Regression Model With Modified Linear Discriminant Analysis Features For Bimodal Emotion Recognition

Gaikwad Kiran Pandhari, Manna Sheela Rani Chetty

Abstract: Now days recognizing the face accurately is becoming more challenging and essential task in the biometric authentication. Use of minimum facial features is important to reduce the complexity of designing the face recognition system. The performance of any emotion recognition system is mostly dependent on efficient design of face recognition system. Recently in the direction of emotion recognition a lot of the work is carried out. It is suggested by some researchers that use of only facial features or speech features are not sufficient to design emotion recognition system. Here in this paper the approach to extract the facial and speech features to recognize the emotion is proposed. Survey suggests that, combining both the features (facial features and speech features) to recognize the emotion improves the emotion state recognition accuracy. Proposed method uses extraction of facial features in both the directions (row and column) using maximum margin criteria with Modified Linear Discriminant Analysis (MLDA). The respective speech data signals are extracted using Mel-Frequency Cepstral Coefficients (MFCC) of the speech. The resultant features (facial features and speech features) are further integrated to construct an informative feature library. The constructed feature library provides a base for recognizing the emotions using advanced regression model, called as Incomplete Sparse Least Square Regression (ISLSR). After experimentation, the proposed approach is found to provide improved recognition accuracy of emotions than existing approaches.

Keywords: Bimodal emotion recognition, MLDA, MFCC, Incomplete Sparse Least Square Regression, Feature Library

—————————— ◆ ——————————

I. INTRODUCTION

Facial feature or speech feature extraction is the key step in emotion recognition. The two techniques Principle component analysis (PCA) and linear discriminant analysis (LDA) describes feature extraction and data representation techniques of pattern recognition. PCA design a feature space such that the feature samples have the minimum reconstruction error. LDA design a feature space where the samples within the class are close together, and the samples between the classes are sufficiently away from each other. PCA and LDA transforms 2-D matrices into 1-D. As a result, the feature vectors lie in a high-dimensional space. So, it requires the more data samples This is known as small sample size (SSS) problem [15]. LDA has the SSS problem in which the within-class covariance matrix becomes singular, and thus, the traditional LDA algorithm fails. As a refinement to this problem, number of approaches suggested, including pseudo-inverse LDA [17], direct LDA (DLDA) [19], regularized discriminant analysis [20], inverse Fisher [21], maximum margin criterion (MMC) [22], weighted piecewise LDA [23]. PCA plus LDA approach (PCA + LDA. PCA+LDA is method uses feature extracting methods. It does not solve the SSS problem. The 2DDA [24] introduced as a computation of between class, within-class, and total scatter matrices. Due extraction of discriminative features in 2DDA, its complexity is less. 2DDA uses the data structures for data analysis while 1DDA ignores the underlying structure. As 2DDA derives within-class scatter matrix non-singular [19], small sample size problem is avoided in 2DDA. As 2D-PCA 25] and 2D-LDA extracts features in one direction, feature extraction time required is less than traditional PCA and LDA. 2D-PCA and 2D-LDA use more discriminative features and it takes longer to test than 1-D environments. 2DPCA [25] and 2DLDA [26] are working in the row direction of images. Combining row and column directions, two-directional 2-D PCA ((2D)2PCA) [27] and two-directional 2-D LDA((2D)2LDA) [28] methods become more efficient representation. Extracting efficient discriminant features in less amount of time is a challenging task. 2DDA extracts the features in both directions of images. So, it is quite possible that the obtained features have redundant information. These redundant features can be reduced by combining 2DDA and 1DDA. Speech emotion recognition can be categorized into two parts as Speech Feature Extraction and Emotion Classification. In the Speech Feature Extraction, extraction of the speech features that are related with the emotions is carried out. And emotion classification stage is to determine the emotion based on the obtained speech features. During the last decade, many speech emotion recognition methods had been proposed [31], among which the regression-based approaches had been very popular in recent years. In speech emotion recognition, trained model used to decide emotion category. Using this technique, following problem occurs:

i) Due to large dissimilarities between the speech data under training and the testing phase, it predicts out of speech sample data.
ii) Another problem is that, different speech emotion features contribute differently to the emotion recognition. Some speech features contribute more speech emotion information whereas some speech feature contribute less speech emotion information. Hence, it would not be a good choice to use all speech features since some contain less useful information.

