Accuracy Analysis of Expression Recognition Rates Using Subspace Based Approaches

G.P.Hegde[1], M. Seetha[2], Nagaratna Hegde[3],

[1]Department of Computer Science, SDMIT, Ujire, India,
Affiliated to VTU, Belgaum
[2]Department of Computer Science, GNITS, Hyderabad, India,
Affiliated to JNTU, Hyderabad
[3]Department of Computer Science, VCE, Hyderabad, India,

Abstract— *This paper proposes a new approach for facial expression recognition by extracting and concatenating holistic and geometrical face features. High dimension feature dataset has been reduced by symmetrical weighted kernel locality preserving Fisher discriminant analysis (SWKLFDA). The main objectives of this research work are solving singularity problem of linear discriminant analysis, maximizing the Fishers ratio in nonlinear region of subspace by preserving local discriminant structure by increasing the variability between inter class scatter matrix, dimensional reduction and efficient recognition at larger variations of expressions and illumination. Extracted geometrical features were fused with Gabor magnitude and enhanced Gabor phase part. These isolated feature vector spaces were projected into subspace by SWKLFDA subspace method by enhancing and preserving the kernel discriminant features. Matching score level fusion was applied on projected subspace. Euclidean distance metric (L2) and support vector machine (SVM) classifier was implemented to classify the expressions. Performance analysis was carried out by comparing state of art approaches with proposed approach. Experimental results on JAFFE and YALE expression database demonstrate the effectiveness of the proposed approach.*

Key words— **Discriminant analysis, Gabor filter, Expression recognition, Feature extraction, Subspace. Symmetrical weight.**

1. Introduction

Several face appearance situations makes an omission of expressions like happy, sad, angry, fatigue, confusion, surprise, thinking, fear, pain, wink, fun and disgust etc. 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 can be determined through the expressions. Driver feelings can be determined during vehicle driving to avoid the accidental hazards. In real time election voting systems to know the persons correct identity.

Ekman et. al. (2003) [1] was carried out study of expressions and observed that at least six expressions such as anger, disgust, fear, happiness, sadness and surprise are exhibited by human beings and neutral state is considered a normal expression. In pattern recognition, computer vision and biometrics, facial expression recognition is one of the most challenging tasks due to larger variations of illuminations, noisy environments [1-4]. There are many applications finds in psychological studies, medical diagnosis, during different painful situations, determination of human emotional states for criminal and security issues [4]. Generally, there are two categories of feature extraction approaches: holistic feature based appearance approach and analytical feature based geometrical approach. To synthesize a complete image face under appearance based approach, both shape and textures were found significant domain in several studies [5-7].

Most of the LDA extension methods or algorithms were proposed in earlier studies unable to optimize the singularity matrix problems. This problem was resolved by Belhumeur et. al. [9] and proposed a Fisherface method (FF) in 1997, which uses a principal component analysis (PCA), based projection and a change of matrix size so that the matrix is nonsingular. But still for Fisher LDA algorithm it is needed to improve the singularity problem scenarios. In this paper this problem was considered as one of the main issue. 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) it is based on regression and spectral graph analysis. They have suggested that when the sample vectors are non-linear SRKDA can efficiently give better solutions than ordinary subspace learning approaches. Wang et. al [50] proposed Semi supervised kernel marginal Fisher analysis (SKMFA) in which author 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 Local Fisher discriminant analysis(LFDA). Author was suggested there was a challenge to work with encrypted domain even if there was not good recognition rate for unencrypted domain. This method was applied to JAFFE database and achieved 94.37% recognition rates. There are several researchers implemented subspace projection methods directly on input images to achieve feature extraction and dimension reduction. In [46] different earlier subspace methods were implemented on feature dataset for dimensional reduction and compared the weakness of subspace methods. From literature survey it was 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], linear discriminant analysis based and Fisher LDA [16-22], locality preserving projection (LPP) [23-27] are linear approaches. Nonlinear approaches include isomap mapping (Isomaps) [28 -29], KPCA [34], KFDA [45], KLFDA [44]. The common drawback of nonlinear embedding methods is that these approaches are more time consuming to compute high dimensional feature datasets. Yu et. al [32] proposed a direct LDA algorithm for face recognition which incorporates the concept of null space for high dimensional data. Mlakar and Potočnik [62] used histogram of oriented gradient (HOG) descriptor for expression recognition. They considered differences on a level of feature descriptors between a neutral expression and a peak expression of an observed person. SVM classifier was implemented using leave-one-subject-out validation strategy and mean recognition rate of 95.64 % was obtained using the Cohn Kanade database. Zhalehpou et al. [63] presented automatic multmodal emotion recognition system which is based on selection of peak frames with suitable vidoe channel. Mohammadian et al. [64] has been proposed a new approach to address the person dependent problem by employing the person's neutral samples as prior knowledge and a synthesis method based on the subspace learning to generate virtual expression samples. These samples have been incorporated in learning process to learn the style of the new person. Their results shows that task of facial expression recognition have been carried out effectively with better robustness for corrupted data.

A complete kernel Fisher discriminant framework for feature extraction and recognition using KPCA and LDA is proposed in [34]. Few of LDA based approaches are listed in Table 3. The rest of the paper is organized as follows: section 2 presents proposed framework by comparing related earlier works. In section 3, results and discussions are made. In section 4, Conclusions are drawn.

