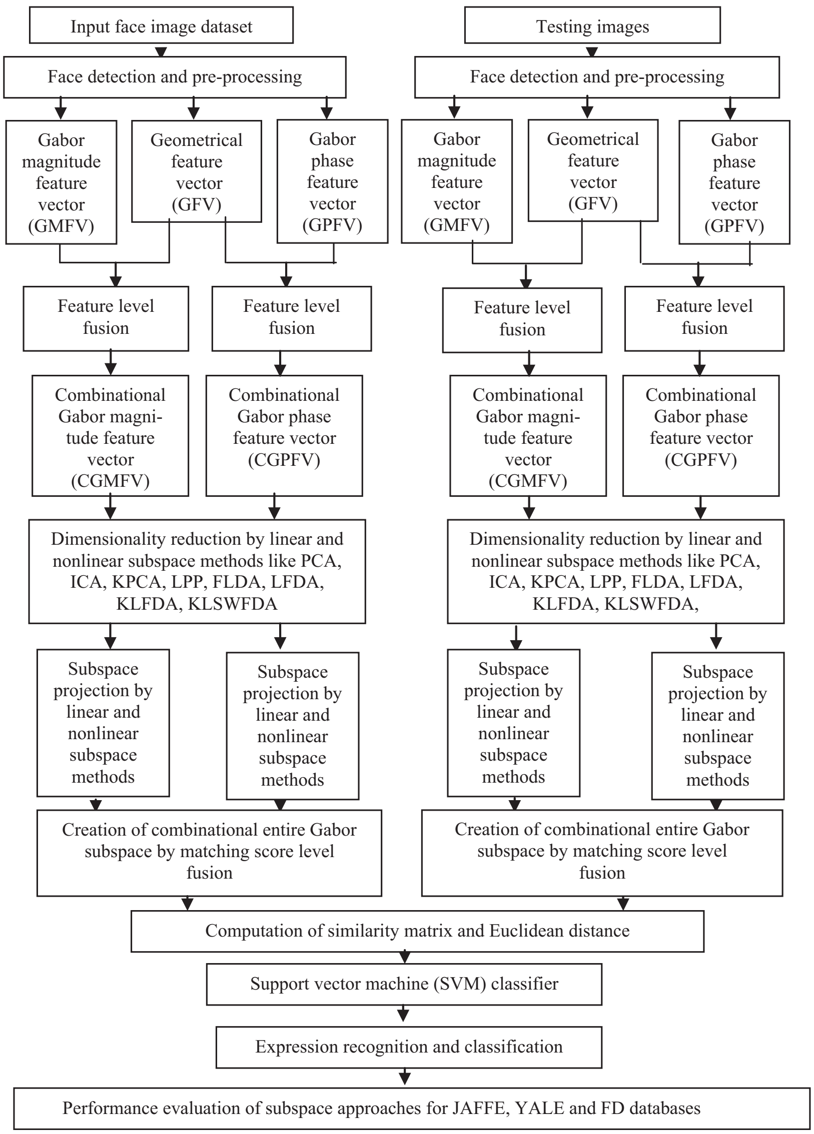

Fig. 1. Schematic diagram of expression recognition system using subspace approaches.

Table 3
Gabor filter parameters and feature vector dimension of JAFFE database.

Number of scales (m)	Number of orientations (n)	Gabor filter size (GF_{mn})	Gabor filter feature vector dimension (GF_{FVD})	Geometrical feature vector dimension (G_{FVD})	Combinational Gabor feature vector dimension (CG_{FVD})
5	4	20	279,720	16	279,736
3	8	24	335,664	16	335,680
3	4	12	167,832	16	167,848
5	8	40	559,440	16	559,456

Table 4
Gabor filter parameters and feature vector dimension of YALE database.

Number of scales(m)	Number of orientations (n)	Gabor filter size (GF_{mn})	Gabor filter feature vector dimension (GF_{FVD})	Geometrical feature vector dimension (G_{FVD})	Combinational Gabor feature vector dimension (CG_{FVD})
5	4	20	81,920	16	81,936
3	8	24	98,304	16	98,320
3	4	12	49,152	16	49,168
5	8	40	163,840	16	163,856

Gabor filter extracts rich texture content from face images. In general, Gabor filters are also called Gabor wavelets [39]. Gabor magnitude information can capture the facial structure and phase information can give a detailed description of facial texture [37,38].

3.1. Brief description about SWFLDA

PCA is a standard eigenface based popular algorithm used for dimensional reduction and feature extraction [17]. To find the symmetrical weighted FLDA the given feature dataset was subjected to PCA transformation. Larger variations with principal components and unequal distribution of components were observed in PCA subspace. It causes the class discrimination problem of feature data in LDA subspace. Let $G = (g_1, g_2, \ldots, g_i, \ldots, g_N)$ represent the $n \times N$ combination Gabor feature dataset matrix, where g_i is a combinational Gabor face vector of dimension n of $a \times b$ face vector matrix, m is a mean of combinational Gabor face vector and N is the number of different combinational Gabor feature data of input samples in the image dataset. Projection matrix of PCA [22] is given as

$$S_P = \sum_{i=1}^{N} (g_i - m)(g_i - m)^T \tag{1}$$

The value of S_P is made with equal distribution of eigen values by applying symmetrical weights to principal components for equalizing the variance of principal components in order to solve the singularity problems. Symmetrical weighted PCA representation is made in [46].

In this work odd–even rule is implemented as given in [46] to decompose face image. Combination Gabor dataset images are $g_i = [g_1, g_2, \cdots g_M]$ and mirror symmetrical combinational Gabor image set is $g_i^s = [g_1^s, g_2^s, \cdots g_M^s]$.. So the ith image can be decomposed as. $g_i = g_{oi} + g_{ei}$.. Where odd symmetrical image can be denoted by $g_{oi} = (g_i - g_i^M)$ and even symmetrical image can be denoted as $g_{ei} = (g_i + g_i^M)$.. Here $i = 1, 2, 3, \ldots M$. Odd symmetrical sample set is $(g_{o1}, g_{o2}, g_{o3}, \ldots, g_{oM})$ and even symmetrical sample set is $(g_{e1}, g_{e2}, g_{e3}, \ldots, g_{eM})$ both are derived from original training samples set by mirror symmetrical transformation. Odd and even sets can be defined as follows [46]

$$S_{oP} = \sum_{i=1}^{N} (g_{oi} - m_o)(g_{oi} - m_o)^T \tag{2}$$

$$S_{eP} = \sum_{i=1}^{N} (g_{ei} - m_e)(g_{ei} - m_e)^T \tag{3}$$

where, $S_P = S_{oP} + S_{eP}$, hence the eigenvalue decomposition on S_P is equal to the eigen decomposition on S_{oP} and S_{eP}. Hence, image g_i can be reconstructed by the feature vector of S_{oP} and S_{eP}. With respect to eigen theory assume all the non-zero eigen values of S_{oP} and S_{eP} are λ_{oi} and λ_{ej}, and the corresponding eigen vectors are w_{oi} and w_{ej} where $i = 1 \ldots \text{rank}(S_{oP})$ and $j = 1 \ldots \text{rank}(S_{eP})$. Transformation of weight matrix for odd (T_o) and even (T_e) symmetrical sample sets are derived from above demonstration as

$$T_o = [w_{o1}, w_{o2} \cdots w_{or_o}], \wedge_o = diag(\lambda_{o1}, \lambda_{o2}, \cdots \lambda_{or_o}) \tag{4}$$

$$T_e = [w_{e1}, w_{e2} \cdots w_{er_e}], \wedge_e = diag(\lambda_{e1}, \lambda_{e2}, \cdots \lambda_{er_e}) \tag{5}$$

where $r_o = \text{rank}(S_{oP})$, $r_e = \text{rank}(S_{eP})$.