- *Mr. Gaikwad Kiran Pandhari Department of Computer Science and Engineering K L E F, Vaddeswaram, Dist.Guntur Andhra Pradesh, India kiran30march@gmail.com*
- *Dr. Manna Sheela Rani Chetty Department of Computer Science and Engineering K L E F, Vaddeswaram, Dist.Guntur Andhra Pradesh, India sheelarani_cse@kluniversity.in*

Author [Citation]	**Adopted Methodology**	**Features**	**Challenges**
Davood Gharavian *et al.* [1]	PSO algorithm	❖ Good convergence rate ❖ Less computational complexity	❖ Sticking with local optimal under multimodal scenarios
M. Shamim Hossain and Ghulam Muhammad [2]	Ridgelet Transform	❖ More effective in handling direction ❖ It overcomes the shortcoming of the wavelet	❖ Not viable for majority of the applications
Soroosh Mariooryad and Carlos Busso [3]	Bilinear model	❖ Multiple instances can be concurrently used	❖ Designs are suboptimal
Felix Shaw and Barry-John [4]	Independent component analysis	❖ A model is required for all the expression instead of all expression pair	❖ The output animation is still imperfect to the styles for which a model is accessible.
Jingjie Yan *et al.* [5]	Regression model	❖ It implements a statistical model and shows optimal results.	❖ Deployed to design non-linear relationships ❖ It restrict from predicting the numeric results.

The increase of speech features increases the computational complexity. And hence the higher dimensionality of the speech feature vector may cause the over-fitting problem [32]

Here, II describes literature review, III proposed bimodal emotion recognition system, IV Experimental discussion and V conclusion

II. LITERATURE REVIEW

A. Related Works

In 2017, Davood Gharavian *et al.* [1] have worked on emotion recognition using audio data and visual data. Here, the audio and visual systems have combined at feature and decision levels using stacked generalization technique. For describing the improved values in the parameter of choice, the learning rate of Fuzzy ARTMAP Neural Network (FAMNN) and Particle Swarm Optimization (PSO) used. The results obtained from experiments showed that the computations of unimodal systems were improved using the decision level fusions, feature level, and PSO optimization-based FAMNN. In 2016, M. Shamim Hossain and Ghulam Muhammad [2] have proposed an visual - audio emotion recognition system using Multi-Directional Regression (MDR) properties. The MDR was employed for the ridgelet and audio transform coefficients for recognizing the images of the face. MDR captures the four-directional information. Generally, the ridgelet transform was a helpful equipment to discover the ridges in the points of interest. Moreover, two Extreme Learning Machine (ELM) classifiers are utilized uniquely for the purpose of classification. Here, these two classifiers were mainly used for face modality and the speech modality. To make the final decision by means of the Bayesian sum rule, the scores of the two ELMs were combined. Besides, the system was calculated in the eNTERFACE database by means of both the bimodal and single modal. In 2016, Soroosh Mariooryad and Carlos Busso [3] had worked on a new technique based on the asymmetric bilinear design for lexical compensation. In machine learning applications, to separate style from the content, the factor design for separation was established initially. Moreover, the proposed approach does not need phoneme labels, which was an important enhancement over the supervised lexical compensation. Here, two supervised lexical compensation techniques like model-level and feature-level are adopted, which presume that labels of viseme and their alignment of time information are existing for testing and training the models. Finally, the simulations on SEMAINE database by image-based features reveals the efficiency of the implemented method. In 2016, Felix Shaw and Barry-John [4] have worked on expressive visual speech animation that differs from time and frame by frame. Here, the Active Appearance Models (AAM) has provided a description for statistical shape and appearance changes in a regular image set. Moreover, the Independent Component Analysis (ICA) offers a architecture