2. Proposed Framework

This paper mainly focuses and illustrates on projection of high dimensional image space into low dimensional subspace by solving the singularity problems of linear discriminant analysis. This task was carried by proposing symmetrical weighted kernel locality preserving Fishers discriminant analysis algorithm (SWKLFDA). In the beginning of this expression recognition system face detection has been carried out and trained database was created from still raw images by resizing the images as per the previous work as given in [43]. Feature extraction was carried by implementing Gabor filter e as given in [40]. Both Gabor magnitude and phase parts were isolated and features were extracted separately. Then combinational entire Gabor feature dataset was formed by fusing the Gabor magnitude feature vector and Gabor phase vector with geometrical feature vector (from 18 fiducially created) as presented in Fig. 1 for equal distribution of feature dimension. These two vectors were named as combinational Gabor magnitude vector (CGMV) and combinational Gabor phase vector (CGPV). These vectors size were found to be large in dimension and projected into subspace by applying SWKLFDA algorithm. Both these projected subspaces similar matrix were fused by matching score level fusion as introduced in [33]. Using Euclidean distance metric (L2) and SVM [42] classifier technique expressions were classified. Gabor filter was constructed with 5 scales and 8 orientations

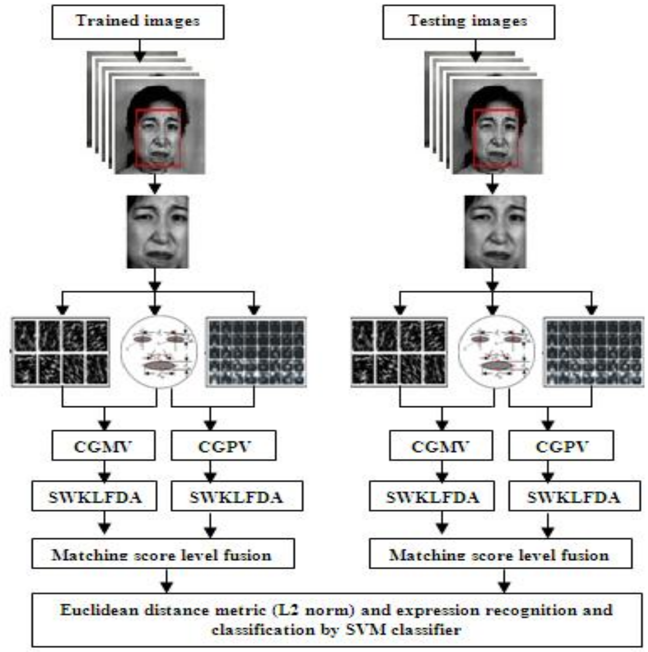

Fig. 1 Schematic diagram of expression recognition system using proposed approach.

[40]. Geometrical feature were extracted as per the procedure mentioned in [43]. In general Gabor wavelets which have been widely used in recognition of several objects like face, palm, skin etc. In face recognition Gabor filter extracts rich texture content from face images. In general, Gabor filters are also called Gabor wavelets [40] designed with different scales and orientations as mentioned in table1. Gabor magnitude information can capture the facial structure and phase information can give a detailed description of facial texture.

2.1 Brief Description about SWKLFDA

In order to solve the singularity problems and unequal distribution of eigen components of PCA algorithm in nonlinear regions for efficient recognition of expressions, KLFDA algorithm [51] is enhanced and proposed as symmetrical weighted KLFDA with a kernel trick as defined in [45] and symmetrical weights principles described as in [47]. This algorithm is designed by combining the properties of symmetrically weighted PCA based kernel FLDA and locality preserving projection properties. LPP mentioned in [27] is a unsupervised linear subspace method which preserves the local structure of neighboring data samples in the original space are kept close in the locality preserving projected embedded space. FLDA based LPP preserves neighborhood relationships in the embedding by implementing a affinity matrix as given in (18) that presented below.

In this section kernel Fishers discriminant analysis (KFDA) finds projection *w* of vectors in a higher dimensional kernel domain space such that it maximizes Fisher's ratio in that space. The idea of KFDA is to solve the problem of FLDA in an implicit feature space *F* constructed by a nonlinear mapping as in (1). Feature space *F* of Gabor face image dataset in kernel region can be defined as

$$\phi(g)=[\phi(g_1),\phi(g_2)\cdots,\phi(g_n)] \quad (1)$$
$$\phi : g \in R^N \rightarrow \phi(g) \in F$$

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 Gabor feature dataset set with n elements and C classes, and i_n represents the number of samples in the i-th 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

$$J(W)=\underset{w}{\arg\max}\frac{w^T S_b^{\phi} w}{w^T S_w^{\phi} w} \quad (2)$$