The representation of the odd and even symmetrical images can be represented as,

$$g_{oi} = T_o P_{oi}, \quad P_{oi} = T_o^t g_{oi}, \quad g_{ei} = T_e P_{ei}, \quad P_{ei} = T_e^t g_{ei} \tag{6}$$

Above, P_{oi} and P_{ei} are the odd symmetrical feature and even symmetrical feature of the ith face combinational Gabor image. In order to reduce the effects made by the principal components which contain the variation due to illumination or face expression, it can treat each component equally and let each component have equal variance through transforming conventional PCA feature space to weighted PCA feature space by the following whitening transformation for odd symmetrical sample set and even symmetrical sample set as given in [46]:

$$Q_o = \wedge_o^{-1/2} T_o^t = (\lambda_{o1}^{-1/2} w_{o1}, \lambda_{o2}^{-1/2} w_{o2} \cdots \lambda_{or_o}^{-1/2} w_{or_o}) \tag{7}$$

$$Q_e = \wedge_e^{-1/2} T_e^t = (\lambda_{e1}^{-1/2} w_{e1}, \lambda_{e2}^{-1/2} w_{e2} \cdots \lambda_{er_e}^{-1/2} w_{er_e}) \tag{8}$$

Here Q_o and Q_e are the transform matrix of odd symmetrical images and even symmetrical images for WPCA feature space. In particular, the representation of the odd or even symmetrical images in WPCA feature space is given as [47]

$$g_{oi} = Q_o z_{oi}, \quad z_{oi} = Q_o^t g_{oi}, \quad g_{ei} = Q_e z_{ei}, \quad z_{ei} = Q_e^t g_{ei} \tag{9}$$

$$z_i = [z_{ei}^t, z_{oi}^t]^t = \begin{bmatrix} z_{ei} \\ z_{oi} \end{bmatrix}, \quad Q = [Q_e, Q_o], \quad \wedge = diag[\wedge_e, \wedge_o] \tag{10}$$

$$g_i = Q z_i = [Q_e, Q_o] \begin{bmatrix} z_{ei} \\ z_{oi} \end{bmatrix}, \quad z_i = Q^t g_i \tag{11}$$

For feature selection in symmetrical weighted PCA [46], sorting the eigenvalues either in ascending or descending order, the largest eigen vectors are selected corresponding to the first largest eigenvalues. Since the variance (corresponding to eigenvalues) of the weighted even symmetrical components is bigger than the variance of the correlative components of weighted odd symmetrical components, it is natural to consider the even symmetrical components first, and then the odd symmetrical components if necessary otherwise discarded. All the zero eigen values of PC components are eliminated.

In several expression classification approaches Fisher linear discriminant analysis (FLDA) finds a significant role for statistical feature extraction. Discriminant feature space is kept maximum by limiting the total number of training samples. When rank of within class scatter matrix is less than the number of features then within the class scatter matrices or intra covariance matrix becomes singular. So that limited training samples and dimensionality reduction problems occur. As mentioned in paper [22] LDA separated the samples of distinct groups by maximizing between class distances while minimizing within class distance. Between class matrix S_b can be given as

$$S_b = \sum_{i=1}^{C} N_i (\overline{m_i} - \overline{m})(\overline{m_i} - \overline{m})^T \tag{12}$$

Within class scatter matrix S_w can be defined as

$$S_w = \sum_{i=1}^{C} (N_i - 1) S_i = \sum_{i=1}^{C} \sum_{j=1}^{N_i} (m_{i,j} - \overline{m_i})(m_{i,j} - \overline{m_i})^T \tag{13}$$

where $m_{i,j}$ is the n-dimensional pattern j from class C_i, and N_i is the number of training pattern from class C_i, and C is the total number of classes or expression groups. The total meanvector is given by

$$\overline{m} = \frac{1}{N} \sum_{i=1}^{C} N_i \tag{14}$$

$$\overline{m_i} = \frac{1}{N} \sum_{i=1}^{C} \sum_{j=1}^{N_i} m_{i,j} \tag{15}$$

Vector $\overline{m_i}$ and matrix i are the unbiased sample mean and sample covariance of matrix of class. In above Eqs. (14) and (15) N is the total number of samples, that is $N = N_1 + N_2 + N_3 + - - - N_C$. It is important to note that the within class scatter matrix S_w defined Eq. (13) is essentially the standard pooled covariance matrix multiplied by the scalar factor $(N - C)$ and symmetrical weighted within scatter matrix becomes

$$S_{sw} = \sum_{i=1}^{C}(N_i - 1)S_i = (N - C)S_p \tag{16}$$

The main objective of FLDA is to find a projection matrix P_{FLDA} that maximizes the ratio of the determinant of the between-class scatter matrix to the determinant of the within-class scatter matrix [32] (Fisher's criterion), that is

$$P_{FLDA} = \arg\max_{P} \frac{|P^T S_b P|}{|P^T S_{sw} P|} \tag{17}$$

where P is the projection of PCA covariance matrix has an unequal variance distribution. Projection in FLDA subspace i.e. P_{FLDA} is in fact the solution of the following eigen system problem [32].

$$S_b P - S_{sw} P \Lambda = 0 \tag{18}$$

Multiplying both sides by S_{sw}^{-1}, Eq. (18) can be rewritten as

$$S_{sw}^{-1} S_b P - S_{sw}^{-1} S_{sw} P \Lambda = 0 \tag{19}$$

$$S_{sw}^{-1} S_b P - P \Lambda = 0 \tag{20}$$

$$(S_{sw}^{-1} S_b) P = P \Lambda \tag{21}$$

where P and Λ are respectively the eigenvectors and eigenvalues of $S_{sw}^{-1} S_b$. In other words, equation (21) states that if S_{sw} is a non-singular matrix then the Fisher criterion described in Eq. (17) is maximized when the projection matrix P_{FLDA} is composed of the eigenvectors of $S_{sw}^{-1} S_b$ with at most $(C - 1)$ nonzero corresponding eigenvalues. This is the standard FLDA procedure. The performance of the standard FLDA can be seriously degraded if there is only a limited number of total training observations N compared to the dimension of the feature space n. Since the within-class scatter matrix S_w is a function of $(N - C)$ or less linearly independent vectors, its rank is $(N - C)$ or less. Therefore, S_w is a singular matrix if N is less than $(n + C)$, or, analogously, might be unstable if N is not at least five to ten times $(n + C)$. The Fisherfaces subspace method is essentially a two-stage dimensionality reduction method. First the face images from the original vector space were projected to a lower dimensional space using Principal Component Analysis (PCA) [39] and then Linear Discriminant Analysis was applied next to find the best linear discriminative features on that PCA subspace. Or in other words PCA subspace is transformed into FLDA subspace. Let us consider P_{pca} is the projection matrix from the original image space to the PCA subspace, and P_{FLDA} is the projection matrix from the PCA subspace to the FLDA subspace obtained by maximizing the ratio as given below [32].

$$P_{FLDA} = \arg\max_{P} \frac{|P^T P_{pca}^T S_b P_{pca} P|}{|P^T P_{pca}^T S_{SW} P_{pca} P|} \tag{22}$$

(22) analogously states that if $P_{pca}^T S_{sw} P_{pca}$ is a non singular matrix then the Fisher criterion is maximized when the projection matrix P_{LDA} is composed of the eigenvectors of $(P_{pca}^T S_{sw} P_{pca})^{-1} (P_{pca}^T S_b P_{pca})$ with $(C - 1)$ at most nonzero corresponding eigen values. The singularity problem of the within-class scatter matrix S_w is then overcome if the number of retained principal components varies from at least C to at most N-C, PCA features. In PCA space unequal distribution of eigen values causes less discrimination among the features between class scatter matrix. Hence symmetrical weighted principal components strategy is used. If within scatter matrix is non singular then eigen vectors correspond to the set of the largest eigenvalues of matrix $(S_b + S_{sw}) - 1.S_b$. The problem of SWFLDA is the preservation of local image information under nonlinear region. This can be extended by constructing KLSWFDA as illustrated in Section 3.1.1.