for calculating this matrix for computing the independent components. Here, the Principal Component Analysis (PCA) was adopted to project the appearance and shape into lower dimensional spaces which involve the simpler computation.
In 2016, J. Yan *et al.* [5] have introduced a bimodal emotion recognition technique based on Sparse Kernel Reduced-Rank Regression (SKRRR) fusion technique. In this paper, the Scale Invariant Feature Transform (SIFT) feature descriptor has obtained efficient properties from facial expression modality and speech modality. Therefore, the SKRRR fusion technique was adopted to combine the features of emotion of both the modalities. In addition, the proposed technique was a nonlinear extension of the conventional Reduced-Rank Regression (RRR). In RRR, the response feature vectors and predictors were kernelized by mapping them to two high-dimensional feature spaces using two nonlinear mappings. Finally, the experiments of bimodal and monomodal emotion recognitions are simulated, and the results have indicated that the proposed technique achieves the largest rate of bimodal emotion recognition .

B. Review

Few research contributions have been made to address the problem in bimodal emotion recognition. The conventional emotion recognition system is shown in Table I. PSO used in [1] has experimentally shown its fast convergence behavior. Nevertheless, it sticks with local optimal under multimodal scenarios. Moreover, ridgelet transform [2] is more effectual, and it overcomes the drawbacks of the wavelet. Conversely, the main issue of this method is its ineffectiveness for certain applications. Besides, the bilinear model [3] can be concurrently utilized for multiple instances. Designs of the bilinear model are sub-optimal. The Independent component analysis in [4] has been experimentally demonstrated that the model is essential for each expression. The Regression model used in [5] has shown optimal results, but it is limited to predicting the numeric output. Hence, it is motivated to deal with the bimodal emotion recognition from both the facial expression and speech.

III. BIMODAL EMOTION RECOGNITION APPRAOCH

From above literature review it is observed that, designing the bimodal emotion system for emotion recognition improves the recognition accuracy. With this conclusion from literature review, here bimodal emotion recognition approach is proposed. With the achievement of Two-Dimensional Maximum Margin criteria and Regression Technique, here combination of these two techniques is proposed to design a bimodal emotion recognition.

et D_v be the database and V_i be the video where $i = 1,2 \cdots N_v$ and N_v is the number of videos in the database, The video V_i includes both image and video sequences represented by Eq. (1), where, N_I is the total number of image frames and N_A is the number of video sequences.

$$V_i = \begin{bmatrix} I_1 & I_2 & \cdots & I_{N_I} \\ A_1 & A_2 & \cdots & A_{N_A} \end{bmatrix} \quad (1)$$

C. Facial Feature Extraction

The facial features are obtained modified maximum margin criteria [33] method. A face image sample set $B = \{B_1, B_2, \cdots B_T\}$, where every individual element addresses a face image represented with a $m \times n$ dimension matrix and T denotes the total number samples used for training. Here $\omega_1, \omega_2 \cdots \omega_c$ are used to define pattern classes and $n_i (i = 1, \cdots c)$ is the total number of the face image samples of the i^{th} class

The N describes the entire mean matrix of B and $N_i (i = 1, \cdots c)$ indicates the mean matrix of the i^{th} class. The 2-D in-class scatter matrix P_w, 2-D among class scatter

matrix P_b, 2-D entire scatter matrix P_t are given by Eq.(2), Eq. (3), Eq. (4) respectively.

$$P_b = \frac{1}{T} \sum_{i=1}^{c} n_i (N_i - N)^t (N_i - N) \quad (2)$$

$$P_w = \frac{1}{T} \sum_{i=1}^{c} \sum_{j \in w_i} (B_j - N_i)^t (B_j - N_i) \quad (3)$$

$$P_t = \frac{1}{T} \sum_{i=1}^{T} \left((B_i - N)^t (B_i - N) \right) = P_b + P_w \quad (4)$$