Fisher ratio can be increased by increasing the variance between the inter class or between class matrix in kernel region. The value of S_b^{ϕ} is made high by applying symmetrical weighted PCA for equalizing the variance of principal components in order to solve the singularity problems. Symmetrical weighted PCA representation is made in [47]. In this work odd-even rule is implemented as given in [47] to decompose face image . Entire Gabor dataset images are $g_i=[g_1,g_2,\cdots g_M]$ mirror symmetrical 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,…..M. Odd symmetrical sample set ($g_{o1}, g_{o2}, g_{o3}, \ldots, g_{oM}$) and even symmetrical sample ($g_{e1}, g_{e2}, g_{e3}, \ldots g_{eM}$) set both are derived from original training sample set by mirror symmetrical transform. Odd and even set can be defined as

$$S_b^{\phi}=\sum_{i=1}^{C} N_i(m_i^{\phi}-m^{\phi})(m_i^{\phi}-m^{\phi}) \quad (3)$$

$$S_{ob}^{\phi}=\sum_{i=1}^{C} N_i(m_{oi}^{\phi}-m_o^{\phi})(m_{oi}^{\phi}-m_o^{\phi}) \quad (4)$$

$$S_{eb}^{\phi}=\sum_{i=1}^{C} N_i(m_{ei}^{\phi}-m_e^{\phi})(m_{ei}^{\phi}-m_e^{\phi}) \quad (5)$$

Where, $S_b^{\phi}=S_{ob}^{\phi}+S_{eb}^{\phi}$, hence the eigen value decomposition on S_b^{ϕ} is equal to the eigen decomposition on S_{ob}^{ϕ} and S_{eb}^{ϕ}. Hence, image g_i can be reconstructed by the feature vector of S_{ob}^{ϕ} and S_{eb}^{ϕ}. With respect to eigen theory assume all the non-zero eigen values of S_{ob}^{ϕ} and S_{oe}^{ϕ} are λ_{oi} and λ_{ej}, and the corresponding eigen vectors are w_{oi} and w_{ej}. Where i=1…. rank (S_{ob}^{ϕ}) and j=1. . . rank(S_{oe}^{ϕ}). Transformation of weight matrix for odd (T_o) and even (T_e) symmetrical sample sets be derived from above demonstration as

$$T_o=[w_{o1,}w_{o2}\cdots w_{or_o}],\ \wedge_o=diag(\lambda_{o1},\lambda_{o2},\cdots\lambda_{or_o}) \quad (6)$$
$$T_e=[w_{e1,}w_{e2}\cdots w_{er_e}],\ \wedge_e=diag(\lambda_{e1},\lambda_{e2},\cdots\lambda_{er_e}) \quad (7)$$

Where r_o= rank(S_{ob}^{ϕ}), r_e=rank(S_{oe}^{ϕ})

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

$$g_{oi}=T_oP_{oi,}\ \ P_{oi}=T_o^t g_{oi}\quad g_{ei}=T_eP_{ei}\ \ P_{ei}=T_e^t g_{ei} \quad (8)$$

In above, P_{oi} and P_{ei} are the odd symmetrical feature and even symmetrical feature of the ith face Gabor image. In order to reduce the effects made by the principal components which contain the variation due illumination or face expression, it can treat as 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:

$$Q_o=\wedge_o^{-1/2}T_o^t=(\lambda_{o1}^{-1/2}w_{o1},\lambda_{o2}^{-1/2}w_{o2}\cdots\lambda_{or_o}^{-1/2}w_{or_o}) \quad (9)$$

$$Q_e=\wedge_e^{-1/2}T_e^t=(\lambda_{e1}^{-1/2}w_{e1},\lambda_{e2}^{-1/2}w_{e2}\cdots\lambda_{er_e}^{-1/2}w_{er_{oe}}) \quad (10)$$

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

$$g_{oi}=Q_oz_{oi}\quad z_{oi}=Q_o^tg_{oi}\quad g_{ei}=Q_ez_{ei}\quad z_{ei}=Q_e^tg_{ei} \quad (11)$$
$$z_i=[z_{ei}^t,z_{oi}^t]^t=\begin{bmatrix} z_{ei} \\ z_{oi}\end{bmatrix}\quad Q=[Q_e,Q_o]\quad \wedge=diag[\wedge_e,\wedge_o] \quad (12)$$
$$g_i=Qz_i=[Q_e,Q_o]\begin{bmatrix} z_{ei} \\ z_{oi}\end{bmatrix}\ z_i=Q^tg_i \quad (13)$$

For feature selection in Symmetrical weighted PCA [47], sorting the eigenvalues either in ascending or descending order, then largest eigen vectors were selected corresponding to 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. So 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 were eliminated. Inter-class kernel space scatter matrix is considered from above (1) and within class kernel space scatter matrix is given as

$$S_w^{\phi} = \sum_{i=1}^{C} \sum_{g \in G_i} (\phi(g) - m_i^{\phi})(\phi(g) - m_i^{\phi}) \tag{14}$$

Where G_i is the number of samples from the i^{th} class, m_i^{ϕ} is the centroid of the i^{th} class, C is the number of classes, m_i^{ϕ} is the centroid of the i^{th} class 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 i^{th} class. S_w^{ϕ} represents the degree of 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 trick" [23] 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. Sugiyama et. al. [41] introduced local Fisher discriminant analysis (LFDA) and author have shown that the effect of properties of combination of both LDA and LPP. LFDA preserves neighborhood relationships in the embedding by employing an "affinity" matrix that is defined below. The optimization solution J(W), corresponding to the largest eigenvalues λ, can be illustrated by the generalized eigenvalue problem.