3.1.1. Construction of KLSWFDA

In this section Kernel Locality Preserving Symmetrically Weighted Fisher Discriminant Analysis (KLSWFDA) finds projection of vectors in a higher dimensional kernel domain space such that it maximizes Fisher's ratio in that space. The idea of KLSWFDA is to solve the problem of SWFLDA and in an implicit feature space F constructed by a nonlinear mapping as in (23). Feature space F of Gabor face image dataset in kernel region can be defined as

$$\phi(g) = [\phi(g_1), \phi(g_2) \cdots, \phi(g_n)] \tag{23}$$

$$\phi : g \in R^N \to \phi(g) \in F$$

In implementation, implicit feature vector ϕ does not need to be computed explicitly, instead it is embodied by computing the inner product of two vectors in F with a kernel function, $k(x, y) = (\phi(x), \phi(y))$. Let g be a vector of the input Combinational Gabor feature dataset set with n elements and C classes, and i_n represents the number of samples in the ith class. The mapping of g_i is noted as $\phi_i = \phi(g_i)$. Performing FLDA in F mean to maximize the following Fisher discriminant function and the objective function of kernel discriminant analysis is given as

$$J(U) = \arg\max_{U} \frac{U^T S_b^{\phi} U}{w^T S_{sw}^{\phi} U} \tag{24}$$

within class kernel space symmetrical weighted scatter matrix is given as

$$S_{sw}^{\phi} = \sum_{i=1}^{C} \sum_{g \in G_i} (\phi(g) - m_i^{\phi})(\phi(g) - m_i^{\phi}) \tag{25}$$

S_{sw}^{ϕ} matrix is a symmetrically weighted within class scatter matrix its concept is illustrated above in Section 3.1. Between the class scatter matrix can be given as

$$S_b^{\phi} = \sum_{i=1}^{C} N_i (m_i^{\phi} - m^{\phi})(m_i^{\phi} - m^{\phi}) \tag{26}$$

where G_i is the number of samples from the ith class, m_i^{ϕ} is the centroid of the ith class, C is the number of classes, and m^{ϕ} is the global centroid, g is a vector for a specific class and G_i is the set of samples of the ith class. S_{sw}^{ϕ} represents the degree of symmetrical weighted scattering within class of expressions and is calculated as the summation of covariance matrices of each class, whereas S_b^{ϕ} represents the degree of scattering between classes of expressions and is calculated as the summation of the covariance matrix of the means of each class. The kernel domain space usually has a much higher dimension than the input space. The "kernel trick" [58] allows for the computation of algorithms in a kernel domain space without explicitly evaluating the mapping, as long as the algorithm can be expressed in terms of dot products of vectors in the input space.

LFDA preserves neighborhood relationships in the embedding by employing an "affinity" matrix that is defined below. The optimization solution $J(U)$, corresponding to the largest eigenvalues λ, can be illustrated by the generalized eigenvalue problem.

$$S_b^{\phi} U_i = \lambda_i S_{sw}^{\phi} U_i \tag{27}$$

Because the eigenvectors are linear combinations of $\phi(g_i)$, there exists a coefficient α_i such that

$$U = \sum_{i=1}^{n} \alpha_i \phi(g_i) \tag{28}$$

Let $\alpha = [\alpha_1, \alpha_2, \ldots, \alpha_3]$, it can be proved that Eq. (24) is equivalent to

$$\alpha_{opt} = \arg\max_{\alpha} \frac{\alpha^T K L^{(b)} K \alpha}{\alpha^T K L^{(w)} K \alpha} \tag{29}$$

And the corresponding generalized eigenvalue problem is

$$K L^{(b)} K \alpha = \lambda K L^{(w)} K \alpha \tag{30}$$

where K is the kernel matrix

$K_{ij} = k(g_i, g_j)$, $L^{(b)}$ and $L^{(w)}$ are defined as:

$$L^{(b)} = D^{(b)} - W^{(b)} \tag{31}$$

$$L^{(w)} = D^{(w)} - W^{(w)} \tag{32}$$

where $W^{(b)}$ is the weight matrix of local between adjacency graph and $W^{(w)}$ is the weight matrix of local within adjacency class matrix. $D^{(b)}$ and $D^{(w)}$ are both diagonal matrices, and their entries are column sums of $W^{(b)}$ and $W^{(w)}$ respectively.

$$D_{ii}^{(b)} = \sum_{j=1}^{n} W_{ij}^{(b)} \tag{33}$$

$$D_{ii}^{(w)} = \sum_{j=1}^{n} W_{ij}^{(w)} \tag{34}$$

So the problem of kernel linear discriminant analysis is converted into finding the leading eigenvectors of $(KL^{(w)}K)^{-1}\ KL^{(b)}K$, each eigenvector α gives a projection function U in the feature space. For a new data point g, its projection onto U in the feature space F can be computed by

$$\{U, \phi(g)\} = \sum_{i=1}^{n} \alpha_i \{\phi(g_i), \phi(g)\} \tag{35}$$

$$= \sum_{i=1}^{n} \alpha_i k(g_i, g) \tag{36}$$

During expression recognition singularity problem is solved because matrix $KL^{(w)}K$ and $KL^{(b)}K$ made to be guaranteed to be non-singular by eliminating zero eigen values. Hence the proposed method solves the problem of singularity matrix concept.

3.1.2. Post level fusion technique

Final scores obtained from projected subspace of both combinational Gabor magnitude space and combinational phase space parts are normalized using z-score normalization and fused together. Maximum normalized score is obtained and utilized for computation of Euclidean distance.

$$NS_{CGMKLSWFDA} = \frac{CGMKLSWFDA_S - \mu(CGMKLSWFDA_S)}{Std(CGMKLSWFDA_S)} \tag{38}$$

$$NS_{CGPKLSWFDA} = \frac{CGPKLSWFDA_S - \mu(CGPKLSWFDA_S)}{Std(CGPKLSWFDA_S)} \tag{39}$$

where $(CGMKLSWFDA)_S$ is similarity score matrix of combinational Gabor magnitude kernel locality preserved Fisher discriminant projected subspace. Similarly computational Gabor phase kernel locality preserved Fisher discriminant $(CGPKLSWFDA)_S$ similarity score matrix have also been computed. The combinational entire Gabor subspace matrix (CEG) is created by post level fusion technique.

$$W_{CEGKLSWFDA} = MAX[(NS_{CGMKLSWFDA} + NS_{CGPKLSWFDA})] \tag{40}$$

For both combinational entire Gabor subspace train and test image dataset final score weighted matrices are computed, then Euclidean distance is evaluated as

$$\varepsilon_i^2 = \|W_{CEGKLSWFDAQ} - W_{CEGKLSWFDAT}\|^2 \tag{41}$$

where $W_{CEGKLSWFDAT}$ and $W_{CEGKLSWFDAQ}$ are projected vector final score weight matrices of training and testing combinational entire Gabor subspace images. The image set with lower Euclidean distance is computed. Re-perform the operation on this image set with a lower threshold value to get the image having the expression closer to the defined image. The image with lowest Euclidean distance in expression images will be represented as the resultant expression image. So that testing expression image is matched with trained image. Based on Euclidean distance metric and RBF kernel based SVM classifier [41] facial expressions are classified.

4. Experimental testing and result analysis

4.1. Databases used

The experiments are performed in order to analyze the performance of the proposed approach on three public databases as given below. Proposed approach has been tested for JAFFE, YALE and FD with different dimensions of baseline feature dataset.

4.1.1. JAFFE database

In this work, Japanese Female Facial Expression (JAFFE) database [61] is used for experiment. This database contains 213 images of 7 facial expressions. In that six expressions are basic and one is neutral facial expression. All the expressions are posed by 10 Japanese female models of 256 * 256 resolution. Each image has been rated on 6 emotion adjectives by 60 Japanese subjects. In this work all the images of this database are pre-processed to obtain pure facial expression images, which have normalized intensity, uniform size and shape. Illumination and lighting effects also removed as given in [43]. The pre-processing procedure used in this work performs detecting facial feature points automatically including eyes, nose and mouth. Finally histogram equalization technique is used to remove illumination effects. Fig. 2 shows pre-processed and resized cropped samples of JAFFE database.