The 2-D in-class scatter matrix is given by A_w, the 2-D relying among class scatter matrix indicated by A_b and the 2-D total scatter matrix referred by A_t which are formulated as and represented by Eq. (5), Eq. (6) and Eq. (7), respectively, where, A_b, A_w and A_t are $m \times m$ matrices.

$$A_b = \frac{1}{T} \sum_{i=1}^{c} n_i (N_i - N)(N_i - N)^t \quad (5)$$

$$A_w = \frac{1}{T} \sum_{i=1}^{c} \sum_{j \in w_i} (B_j - N_i)(B_j - N_i)^t \quad (6)$$

$$A_t = \frac{1}{T} \sum_{i=1}^{T} \left((B_i - N)(B_i - N)^t \right) = A_b + A_w \quad (7)$$

MMC [22] is a 1-D linear discriminant technique that does exhibit the issues caused due to SSS. The aim of MMC is to discover a projection matrix which increases the standard border among the classes as represented by Eq. (8) where, $Vr(z)$ denotes the trace matrix of z, W_w and W_b denotes the in-class scatter matrix and between-class matrix respectively, S refers to weight that is referred for optimization and S^t is the transpose of weight.

$$S_{MMC} = \underset{S \in P^{d \times d'}, S^t S = I_{d'}}{\arg\max} Vr\left(S^t W_b S - S^t W_w S\right) \quad (8)$$

Feature extraction technique depending on MMC principle is efficient, stable and robust. The 2D^{2}MMC method extracts distinguishing face feature vectors from two directions as shown in Eq. (9) and Eq. (10).

$$S_{MMC-P} = \underset{S \in P^{n \times n_r'}, S^t S = I_{n_r}}{\arg\max} Vr\left(S^t P_b S - S^t P_w S\right) = w_1^r \; w_2^r \cdots w_{n_r}^r \quad (9)$$

$$S_{MMC-A} = \underset{S \in P^{m \times n_c'}, S^t S = I_{n_c}}{\arg\max} Vr\left(S^t A_b S - S^t A_w S\right) = w_1^c \; w_2^c \cdots w_{n_c}^c \quad (10)$$

Thus, 2D^{2}MMC should identify two best projectors from both directions, namely, the column projector S_{MMC-A} and the row projector S_{MMC-P}.

The subsequent steps are the procedure for 2D^{2}MMC extraction.

- Evaluate P_b and P_w, which are $n \times n$ matrices as represented in Eq. (4) and A_b and A_w which are $m \times m$ matrices as defined in Eq. (7).
- Using 2D^{2}MMC, calculate two projectors from both the directions.
- Measure the row projector S_{MMC-P} by finding optimization in Eq. (9) and Eq. (11)

s.t. $(w_i^r)^t (P_b - P_w) w_i^r > \max\left(Vr(P_b - P_w)/n, 0\right)$

(11)

Calculate the column projector S_{MMC-A} in Eq. (10) and Eq. (12).

s.t. $(w_i^c)^t (A_b - A_w) w_i^c > \max\left(Vr(A_b - A_w)/m, 0\right)$ (12)

Further, by employing the projectors S_{MMC-P} and S_{MMC-A}, the dimensionality reduction of facial feature vectors can be done using Eq. (13).

$$S_{MMC-P} \oplus S_{MMC-A} : X \in P^{m \times n} \to S_{MMC-A}^t X S_{MMC-P} \in P^{n_c \times n_r} \quad (13)$$

- Convert 2-D matrix into 1-D vector, and then apply Linear Discriminant Analysis (LDA) to take out the features further. Also evaluate the $\tilde{W}_b$ and $\tilde{W}_S$ in the $n_r n_c$ dimensional space.
  - Calculate $S_{LDA} = [\phi_1, \phi_2, \cdots \phi_{nl}]$ which comprises the general eigen vectors of $\tilde{W}_b$ and $\tilde{W}_S$ related to the non zero eigen values, and then employ the projector S_{LDA} as represented in Eq. (14) to minimize the dimensions further

$$S_{LDA} : x \in S^{n_c n_r} \to S_{LDA}^t x \in S^{n_l} \quad (14)$$

- Recognition: For a given face image test sample set B, utilize the three projectors S_{MMC-P}, S_{MMC-A} and S_{LDA} to reduce the dimension, and then use the nearest neighbor classifier to classify the face image test sample set.