$$S_b^{\phi} W_i = \lambda_i S_w^{\phi} W_i \tag{15}$$

Let us consider set of N Gabor images, that is $\{g_1, g_2, g_3, \ldots, g_n\}$ taking in n-dimensional image space and assume that each image belongs to one of c classes $\{C_1, C_2, C_3, \ldots C_c\}$. Let N_i be the number of images in class C_i where (i=1,2,3,. . . .C),

$$\mu_i = \frac{1}{N_i} \sum_{G \in C_i} G \tag{16}$$

Here μ_i be the mean of the samples in class and

$$\mu = \frac{1}{N} \sum_{i=1}^{N} G_i \tag{17}$$

Where μ be the mean of all samples. Define $S_{i,j} \in [0,1]$ as the affinity between g_i and g_j given by

$$S_{i,j} = e^{-\frac{\|g_i - g_j\|^2}{\gamma_i \gamma_j}} \tag{18}$$

$$\gamma_i = \left\| g_i - g_i^{(k)} \right\| \tag{19}$$

Local scaling of data g_i and $g_i^{(k)}$ is the k^{th} nearest neighbor of g_i as presented in (19). Symmetric matrix is $S_{i,j}$ also called as affinity matrix of size nxn which measures the local distance between the data samples in the input space. Local between class $S^{(lb)}$ and local within class $S^{(lw)}$ scatter matrices are defined as

$$S^{lb} = \frac{1}{2} \sum_{i,j=1}^{n} W_{i,j}^{(lb)} (g_i - g_j)(g_i - g_j)^T \tag{20}$$

$$S^{lw} = \frac{1}{2} \sum_{i,j=1}^{n} W_{i,j}^{(lw)} (g_i - g_j)(g_i - g_j)^T \tag{21}$$

Where $W^{(lb)}$ and $W^{(lw)}$ are nxn matrices defined as

$$W_{i,j}^{(lb)} = \begin{cases} S_{i,j}(1/n - (1/n_l)), if\,(y_i = y_j = l) \\ 1/n, \qquad if\,(y_i \neq y_j) \end{cases} \tag{22}$$

$$W_{i,j}^{(lw)} = \begin{cases} S_{i,j}/n_l, if\,(y_i = y_j = l) \\ 0, \qquad if\,(y_i \neq y_j) \end{cases} \tag{23}$$

When these modified scatter matrices are employed in (14), optimizing the modified Fisher ratio results in an LFDA formulation. The weights defined in (22) and (23) give LFDA its neighborhood pixel preserving properties. The KLFDA algorithm can be viewed as a kernel extension of LFDA via the kernel trick. In this letter, the kernel function employed is the radial basis function (RBF) kernel [34], defined as

$$K(g_i, g_j) = e^{-\frac{\|g_i - g_j\|^2}{2\sigma^2}} \tag{24}$$

Where $\sigma > 0$ is a user-defined parameter of the kernel. Sugiyama et. al. [41] invokes the kernel trick and reformulates the LFDA algorithm in kernel-induced spaces. In other words, the local within and between class scatter matrices are defined in the kernel domain space. Projection $\hat{w}$ in the kernel domain space that maximizes the modified Fisher ratio is given by the solution of the generalized eigenvalue problem, i.e.

$$KL^{(lb)} K \hat{w} = \bar{\Lambda}(KL^{(lw)} K + \varepsilon I_n) \hat{w} \tag{25}$$

Here $\hat{\Lambda}$ is the diagonal eigenvalue matrix; ε is a small (regularization) constant, $\hat{w}$ is the eigenvector matrix, K is the kernel matrix defined in (41), $L^{(lw)} = D^{(lw)} - W^{(lw)}$, where $D^{(lw)}$ is a diagonal matrix with the i^{th} diagonal element becomes

$$D_{ii}^{(lw)} = \sum_{j=1}^{n} W_{ij}^{(lw)} \tag{26}$$

Like above $L^{(lb)} = L^{(lw)} - L^{(lw)}$, where $L^{(m)}$ is the local mixture matrix defined as $L^{(m)} = D^{(m)} - W^{(m)}$, and $D^{(m)}$ is a diagonal matrix with the i^{th} diagonal element becomes

$$D_{ii}^{(m)} = \sum_{j=1}^{n} W_{ij}^{(m)} \quad (27)$$

2.2 Matching Score Level Fusion

In the proposed approach Z-score normalization and fusion technique is used by considering the parameters of paper [33]. Experiments of Jain et al.[33], also reveal that the min–max and Z-score normalization techniques are sensitive to outliers in the data, highlighting the need for a robust and efficient normalization procedure like the tanh normalization.

$$NS_{CGMSWKLFDA} = \frac{CGMSWKLFDA_S - \mu(CGMSWKLFDA_S)}{Std(CGMSWKLFDA_S)} \quad (28)$$

$$NS_{CGPSWKLFDA} = \frac{CGPSWKLFDA_S - \mu(CGPSWKLFDA_S)}{Std(CGPSWKLFDA_S)} \quad (29)$$

$$CEGSWKLFDA_S = Max[(NS_{CGMSWKLFDA} + NS_{CGPSWKLFDA})] \quad (30)$$

Where $(CGMSWKLFDA)_S$ is similarity score matrix of Gabor magnitude symmetrical weighted kernel locality preserved Fisher discriminant projected subspace and $(CGPSWKLFDA)_S$ is similarity score matrix of Gabor enhanced phase part subspace. Local structure of LPP is preserved by kernel locality Fisher discriminant analysis. The main concept of LPP is to optimize the maximum locality preservation structures. In other words, the objective function that is minimized in LPP is designed such that it expenses a larger penalty, if neighboring points that are close in the input space are mapped far apart in the projection as shown in (18), (20), and (21). An adjacency or affinity matrix identical to that in (18) is used in this formulation to preserve local neighborhood relationships.