4.1.2. YALE database

This database YALE contains 11 images per person for 15 individuals resulting in a total of 165 images. The images in this database reveal major variations of illumination changes, different facial expressions, and the persons wearing eyeglasses/no eyeglasses. The original size of the images in this database is 243×320 pixels with 256 gray levels. For experiments, the size of these images was scaled down to 64 * 64 pixel size. In this work six expressions were used for experiment such as happy, surprise, sad, wink, sleep and neutral. Totally 90 images were considered for experiment without doing histogram equalization operation. Few samples of YALE database is shown in Fig. 3.

4.1.3. FD database

Another database used in this work is Facial expression face database (FD) that consists of 13 subjects and each subject has 75 images with different expressions. This database has total 975 images. In this work 500 images are used with 10 subjects, five expressions such as happy, surprise, angry, sad and neutral. Each class of expression has 100 images. For experiments, all the images are pre-processed and the size of these images is scaled down to 92 * 92 pixel size shown in Fig. 4.

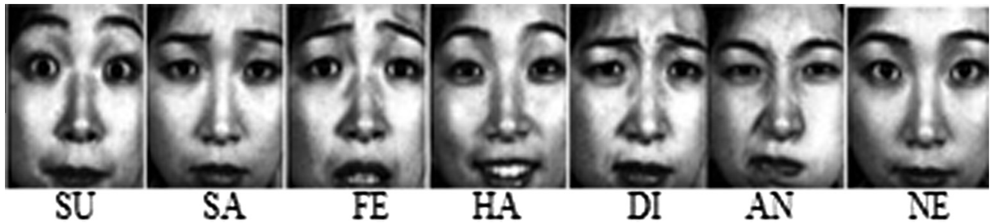

Fig. 2. Detected and cropped face samples of JAFFE database.

Fig. 3. Detected and cropped face samples of YALE database of size 64 * 64.

Fig. 4. Sample images of FD expression database of size 92 * 92 (preprocessed).

4.2. Testing results

In this work support vector machine classifier (SVM) [41] using Radial Basis Function (RBF) kernel technique is implemented to classify the expressions. To create input dataset, all 210 images of JAFFE database, 90 images of YALE database and 500 images of FD database are considered. In this work specific expression image is recognized using Euclidean distance metric between trained and testing images. Using "Leave One Out" SVM strategy all the expression classes of images are classified. All the public databases are tested with all the subspace models and proposed approach. In addition to a drastic reduction in the feature vector dimension for highest recognition rates are considered, it has been noted and observed that a considerable improvement in the recognition rate concerned with expression recognition. Performance of proposed approach is compared with state of art approaches listed in Tables 1 and 2.

The work mentioned in literature about Gabor + PCA + LDA is related to this approach but it is not concentrating on dimensional reduction after finding suitable solution to singularity problem in within scatter matrix. Simply Gabor features and PCA and LDA features were fused, and recognition accuracy was found to be high. Baseline method does not solve the efficient discrimination of face features of fused dataset of geometrical and Gabor filter texture information. Base line methods like PCA and LDA can preserve global properties of feature while dimensional reduction or projection of high dimensional data to subspace. But local properties of pixels cannot be preserved by PCA and LDA baseline methods. This can be made effective by proposed approach. Phase part of Gabor filter can compensate the poor recognition of expressions when it is combined with geometrical features. Proposed algorithm can reduce the redundant data which is created during fusion of phase and magnitude part with geometrical features. Adopted representation can enhance discriminative information compared to base line methods. Adopted approach is compared with related approaches as shown in Table 15.

In this work baseline features are created by fusing the 8 dimensional geometrical feature vectors (upper face part) with magnitude part of Gabor Filter. And remaining 8 dimensional geometrical vectors (lower face part) are fused with phase part of Gabor filter. Then total dimension is calculated for each input dataset. The feature dimension is found to be high as mentioned in Tables 3–5 respectively. Recognition accuracy is not considered during this stage because, if it comes more or high our goal will not complete. As per the objectives of this work we need the efficiency of recognition only during dimensional reduction at subspace scores fusion level. Hence base line method is not taken to account.

In this work 3 public datasets are considered for experiment. All the databases are having different dimensional images. We checked the subspace approaches for different image dimensional feature dataset with different Gabor filter parameters i.e. number of scales and number of orientations. Gabor filter with a high value of m and n yields higher texture features which influences and causes the improvement of recognition accuracy. Implementation part of this work has been carried out with SVM classifier by considering Euclidean distance as computing matching features metric. SVM leave one out technique finds better recognition accuracy, and this will suite the best classification technique for this whole subspace approach.

The recognition accuracy of base line approach might give the same recognition accuracy as obtained from proposed approach.

Table 5
Gabor Filter parameters and feature vector dimension of FD database.

Number of scales (m)	Number of orientations (n)	Gabor filter size (GF_{mn})	Gabor filter feature vector dimension (GF_{FVD})	Geometrical feature vector dimension (G_{FVD})	Combinational Gabor feature vector dimension (CEG_{FVD})
5	4	20	169,280	16	169,296
3	8	24	203,136	16	203,152
3	4	12	101,568	16	101,584
5	8	40	338,560	16	338,576

But regarding individual expression recognition rates of different expressions the proposed approach gives good accuracy of correct recognition for different expressions. This is due to elimination of null space from between scatter matrix and preservation of high discriminative data structure information in proposed approach during subspace conversion.

In proposed approach classification time performance is also improved compared to base line approach. But base line approach dimensional reduction vector found to be larger than proposed approach may consume more space in memory and larger classification time. In most of the base line approach features of LDA and PCA are added to Gabor features hence base line approach can give a yield nearer accuracy rate of proposed approach. But Gabor + PCA + LDA approach not solving the problem of singularity by eliminating zero eigen components completely from between scatter matrix and preserving local information of both scatter matrices completely.

In this section to investigate the performance of proposed approach for expression recognition from JAFFE, YALE and FD database for CEGPCA, CEGICA, CEGKPCA, CEGLPP, CEGFLDA, CEGLFDA, and CEGKLFDA approaches are compared. These subspace approaches have been framed for dimensionality reduction of higher dimensional baseline feature dataset obtained from concatenating of Gabor filter feature vector and geometrical feature vector dataset dimensions as given in Tables 3–5 respectively. For CEGLPP, CEGKPCA, CEGLFDA, CEGKLFDA and CEGKLSWFDA algorithms nearest neighbor number k is set to 7 where the value of σ was set to be 0.5. Performances of subspace approaches (time and space measures) are given in Tables 6–8 for FD, JAFFE and YALE database respectively. Figs. 5–7 show the comparison of subspace approaches in terms of overall expression recognition rates for JAFFE, YALE and FD database respectively. Overall facial expression recognition rate for JAFFE database is found to be 97.14% and overall expression recognition rate for YALE database is 83.84% and overall facial expression recognition rate for FD database is 93.33% using proposed approach. From the results it is noted that CEGKLSWFDA proposed approach consistently outperforms the CEGPCA, CEGICA, CEGKPCA, CEGLPP, CEGFLDA, CEGLFDA and CEGKLFDA expression recognition approaches. Comparison of individual facial expression recognition accuracy rates for JAFFE database, YALE database and FD database is presented in Figs. 8–10 respectively. From Table 9 it has been noted that for JAFFE database disgust, happy, sad, surprise and neutral expressions recognition rate is 100%. But anger, happy and fear expression recognition rate is 93.33%. Probably it is due to confusion with sad and disgust expressions. Confusion matrix of YALE database is shown in Table 10. From these results it has been noted that overall recognition rate is 83.84%. For happy, and surprise expressions correct recognition rate (CRR) is 100%. For sad and sleep expression (closed eyes) recognition rate is 77.78%. For wink expression CRR is 66.66% and neutral 80.88% respectively.

Table 6
Performance of subspace approaches for FD database at $m = 5$ and $n = 8$.

Subspace approaches	Overall facial expression recognition rate in (%) (OFERR)	Classification time in (s) (CT)	Dimension reduction feature vector (DR_{FV})
CEGPCA	79.46	1.967	175
CEGICA	80.80	1.935	175
CEGKPCA	82.02	1.781	175
CEGFLDA	85.94	1.209	175
CEGLPP	84.28	1.126	175
CEGLFDA	89.46	1.098	175
CEGKLFDA	91.20	1.012	150
CEGKLSWFDA	93.33	0.989	150

Table 7
Performance of subspace approaches for JAFFE database at $m = 5$ and $n = 8$.