After obtaining facial feature vectors by 2D^{2}MMC, the weight S in Eq. (8) should be optimized by Non-Linear Image

Programming Model [34]. And hence the unique features, which clearly distinguish the emotions, can be obtained. The obtained face feature vectors are represented with F_1 where $F_1 = \{a_1, a_2 \cdots a_{N_D}\}$. Here N_D indicates the number of features obtained using 2D²MMC.

D. Speech Feature Data Extraction

Cepstral analysis (MFCC) [36] is adopted for obtaining the audio feature. The speech data signal $s(n)$, is obtained from the two signals $G(n)$ and $I(n)$ represented in Eq. (15). Composite cepstrum $\hat{s}(n)$ is represented in Eq. (22), where, $G(n)$ and $I(n)$ represents the actual speech portion data and No speech data portion respectively.

$$s(n) = G(n) + I(n) \rightarrow \hat{s}(n) = \hat{G}(n) + \hat{I}(n) \tag{15}$$

The cepstral coefficients can be found by Eq. (16). The log of the speech data signal $L(z)$ is represented in Eq. (17)

$$s(n) = s_1(n) * s_2(n) \Leftrightarrow L(z) = L_1(z)L_2(z) \tag{16}$$

$$\log\{L(z)\} = \log\{L_1(z)\} + \log\{L_2(z)\} = \hat{L}(z) \tag{17}$$

The convolved signals $\hat{s}_1(n)$ and $\hat{s}_2(n)$ could be summed if the complex-log is discriminative. The z transform can be as follows:

$$\hat{s}(n) = \hat{s}_1(n) + \hat{s}_2(n) \tag{18}$$

It is observed that, having poles and zeros into the unit circle of the signal $s(n)$ is not feasible. So, to avid Eq. (19) is used.

$$\log\{L(v)\} = \log\{|L(v)|\} + j \sphericalangle L(v) \tag{19}$$

If $L(v) = L_1(v)L_2(v)$ then $\log\{|L(v)\}$ can be computed using following equation:

$$\log\{|L(v)|\} = \log\{|L_1(v)L_2(v)\} = \log\{|L_1(v)|\} + \log\{|L_2(v)|\} \tag{20}$$

The real cepstrum, $D_x(n)$ is represented as follows:

$$D_x(n) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \log|L(e^{lv})|\, e^{lvn} dv \tag{21}$$

The magnitude of $D_x(n)$ is real non-negative value.

The risky cepstrum $\hat{s}(n)$ is computed as follows:

$$\hat{s}(n) = \frac{1}{2\pi}\left[\left[\log|L(e^{lv})|\right] + l \arg\left(L(e^{lv})\right)\right] e^{lvn} dv \tag{22}$$

The phase is denoted as $\arg(\ )$, $\log\{|L(v)\}$ and $\log|L(e^{lv})$ are the log spectrum of the signal. It is complex logarithm and risky cepstrum of the actual sequence is real.

The real cepstrum is the even part of $\hat{s}(n)$. $D_x(n)$, is computed using an Inverse Fourier Transform (IFT) applying on the log spectrum of the speech signal as follows:

$$D_x(n) = \frac{\hat{s}(n) + \hat{s}(-n)}{2} \tag{23}$$

Instead of using Fourier transform on speech signal data, Discrete Fourier transform (DFT) is applied on speech signal data as follows:

$$\begin{aligned} L_t(h) &= \sum_{n=0}^{H-1} s(n) e^{-l\frac{2\pi}{N}hn} \quad 0 \le h \le H-1 \\ \hat{L}_t(h) &= \log\{L_t(h)\} \quad 0 \le h \le H-1 \\ \hat{s}_t(n) &= \frac{1}{H}\sum_{h=0}^{N-1} \hat{L}_t(h) e^{l\frac{2\pi}{H}hn} \quad 0 \le n \le H-1 \end{aligned} \tag{24}$$