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

$$\varepsilon_i^2 = \left\| W_{CEGSWKLFDAQ} - W_{CEGSWKLFDAT} \right\|^2 \quad (31)$$

Where $W_{CEGSWKLFDAT}$ and $W_{CEGSWKLFDAQ}$ are projected vector final score weight matrices of training and testing Gabor dataset images. If ε_i is less than some predefined threshold value θ_i, then test image belongs to class i. So that testing expression image is matched with trained image. Based on Euclidean distance and RBF kernel based SVM classifier [33] facial expressions were classified.

3. Results and Discussions

3.1 Database used

The experiments were performed in order to analyze the performance of the proposed approach on two public databases as given below. Proposed approach was tested for JAFFE database with and without noise. But YALE database illumination variations were considered while testing the proposed approach.

3.2 JAFFE database

In this work, Japanese Female Facial Expression (JAFFE) database was used for experiment [43]. Figure 2 shows cropped samples of JAFFE database. Total 210 images were cropped into 111x126 size. Only required areas like mouth nose, eyes and chin areas has been considered during face detection for extraction of texture features and rest of the part was removed.

Fig. 2 Preprocessed and resized samples of JAFFE database

3.3 YALE database

YALE database contains 11 images per person for 15 individuals resulting into 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 not considered. 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 pixels size. In this work six expressions were used for experiment such as happy, surprise, sad, wink, sleep and neutral as shown in Figure 7.

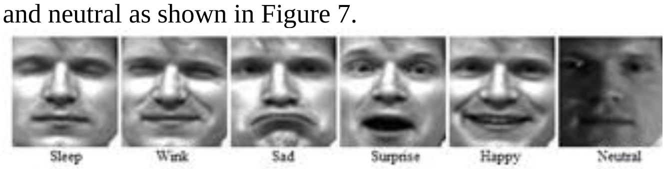

Fig 3. Resized samples of YALE database at different lights

Table 1 Gabor Filter input parameters common to two database features

Number of scales (m)	Number of orientations (n)	Gabor filter size (GF_{mn})
5	4	20
3	8	24
3	4	12
5	8	40

Table 2 Comparison of Proposed approach with state of Art Approach for YALE Database

Author	**Approaches**	**FERR**
Ziqiang Wang[50]	SKMFA	73.6%
Yan Wang [48]	Gabor+PCA+NN	86.64%
Ours	CEGSWKLFDA	83.84%

Table 3 comparison of performance of proposed and state of art approaches for JAFFE database

Author /reference no.	Approaches	FERR
Shiqing Zhang et. al. [3]	LBP based LDA	73.4% ± 5.6
Shiqing Zhang et. al. [3]	Boosted LBP based LDA	77.67 %±5.7
Wang Z et. al. [52].	Orthogonal LDA	86.33%
I. Cohen et. al. [4]	LFDA	90.70%
Hong-Bo Deng et. al.[53]	Local Gabor +PCA+LDA	97.33%
F. Y. Shih et. al.[54]	2DLDA+SVM	94.13%
Shi Dongcheng et. al.[55]	Gabor+PCA, Gabor+2DPCA	91% and 94%
Gang Bai et. al. [56]	Gabor+LBP+LDA	92% to 97%
R.Zhi and Ruan [57]	2D discriminant LPP	95.91%
Z.Zhang et.al.[58]	Multilayer Perceptron	90.34%
W.Liejun et.al.[59]	SVM based	95.7%
L.Zhao et. al. [60]	PCA and NMF	93.72%
Chien-Cheng Lee [61]	RDAB	96.67%
Ours	**CEGSWKLFDA**	**97.14%**

Table 4 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 (sec) (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 5 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 (sec) (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

Table 6 Confusion Matrix of JAFFE Database Using CEGSWKLFDA Subspace Approach

	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	93.3	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

Table 7 Confusion Matrix of YALE Database Using CEGSWKLFDA Subspace Approach

	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

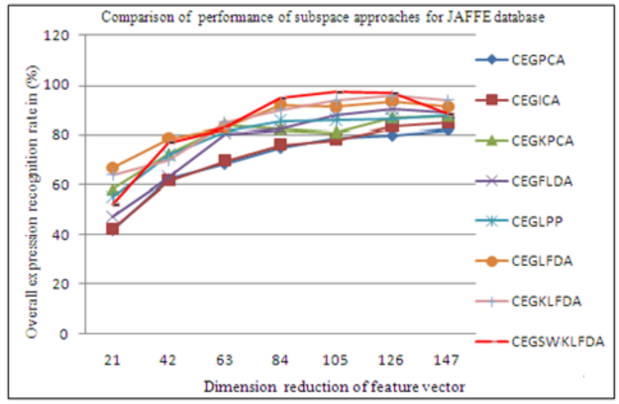

Fig. 4 Comparison of overall expression recognition rates for JAFFE database for different subspace approaches.