Subspace approaches	Overall facial expression recognition rate in (%) (OFERR)	Classification time in (s) (CT)	Dimension reduction feature vector (DR_{FV})
CEGPCA	82.35	1.012	147
CEGICA	85.03	1.245	147
CEGKPCA	87.52	1.045	147
CEGFLDA	90.45	0.874	126
CEGLPP	88.08	1.010	147
CEGLFDA	93.45	0.997	147
CEGKLFDA	95.83	0.982	126
CEGKLSWFDA	97.14	0.929	105

Table 8
Performance of subspace approaches For YALE Database at $m = 5$ and $n = 8$.

Subspace approaches	Overall facial expression recognition rate in (%) (OFERR)	Classification time in (s) (CT)	Dimension reduction feature vector (DR_{FV})
CEGPCA	61.08	0.997	63
CEGICA	64.80	0.912	63
CEGKPCA	68.52	0.929	63
CEGFLDA	75.78	0.929	63
CEGLPP	72.27	0.802	63
CEGLFDA	77.15	0.797	63
CEGKLFDA	81.38	0.758	54
CEGKLSWFDA	83.84	0.745	54

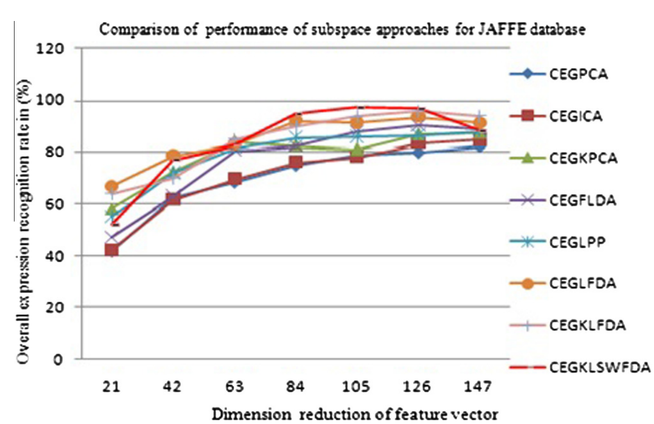

Fig. 5. Comparison of overall facial expression recognition rate for JAFFE database with feature vector dimensional reduction of different subspace approaches at $m = 5$ and $n = 8$.

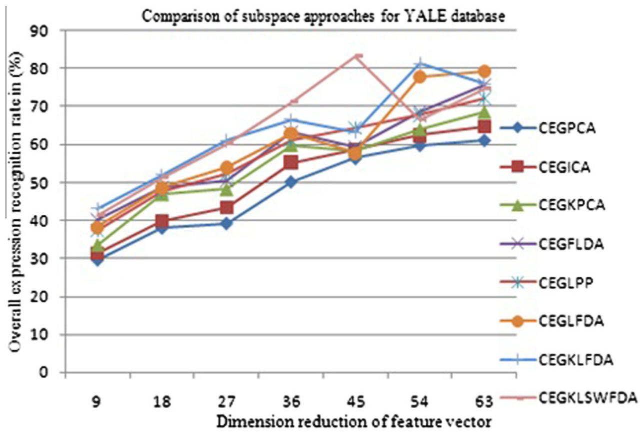

Fig. 6. Comparison of overall facial expression recognition rate for YALE database with feature vector dimensional reduction of different subspace approaches at $m = 5$ and $n = 8$.

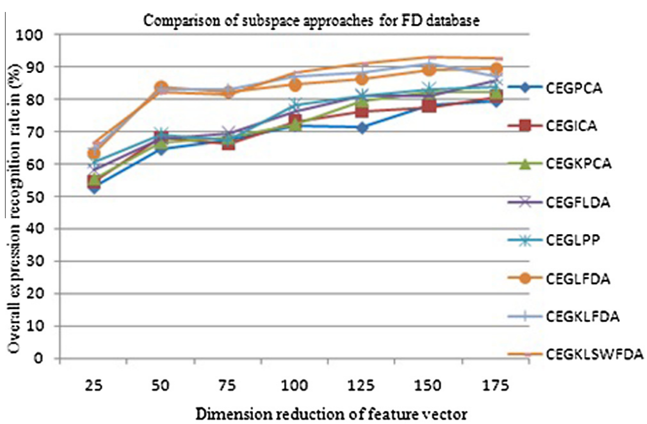

Fig. 7. Comparison of overall facial expression recognition rate for FD database with feature vector dimensional reduction of different subspace approaches at $m = 5$ and $n = 8$.

Similarly for FD database individual expression accuracy rate is presented in Table 11. From this table it has been noted that for happy and anger expression recognition rate is 96.67%, for surprise expression it is 100%. Sad and neutral expression CRR is 86.67%. This work clearly analyzes that CEGFLDA algorithm performs comparatively better than CEGPCA, CEGICA, CEGKPCA and CEGLPP approaches. It demonstrates that it is always necessary to discriminate the feature for efficient recognition using class label information. Although the CEGLFDA algorithm outperforms, CEGFLDA, algorithm approach by using both local subspace structure and class label information, it is still a linear algorithm and is inadequate to describe the nonlinear face image space due to high variability of the image content and style. Therefore it performs worse and is weaker than the kernel based KLSWFDA algorithm. Confusion matrix is derived from SVM_RBF kernel based using "Leave One Out "strategy. It demonstrates the correct and misclassification of expressions. It is also noted that classification time (CT) of proposed approach is less compared to other approaches.

Adopted proposed method in this paper can handle SSS problem or singularity problem issues of within scatter matrix. Subspace learning methods are compared with respect to their weakness and strengthening effectiveness on databases regarding higher value of recognition efficiency at specific features dimension. This feature dimension depends on how many images are considered for training. For example for YALE database 90 images are considered for experiment. Dimensionality reduction feature vector is 54 for proposed approach indicating that, for this value the highest recognition is achieved. Hence at 54 dimensional reduced feature vectors for proposed approach, it can give higher efficiency and less classification time. If the number of features goes on multiples or increases then SSS problem appears. This work estimates this problem in proposed approach. Hence each subspace approach is tested with database by considering the number of trained samples is greater than testing samples. Even in larger feature samples proposed approach yields good efficiency. Other approaches like CEGFLDA, CEGLFDA and CEGKLFDA are popular class discriminant approaches but still gives less accuracy for higher features due to incomplete solution of SSS problem. But GEGPCA, CEGICA, CEGKPCA are not class discriminative approaches. Hence it can support the SSS solution at certain extent. But due to its non discriminant nature of classification of features the recognition accuracy has been reduced. Proposed approach is made different from the approaches mentioned in the paper is

1. To solve the singularity problem issues in the presence of higher features by eliminating completely zero eigen components.

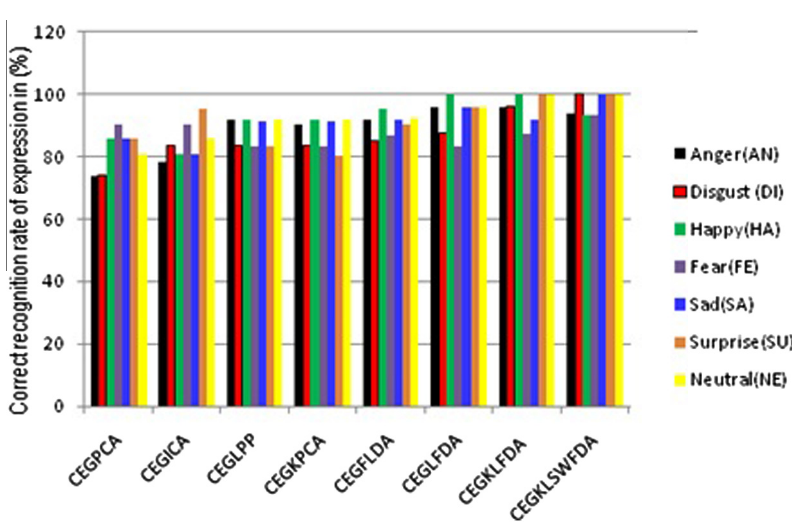

Fig. 8. Comparative analyses of subspace approaches for individual expression correct recognition rate of JAFFE database at $m = 5$ and $n = 8$.