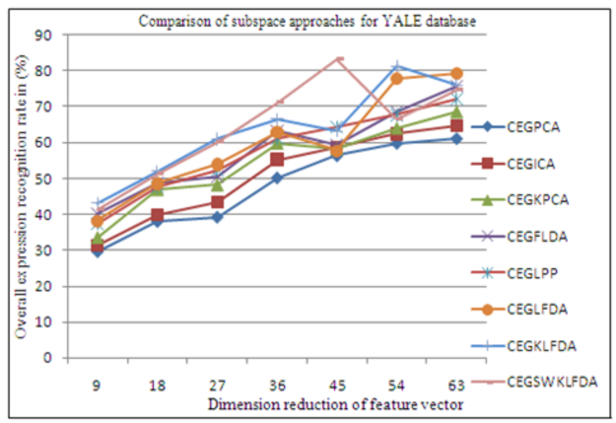

Fig. 5 Comparison of overall expression recognition rates for YALE database for different subspace approaches.

In this section to investigate the performance of proposed approach for expression recognition of JAFFE database for CEGPCA, CEGLPP, CEGFLDA, CEGLFDA approaches were compared with proposed CEGSWKLFDA approach. Similarly CEGKPCA, CEGLPP, CEGFLDA, CEGLFDA approaches were compared with proposed CEGSWKLFDA for YALE database. Over all facial expression recognition rates (FERR) of linear and nonlinear subspace methods are presented in Table 4 and Table 5 for JAFFE and YALE database respectively. These are all most useful dimensionality reduction algorithms used in face and expression recognition. For CEGLPP, CEGKPCA CEGLFDA and CEGSWKLFDA algorithms nearest neighbor number k is set to 7. Where the value of σ was set to be 0.5. The overall recognition results are shown in Fig. 4 and Fig 5 respectively. Overall expression recognition rate for JAFFE database is 97.14% and overall expression recognition rate for YALE database is 83.84% using proposed approach. From the results it has been noted that CEGSWKLFDA proposed approach consistently outperforms the CEGKPCA, CEGLPP, CEGFLDA, CEGLFDA expression recognition approaches. It was found that for JAFFE database anger, fear and happy expressions recognition rate is 93.33%. But disgusted, sad, surprise and neutral expression accuracy rate is 100% respectively as shown in Table 6. Probably it is due to confusion with sad and disgust expressions. When confusion matrix of YALE database is referred it was noted that 100% for happy and surprise expressions, sad and sleep expressions rate has been found to be 77.78%, for wink expression 66.60% accuracy rate and for neutral 80.88% accuracy has been achieved respectively as shown in Table 7. This work clearly analyzes that CEGFLDA algorithm performs comparatively CEGLPP algorithm. It demonstrates that it is hard to evaluate whether local subspace structure or class label information is more important. Although the CEGLFDA algorithm outperforms CEGKPCA, CEGFLDA, CEGPCA, CEGLPP algorithms by using both local subspace structure and class label information, it is still a linear algorithm and is in adequate to describe the nonlinear face image space due to high variability of the image content and style. Therefore it performs worse and weak than the kernel based SWKLFDA algorithm. Confusion matrix was derived from SVM_RBF kernel based using leave one out strategy. It demonstrates the correct and misclassification of expressions.

4. Conclusions

Performance of expression recognition depends on face detection, feature extraction and feature dimension. In this work Gabor face features were isolated and fused with geometrical face features but not discussed in this paper. Extracted feature dataset dimension was found to be too large. The main goal of this work is dimensional reduction of feature dataset by preserving local discriminative structure of data by resolving the singularity problems. Gabor magnitude and phase part vectors are having rich set of texture information, in order to utilize these features sufficiently both the vectors were fused separately with geometrical features. Higher dimension feature dataset was projected into subspace by several linear and non linear subspace methods. In this paper only few of them are given and compared with proposed subspace approach. Proposed SWKLFDA algorithm reduces the higher dimension feature dataset which was framed by combination of Gabor filter and geometrical distance vector features and termed as CEGSWKLFDA approach. Unlike most of the traditional dimensionality reduction algorithms which seek the data independent nonlinear structure of the face image space. Proposed algorithm explicitly considers the both the intrinsic subspace structure and discriminative information. From the experimental results it was found 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% and for YALE database 83.84% was achieved. Accuracy of proposed approach for different expression of JAFFE and YALE database also have been improved compared to approaches listed in Table 2 and Table 3. The results for CEGKPCA were found less compare to other approaches.

5. References

1. Ekman P. and Friesen W., Unmasking the Face A Guide to Recognizing Emotions from Facial Clues, Cambridge MA Malor Books, CA, 2003.
2. St.George, B., Chang, C. C., Lin, C.J., “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 Transaction on Signal Processing E-ISSN: 2224-348821, Issue 1, Volume 8, January 2012.
4. I. Cohen, N. Sebe, A. Garg, L. Chen, T.S. Huang, Facial expression recognition from video sequences: temporal and static modeling, Computer Vision and Image Understanding 91, 160–187, 2003.
5. Rahulamathavan, Yogachandran, RC-W. Phan, Jonathon A. Chambers, and David J. Parish, "Facial expression recognition in the encrypted domain based on local fisher discriminant analysis," Affective Computing, IEEE Transactions on 4, no. 1, 83-92, 2013.
6. Frank Y. Shih, Chao-Fa Chuang, Patrick S. P. Wang, “Performance Comparisons of Facial Expression Recognition in Jaffe Database,” International Journal of Pattern Recognition and Artificial Intelligence Vol. 22, No. 3, 445–459, 2008.
7. GRS Murthy, R.S. Jadon, “Effectiveness for Eigenspaces for Facial Expressions Recognition,” International Journal of Computer Theory and Engineering, Vol 1, No. 3. 1793-8201, 2000.
8. Jamal Hussain Shah, Muhammad Sharif, Mudassar Raza, and Aisha Azeem “A Survey: Linear and Nonlinear PCA Based Face Recognition Techniques” The International Arab Journal of Information Technology, Vol. 10, No. 6, 2013.
9. P. Belhumeur, J. Hespanda, and D. Kiregeman, “Eigenfaces versus Fisherfaces: Recognition using class specific linear projection,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 7, pp. 711–720, Jul. 1997.
10. Z. Li, D. Lin, and X. Tang, “Nonparametric discriminant analysis for face recognition,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 31, no. 4, pp. 755–761, Apr. 2009.
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, and Ioannis Pitas, “Improving the Robustness of Subspace Learning Techniques for Facial Expression Recognition,” pp. 470–479, Springer-Verlag Berlin Heidelberg 2010.
13. A. Adeel Mohammed, M. A. Sid-Ahmed Rashid Minhas, and 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. Abegaz, T., Adams J, Bryant, K., Dozier, G., Popplewell, K., Ricanek, K., Shelton, J., and Woodard, D.L., “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, and 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, and D. Jacobs, “Face verification across age progression using discriminative methods,” IEEE Trans. Inf. Forensic Security, vol. 5, no. 1, pp. 82–91, Mar. 2010.
17. J. Li, B. Zhao, and 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 and X. Tang, “Dual-space linear discriminant analysis for face recognition,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2004, pp. 564–569.
19. X. Wang and X. Tang, “Random sampling LDA for face recognition,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognition., pp.259–265, 2004.
20. Jiadong Song, Xiaojuan Li, Pengfei Xu, Mingquan Zhou “A New Face Recognition Framework: Symmetrical Bilateral 2DPLS plus LDA,” Journal of Multimedia, Vol. 6, NO. 1, February 2011.
21. M. Loog, R. P. W. Duin and R. Hacb-Umbach, “Multi-class Linear Dimension Reduction by Weighted Pair Wise Fisher Criteria,” IEEE

Transaction on Pattern Recognition and Machine Intelligence, Vol. 23, No. 7, pp. 762-766, 2001.