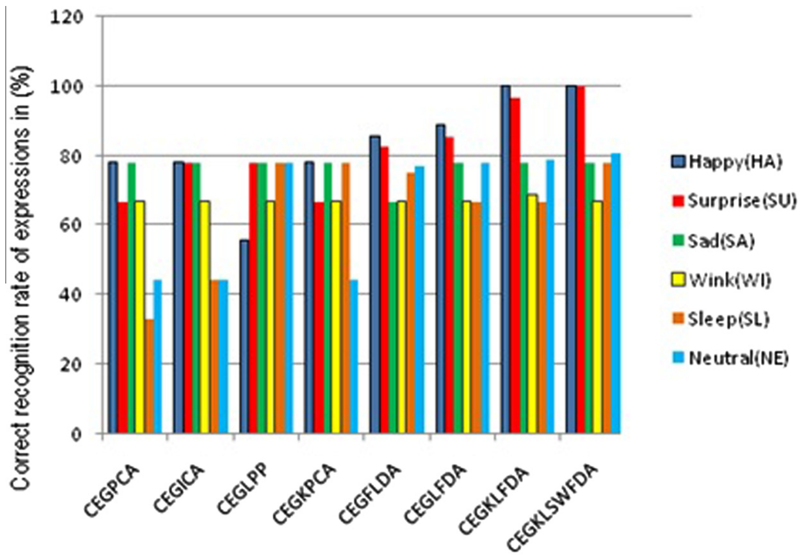

Fig. 9. Comparative analyses of subspace approaches for individual expression correct recognition rate of YALE database at $m = 5$ and $n = 8$.

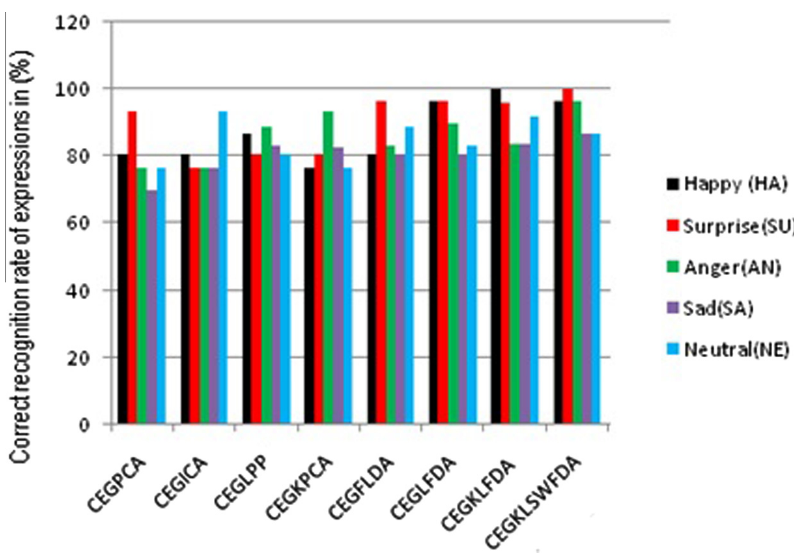

Fig. 10. Comparative analyses of subspace approaches for individual expression correct recognition rate of FD database at $m = 5$ and $n = 8$.

Table 9
Confusion Matrix of JAFFE database using proposed subspace approach using SVM, Leave One Out Technique in (%) at $m = 5$ and $n = 8$.

	AN	DI	HA	FE	SA	SU	NE
AN	93.33	6.67	0	0	0	0	0
DI	0	100	0	0	0	0	0
HA	0	0	93.33	0	6.67	0	0
FE	0	0	6.67	93.33	0	0	0
SA	0	0	0	0	100	0	0
SU	0	0	0	0	0	100	0
NE	0	0	0	0	0	0	100

2. Preservation of local image information as much as possible and achieving highly discriminative class applications. Distance between within class variables is minimized and distance between different variable classes of expressions is maximized. The Tables 12–14 illustrate the results of proposed approach at different values of texture parameters used for Gabor filter design. From this table it has been noted that for a smaller

Table 10
Confusion Matrix of YALE database using proposed subspace approach using SVM, Leave One Out Technique in (%) at $m = 5$ and $n = 8$.

	HA	SU	SA	WI	SL	NE
HA	100	0	0	0	0	0
SU	0	100	0	0	0	0
SA	0	0	77.78	0	22.22	0
WI	16.70	16.70	0	66.60	0	0
SL	0	0	0	22.22	77.78	0
NE	0	0	0	19.12	0	80.88

Table 11
Confusion Matrix of FD database using proposed subspace approach using SVM, Leave One Out Technique in (%) at $m = 5$ and $n = 8$.

	HA	SU	AN	SA	NE
HA	96.67	0	3.33%	0	0
SU	0	100	0	0	0
AN	0	0	96.67	3.33%	0
SA	0	0	0	86.67	13.33
NE	0	0	0	13.33%	86.67

Table 12
Performance of proposed subspace approach at different m and n values for JAFFE database.

Number of scales and orientations	Overall facial expression recognition rate in (%) (OFERR)	Classification time in (s) (CT)	Dimension reduction feature vector (DR_{FV})
$m = 5$ and $n = 4$	89.68	1.636	126
$m = 3$ and $n = 8$	91.76	1.611	126
$m = 3$ and $n = 4$	83.33	3.532	126

Table 13
Performance of proposed subspace approach at different m and n values for YALE database.

Number of scales and orientations	Overall facial expression recognition rate in (%) (OFERR)	Classification time in (s) (CT)	Dimension reduction feature vector (DR_{FV})
$m = 5$ and $n = 4$	80.72	0.762	54
$m = 3$ and $n = 8$	77.77	0.707	63
$m = 3$ and $n = 4$	73.77	1.107	63

Table 14
Performance of proposed subspace approach at different m and n values for FD database.

Number of scales and orientations	Overall facial expression recognition rate in (%) (OFERR)	Classification time in (s) (CT)	Dimension reduction feature vector (DR_{FV})
$m = 5$ and $n = 4$	88.76	1.078	150
$m = 3$ and $n = 8$	85.45	1.982	150
$m = 3$ and $n = 4$	78.35	2.533	150

Table 15
Comparison of performance of different approaches related to singularity problems of scatter matrix (SSS problem).

Literature	Approaches	Performance of recognition (%)
Zuo et al. [67]	BDPCA + LDA	87.14
Yang et al. [68]	YLDA	96
Yu et al. [69]	DLDA	94.9
Chen et al. [70]	CLDA	95.14
Thomaz et al. [71]	MLDA	95.81

number of m and n subspace approaches are not effective in improvement of recognition accuracy even though dimension of feature vector has been reduced.

5. Conclusions

Subspace approaches find vital role in dimensional reduction and expression recognition in several fields. Performance of expression recognition depends on face detection, feature extraction and feature vector dimension. In this work Gabor features are isolated and fused with geometrical features but not discussed in this paper. Less number of geometrical vectors having dimension 16 has been utilized for making combinational Gabor feature dataset. This geometrical feature does not affect much on feature variations due to a few geometrical feature vectors in illumination variations. Addition of geometrical features on magnitude and phase part enhances the image information but extracted feature dataset dimension is found to be too large. This work concludes that higher dimensional combinational Gabor feature vector dimension is reduced by discriminative subspace methods by preserving local discriminative structure of data by resolving the singularity problem at non linear region. Gabor magnitude and phase part vectors are having rich set of texture information, in order to utilize these features sufficiently both these vectors are fused separately with geometrical features. Higher dimension feature dataset is projected into subspace by several linear and non linear subspace methods. Proposed CEGKLSWFDA algorithm reduces the higher dimension feature dataset which has been framed by combination of Gabor filter and geometrical distance vector features and improves the expression recognition efficiency. For JAFFE dataset it has been observed that reduced dimensionality feature vector is 105 for YALE dataset it is found to be 54 and for FD database dimension reduction of feature vector is found to be 150 by achieving higher recognition rate. Unlike most of the traditional dimensionality reduction algorithms which seek the data independent nonlinear structure of the face image space, proposed algorithm explicitly considers both the intrinsic subspace structure and discriminative information. From the experimental results it has been noted that individual expression recognition rate has been improved compared to earlier subspace approaches. For JAFFE database over all accuracy rate of recognition is 97.14%, for YALE database it is 83.84% and for FD database 93.33% accuracy rate is achieved. Accuracy of proposed approach for different expression of JAFFE and YALE database also improved compared to state of art approaches. Accuracy rates for FD dataset are also increased. The results for CEGFLDA is found less compared to other approaches related to discriminative nature algorithms due to incomplete solution for singularity problem with lesser Fisher ratio value. The classification time of various expressions is also found to be less in proposed approach due to elimination of correlated data structure compared to other subspace approaches mentioned above.

References

[1] P. Ekman, W. Friesen, Unmasking the Face a Guide to Recognizing Emotions from Facial Clues, Malor Books, Cambridge, MA, 2003.
[2] B. St. George, C.C. Chang, C.J. Lin, Facial expression recognition with SMS alert, IEEE International Conference on Optical Imaging Sensor and Security (ICOSS), July 2013.
[3] Shiqing Zhang, Xiaoming Zhao, Bicheng Lei, Facial expression recognition based on local binary patterns and local fisher discriminant analysis, WSEAS Trans. Signal Process. 8 (1) (2012).

[4] I. Cohen, N. Sebe, A. Garg, L. Chen, T.S. Huang, Facial expression recognition from video sequences: temporal and static modeling, Comput. Vis. Image Underst. 91 (2003) 160–187.
[5] Yogachandran Rahulamathavan, R.C.-W. Phan, Jonathon A. Chambers, David J. Parish, Facial expression recognition in the encrypted domain based on local fisher discriminant analysis, IEEE Trans. Affective Comput. 4 (1) (2013) 83–92.
[6] Frank Y. Shih, Chao-Fa Chuang, Patrick S.P. Wang, Performance comparisons of facial expression recognition in jaffe database, Int. J. Pattern Recogn. Artif. Intell. 22 (3) (2008) 445–459.
[7] G.R.S. Murthy, R.S. Jadon, Effectiveness for eigenspaces for facial expressions recognition, Int. J. Comput. Theory Eng. 1 (3) (2000) 1793–8201.
[8] Jamal Hussain Shah, Muhammad Sharif, Mudassar Raza, Aisha Azeem, A survey: linear and nonlinear PCA based face recognition techniques, Int. Arab J. Inf. Technol. 10 (6) (2013).
[9] P. Belhumeur, J. Hespanda, D. Kiregeman, Eigenfaces versus fisherfaces: recognition using class specific linear projection, IEEE Trans. Pattern Anal. Mach. Intell. 19 (7) (1997) 711–720.
[10] Z. Li, D. Lin, X. Tang, Nonparametric discriminant analysis for face recognition, IEEE Trans. Pattern Anal. Mach. Intell. 31 (4) (2009) 755–761.
[11] Yue Ming, Qiuqi Ruan, Xiaoli Li, Meiru Mu, Efficient kernel discriminate spectral regression for 3D face recognition, Proceedings Of ICSP 2010.
[12] Dimitris Bolis, Anastasios Maronidis, Anastasios Tefas, Ioannis Pitas, Improving the Robustness of Subspace Learning Techniques for Facial Expression Recognition, Springer-Verlag, Berlin Heidelberg, 2010. pp. 470–479.
[13] A. Adeel Mohammed, M.A. Sid-Ahmed Rashid Minhas, Q.M. Jonathan Wu, A face portion based recognition system using multidimensional PCA, IEEE 54th International Midwest Symposium on Circuits and Systems, pp. 1–4, August 2011.
[14] T. Abegaz, J. Adams, K. Bryant, G. Dozier, K. Popplewell, K. Ricanek, J. Shelton, D.L. Woodard, Hybrid GAs for Eigen-based facial recognition, IEEE Workshop on Computational Intelligence in Biometrics and Identity Management, pp. 127–130, April 2011.
[15] Chengliang Wang, LibinLan, MinjieGu, Yuwei Zhang, Face recognition based on principle component analysis and support vector machine, 3rd International Workshop on Intelligent Systems and Applications, pp. 1–4, May 2011.
[16] H. Ling, S. Soatto, N. Ramanathan, D. Jacobs, Face verification across age progression using discriminative methods, IEEE Trans. Inf. Forensic Secur. 5 (1) (2010) 82–91.
[17] J. Li, B. Zhao, H. Zhang, Face recognition based on PCA and LDA combination feature extraction, The 1st International Conference on Information Science and Engineering, 2009, pp. 1240–1243.
[18] X. Wang, X. Tang, Dual-space linear discriminant analysis for face recognition, in: Proc. IEEE Conference on Computer Vision and Pattern Recognition, 2004, pp. 564–569.
[19] X. Wang, X. Tang, Random sampling LDA for face recognition, in: Proc. IEEE Conf. Computer Vision Pattern Recognition, pp. 259–265, 2004.
[20] Jiadong Song, Xiaojuan Li, Xu Pengfei, Mingquan Zhou, A new face recognition framework: symmetrical bilateral 2DPLS plus LDA, J. Multimedia 6 (1) (February 2011).
[21] M. Loog, R.P.W. Duin, R. Hacb-Umbach, Multi-class linear dimension reduction by weighted pair wise Fisher criteria, IEEE Trans. Pattern Recognit. Mach. Intell. 23 (7) (2001) 762–766.
[22] E.K. Tang, P.N. Suganthan, X. Yao, A.K. Qin, Linear dimensionality reduction using relevance weighted LDA, Pattern Recognit. 38 (4) (2005) 485–493.
[23] W. Li, S. Prasad, J.E. Fowler, L.M. Bruce, Locality preserving dimensionality reduction and classification for hyperspectral image analysis, IEEE Trans. Geo-sci. Remote Sens. 50 (4) (2012) 1185–1198.
[24] X. He, P. Niyogi, Locality preserving projections, in: Advances in Neural Information Processing Systems, MIT Press, Cambridge, MA, 2003.
[25] Lin Kezheng, Lin Sheng, Chen Dongmei, Enhanced locality preserving projections, CSSE 1 (2008) 985–988.
[26] Xiaofei He, Shuicheng Yan, Hu Yuxiao, Partha Niyogi, Hong-Jiang Zhang, Face recognition using Laplacianfaces, IEEE Trans. Pattern Anal. Mach. Intell. 27 (3) (March 2005).
[27] Yi Jin, Qiu-Qi Ruan, An image matrix compression based supervised LPP for face recognition, International Symposium on Intelligent Signal Processing and Communication Systems, 2007.
[28] D. Guangjun, Z. Yongsheng, J. Song, Dimensionality reduction of hyperspectral data based on ISOMAP algorithm, in: Proceedings of the 8th International Conference on Electronic Measurement and Instruments (ICEMI'07), pp. 3935–3938, Xi'an, China, August 2007.
[29] Shaowei Wang, Hongyu Yang, Haiyun Li, Facial expression recognition based on incremental isomap with expression weighted distance, J. Comput. 8 (8) (2013).
[30] J. Kong, S. Wang, J. Wang, L. Ma, B. Fu, Y. Lu, A novel approach for face recognition based on supervised locality preserving projection and maximum margin criterion, in: Proceedings of the International Conference on Computer Engineering and Technology (ICCET'09), vol. 1, pp. 419–423, Singapore, January 2009.
[31] M. Kirby, L. Sirovich, Application of the Karhunen–Lóeve procedure for the characterization of human faces, IEEE Trans. Pattern Anal. Mach. Intell. 12 (1) (1990) 103–108.
[32] Yu Hua, Jie Yang, A direct LDA algorithm for high dimension data with application to face recognition, Pattern Recognit. 34 (10) (2001) 2067–2070.
[33] Anil Jain, Karthik Nandakumar, Arun Ross, Score normalization in multimodal biometric systems, Int. J. Pattern Recognit. 38 (2005) 2270–2285 (Elsevier Publishers).
[34] J. Yang, A.F. Franji, J.Y. Yang, D. Zhang, Z. Jin, KPCA plus LDA: a complete kernel Fisher discriminant framework for feature extraction and recognition, IEEE Trans. Pattern Anal. Mach. Intell. 27 (2) (2005) 230–244.
[35] Y. Tian, Evaluation of face resolution for expression analysis, CVPR Workshop on Face Processing in Video, 2004.
[36] Y. Tian, T. Kanade, J. Cohn, Handbook of Face Recognition, Facial Expression Analysis, Springer, 2005 (Chapter 11).
[37] C. Liu, H. Wechsler, Gabor feature based classification using the enhanced Fisher linear discriminant model for face recognition, IEEE Trans. Image Processing 11 (4) (2002) 467–476.
[38] Muzammil Abdulrahman, Tajuddeen R. Gwadabe, Fahad J. Abdu, Alaa Eleyan, Gabor wavelet transform based facial expression recognition using PCA and LBP, in: Signal Processing and Communications Applications Conference and Communications Applications, 22nd, IEEE, 2014, pp. 2265–2268.
[39] Vitomir Struc, Nikola Pavesic, The complete Gabor-Fisher classifier for robust face recognition, EURASIP J. Adv. Signal Processing (2010).
[40] M. Sugiyama, Dimensionality reduction of multimodal labeled data by local Fisher discriminant analysis, J. Mach. Learn. Res. 8 (2006) 1027–1061.
[41] C.W. Hsu, C.C. Chang, G.J. Lin, A practical guide to support vector classification [Online], Available: <http://www.csie.ntu.edu.tw/cjlin/papers/guide/guide.pdf> (2009).
[42] Shuaishi Liu, Ning Ding, Mao Yang, Keping Liu, Facial expression recognition method based on the fusion of geometry features and texture features, Int. Conf. Electr. Comput. Eng., Adv. Biomed. Eng. 11 (2012).
[43] Hasimah Ali, Muthusamy Hariharan, Sazali Yaacob, Abdul Hamid Adom, Facial emotion recognition using empirical mode decomposition, Expert Syst. Appl. 42 (2015) 1261–1277 (Elsevier Publications).
[44] Qingshan Liu, Xiaoou Tang, Hanqing Lu, Songde Ma, Kernel Scatter-difference based discriminant analysis for face recognition, Proceedings of the 17th International Conference on Pattern Recognition (ICPR'04), IEEE Conference 2004.
[45] Ashok Rao, S. Noushath, Subspace Methods for Face Recognition – Survey, Elsevier, 2010 (Computer Science Review by Science Direct, pp. 1–17, Elsevier Series).
[46] Guoxia Sun, Liangliang Zhang, Face recognition based on symmetrical weighted PCA, IEEE Conference, pp 2249–2252, 2011.
[47] Yan Wang, Weijun Gong, Facial expression recognition based on local sensitive feature extraction with Gabor wavelet and LE+LDA, J. Comput. Inf. Syst. 7 (8) (2011) 2745–2752. Available at <http://www.Jofcis.com>.
[48] G. Baudat, F. Anouar, Generalized discriminant analysis using a kernel approach, Neural Comput. 12 (10) (2000) 2385–2404.
[49] Ziqiang Wang, Xia Sun, Lijun Sun, and Yuchun Huang, Semisupervised kernel marginal Fisher analysis for face recognition, Sci. World J., vol. 2013.
[50] Ziqiang Wang, Xia Sun, Manifold adaptive kernel local discriminant analysis for face recognition, J. Multimedia 7 (6) (2012).
[51] Z. Wang, Q. Ruan, Facial expression recognition based orthogonal local fisher discriminant analysis, Proceedings of the International Conference on Signal Processing (ICSP); Beijing, China, 24–28 October 2010, pp. 1358–1361.
[52] Hong-Bo Deng, Lian-Wen Jin, Li-Xin Zhen, Ian-Cheng Huang, A new facial expression recognition method based on local Gabor filter bank and PCA plus LDA, Int. J. Inf. Technol. 11 (11) (2005).
[53] F.Y. Shih, C.-F. Chuang, P.S.P. Wang, Performance comparisons of facial expression recognition in JAFFE database, Int. J. Pattern Recognit. Artif. Intell. 22 (3) (2008) 445–459.
[54] Shi. Dongcheng, Cai. Fang, Facial expression recognition based on Gabor wavelet phase features, Image (2013) 520–523.
[55] Gang Bai, Facial expression recognition based on fusion features of LBP and Gabor with LDA, 2nd International Congress on Image and Signal Processing, E-ISBN: 978-1-4244-4131-0, pp. 1–5, 2009.
[56] R. Zhi, Q. Ruan, Facial expression recognition based on two dimensional discriminant locality preserving projections, Neuro Comput. 71 (2008) 1730–1734.
[57] Z. Zhang, M. Lyons, M. Schuster, S. Akamatsu, Comparison between geometry based and Gabor wavelet based facial expression recognition using multi layer perceptron, (Proceeding 3rd International Conference on Automatic Face and Gesture Recognition, Nara, Japan, 1998), pp. 454–459.
[58] W. Liejun, Q. Xizhong, Z. Taiyi, Facial expression recognition using improved support vector machine by modifying kernels, Inf. Technol. 8 (4) (2009) 595–599.
[59] L. Zhao, G. Zhuang, X. Xu, Facial expression recognition based on PCA and NMF, Proceeding of the 7th World Congress on Intelligent Control and Automation, Chongqing, China, 2008, pp. 6822–6825.
[60] Chien-Cheng Lee, Shin-Sheng Huang, Cheng-Yuan Shih, Facial affect recognition using regularized discriminant analysis-based algorithms, EURASIP J. Adv. Signal Process. (2010).
[61] http://www.kasrl.org/jaffe.html.
[62] S.L. Happy, Aurobinda Routray, Automatic facial expression recognition using features of salient facial patches, IEEE Trans. Affective Comput. 6 (1) (Jan 2015).
[63] Stefanos Eleftheriadis, Ognjen Rudovic, Maja Pantic, Discriminative shared gaussian processes for multiview and view-invariant facial expression recognition, IEEE Trans. Image Process. 24 (1) (2015) 189–204.