22. E. K. Tang, P. N. Suganthan, X. Yao and A. K. Qin, "Linear Dimensionality Reduction Using Relevance Weighted LDA," *Pattern Recognition*, Vol. 38, No. 4, pp. 485-493,2005.
23. W. Li, S. Prasad, J. E. Fowler, and L. M. Bruce, "Locality preserving dimensionality reduction and classification for hyperspectral image analysis," IEEE Transactions on Geo-science and Remote Sensing, vol. 50, no. 4, pp. 1185–1198, 2012.
24. X. He and P. Niyogi, "Locality preserving projections, "In Advances in Neural Information Processing Systems. Cambridge, MA: MIT Press, 2003.
25. Lin Kezheng, Lin Sheng and Chen Dongmei: "Enhanced Locality Preserving Projections". CSSE (1), 985-988, 2008.
26. Xiaofei He, Shuicheng Yan, Yuxiao Hu, Partha Niyogi, and Hong-Jiang Zhang, "Face Recognition Using Laplacianfaces" IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 27, No. 3, March 2005.
27. Yi Jin and 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, and 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 and Haiyun Li, Facial Expression Recognition Based on Incremental Isomap with Expression Weighted Distance "Journal of Computers, Vol. 8, No. 8, 2013.
30. J. Kong, S. Wang, J. Wang, L. Ma, B. Fu, and 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 and L. Sirovich, "Application of the Karhunen-Lo´eve procedure for the characterization of human faces," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 12, no. 1, pp. 103–108, Jan. 1990.
32. Hua Yu and Jie Yang, "A Direct LDA Algorithm for High Dimention Data with Application to Face Recognition," Pattern Recognition, vol.34 (10), pp.2067-2070, Oct 2001.
33. Anil Jain, Karthik Nandakumar, Arun Ross, Score normalization in multimodal biometric systems, International Journal of Pattern Recognition 38 , 2270 – 2285. – ELSEVIER Publishers-2005.
34. J. Yang, A. F. Franji, J. Y. Yang, D. Zhang, and Z. Jin, "KPCA plus LDA: A complete kernel fisher discriminant framework for feature extraction and recognition," IEEE Trans. Pattern Anal. Mach. Intell., vol. 27, no. 2, pp. 230– 244, Feb. 2005.
35. Kotsia, I.; Pitas, I.; Zafeiriou, S.; Zafeiriou, S., "Novel Multiclass Classifiers Based on the Minimization of the Within-Class Variance," *Neural Networks, IEEE Transactions on,* pp14 - 34 Volume: 20, Issue: 1, 2009.
36. Y. Tian, "Evaluation of face resolution for expression analysis," CVPR Workshop on Face Processing in Video, 2004.
37. Y. Tian, T. Kanade, J. Cohn, Handbook of Face Recognition, Springer, 2005 (Chapter 11. Facial Expression Analysis)
38. Liu, C., Wechsler, H., "Gabor Feature Based Classification Using the Enhanced Fisher Linear Discriminant Model for Face Recognition," IEEE Trans. Image Processing, Vol. 11, No. 4, pp. 467–476, 2002.
39. Abdulrahman, Muzammil, Tajuddeen R. Gwadabe, Fahad J. Abdu, and Alaa Eleyan. "Gabor wavelet transform based facial expression recognition using PCA and LBP." In Signal Processing and Communications Applications Conference and Communications Applications 2014, 22nd, pp. 2265-2268. IEEE, 2014.
40. Vitomir Struc and Nikola Pavesic, "The Complete Gabor–Fisher Classifier for Robust Face Recognition," EURASIP Journals on Advances in Signal Processing, 2010.
41. M.Sugiyama,"Dimensionality reduction of multimodal labeled data by local Fisher discriminant analysis," Journal of Machine Learning Research, vol.8,pp.1027- 1061,July
42. C.W.Hsu,C.C. Chang and G.J.Lin. (2009), A practical guide to support vectorclassification[Online],Available:http//www.csie.ntu.edu.tw/cjlin/papers/guide/guide.pdf.
43. Ahmad Poursaberi1, "Gauss–Laguerre wavelet textural feature fusion with geometrical information for facial expression identification," EURASIP Journal on Image and Video Processing 2012.
44. Hasimah Ali, Muthusamy Hariharan, Sazali Yaacob, Abdul Hamid Adom, "Facial emotion recognition using empirical mode decomposition," Expert Systems with Applications 42 (2015) 1261–1277, Elsevier Publications.
45. 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.
46. Ashok Rao, S. Noushath, Subspace methods for face recognition – survey, Elsevier, Computer Science Review by Science Direct pp 1-17 (2010) –ELSIVER Series.
47. Guoxia Sun, Liangliang Zhang,"Face Recognition Based on Symmetrical Weighted PCA," IEEE conference, pp-2249-2252, 2011.
48. Yan Wang, Weijun Gong, "Facial Expression Recognition Based on Local Sensitive Feature Extraction with Gabor Wavelet and LE+LDA" Journal of Computational Information Systems 7: 8 (2011) 2745-2752 Available at http://www.Jofcis.com.
49. G. Baudat and F. Anouar, "Generalized discriminant analysis using a kernel approach," Neural Computation, vol. 12, no. 10, pp. 2385–2404, Oct. 2000.
50. Ziqiang Wang, Xia Sun, Lijun Sun, and Yuchun Huang, "Semisupervised Kernel Marginal Fisher Analysis for Face Recognition," The Scientific World Journal, Volume 2013.
51. Ziqiang Wang, Xia Sun,"Manifold Adaptive Kernel Local Discriminant Analysis for face Recognition," Journal of Multimedia Vol 7. No. 6 Dec 2012.
52. Wang Z., Ruan Q. 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.
53. 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",International Journal of Information Technology Vol.11,No.11.2005.
54. F. Y. Shih, C.-F. Chuang, and P. S. P. Wang, "Performance comparisons of facial expression recognition in JAFFE database," International Journal of Pattern Recognition and Artificial Intelligence, vol. 22, no. 3, pp. 445–459, 2008.
55. Shi Dongcheng , Cai Fang ,"Facial Expression Recognition Based on Gabor Wavelet Phase Features, "Image 2013, pp 520 – 523.
56. 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.
57. R Zhi, Q Ruan, Facial expression recognition based on two dimensional discriminant locality preserving projections. Neuro Comput. 71, 1730–1734 (2008)
58. 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
59. W Liejun, Q Xizhong, Z Taiyi, Facial expression recognition using improved support vector machine by modifying kernels. Inf. Technol. J. 8(4), 595–599 (2009).
60. 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.
61. Chien-Cheng Lee, Shin-Sheng Huang and Cheng-Yuan Shih, "Facial Affect Recognition Using Regularized Discriminant Analysis-Based Algorithms, *EURASIP Journal on Advances in Signal Processing* 2010.
62. Sara Zhalehpour, Zahid Akhtar, "Multimodal emotion recognition based on peak frame selection from video," Signal Image and Video Processing 10(5) · September 2015.
63. U Mlakar, B Potočnik -Automated facial expression recognition based on histograms of oriented gradient feature vector differences, Signal, Image and Video Processing, 2015.
64. Amin Mohammadian, Hassan Aghaeinia, Farzad Towhidkhah,"Incor porating prior knowledge from the new person into recognition of facial expression," Signal, Image and Video Processing, 2014 .