[64] Mahmoud A.M. Albreem, Mahmoud A.M. Albreem, Shahrel A. Suandi, A novel computation technique for single and multiface detection using Euclidean distance and principal component analysis (pca), J. Eng. Sci. Technol. 7 (5) (2012) 601–613.
[65] Eu T Sun Chin, Wei Jen Chew, Florence Choong, Automated attendance capture and tracking system, J. Eng. Sci. Technol., EURECA 2014 (Special Issue January) (2015) 45–59.
[66] Abduljalil Radman, Nasharuddin Zainal, Mahamod Ismail, Efficient iris segmentation based on eyelid detection, J. Eng. Sci. Technol. 8 (4) (2013) 399–405.
[67] Wangmeng Zuo, Kuanquan Wang, Hongzhi Zhang, Subspace methods for face recognition: singularity, regularization, and robustness, in: I. Mario Dr., M. Chacon (Eds.), State of the Art in Face Recognition, I-Tech, Vienna, Austria, 2009. ISBN-3-902613-42-4, p. 250.
[68] J. Yang, J. Yang, Why can LDA be performed in PCA transformed space?, Pattern Recognit 36 (2003) 563–566.
[69] H. Yu, J. Yang, A direct LDA algorithm for high dimensional data – with application to face recognition, Pattern Recognit. 34 (2001) 2067–2070.
[70] L. Chen, H. Liao, M. Ko, J. Lin, G. Yu, A new LDA-based face recognition system which can solve the small sample size problem, Pattern Recognit. 33 (10) (2000) 1713–1726.
[71] C.E. Thomaz, D.F. Gillies, R.Q. Feitosa, A new covariance estimate for Bayesian classifiers in biometric recognition, IEEE Trans. Circuits Syst. Video Technol. 14 (2) (2004) 214–223 (Special Issue on Image- and Video-Based Biometrics).