$\hat{L}_t(h)$ is denoted as the sampled version of $\hat{X}(e^{jw})$ as follows:

$$\hat{s}_t(n) = \sum_{p=-\infty}^{\infty} \hat{s}(n + pN) \tag{25}$$

The cepstral feature $D_t(n)$ of $s(n)$ can be computed as follows:

$$D_t(n) = \frac{1}{H}\sum_{h=0}^{H-1} \log|L_t(h)|\, e^{l\frac{2\pi}{H}hn} \quad 0 \le n \le H-1 \tag{26}$$

where H is the period

Thus, the speech feature data obtained using cepstral method (MFCC) is represented by F_2 where $F_2 = \{b_1, b_2 \cdots b_{N_C}\}$, here N_c represents the total number of speech data features obtained from cepstral analysis.

IV. RESULTS

The proposed bimodal emotion recognition approach implemented using matlab and simulation results observed. For the experimental purpose, video data of a person with various emotions (angry, sad, surprise, fear, happy and disgust) is captured with some convarsation. Video data is caputured under certain restricted conditions. In noise free and same lighting conditions the videos are captured. The person is asked to speak certain dialogue and video is captured. The videos of 50 persons for 10 seconds with different facial expressions are captured. Assuming that 50 persons are not wearing glasses and no major facial changes videos are captured. Using the audio extractor and frame sequence extractor, the speech data and facial image data obtainned. For every person and for each of the emotion speech data feature vectors and face image feature vectors are created. The features in image sequences in video were extracted based on the 2D²MMC technique and the audio sequences were extracted using the cepstral analysis. Every emotion category contains the possible features obtained with 2D²MMC and cepstral analysis. In this way training database to recognize the emotion category is designed. Then one of the untrained video data of a person is used for testing. Applying 2D²MMC and cepstral analysis facial feature vector and speech data feature vector obtained of untrained data. And using ISLSR classification technique emotions are recognized. Proposed combination of 2D²MMC- ISLSR (DI) for recognizing emotions is compared with four existing combinations such as LDA- LDCA (LL), Row-MDC- LDCA (RML), Column-MDC-LDCA (CML), DDMMC-LDCA (DL).

For experimental purpose, the videos of 10 seconds of 50 persons are created. From Table II, the emotion recognition accuracy of the proposed bimodal emotion recognition system is 5% better than LL, 2.5% better than RML, 5% better than CML, 10.4% better than DL

TABLE II Facial Feature Analysis of Proposed and Conventional Technique

Approach	Accuracy
LL	0.722222
RML	0.740741
CML	0.722222
DL	0.685185
DI	**0.759259**

Similarly, the cepstral analysis for the proposed method is analysed with Cepstral-LDCA (CL) and the results are obtained. From Table III, the proposed of CI is 2.5% better in terms of accuracy than the conventional CLcombination

TABLE-III Speech Feature Analysis of Proposed and conventional Technique

Approach	CL	CI
Accuracy	0.7037	0.722222

V. CONCLUSION

The proposed emotion recognition method obtains facial data features and speech data features. Facial data features were obtained using the proposed $2D^2$MMC features, which makes use of MLDA. The key feature of this method is to avoid small sample size problem. Hence, complexity of designing emotion recognition system becomes easier due to reduced features. Further, cepstral analysis (MFCC) is used for extracting features from speech data. Also, with the reduction of facial data features, speech data features are integrated to improve overall recognition rate. As here, regression technique is used to classify facial features and speech data features, the emotion recognition accuracy is improved. Experimental results show that emotion features classified with ISLSR classification technique is 5% better than LL, 2.5% better than RML, 5% better than CML, 10.4% better than DL. The cepstral analysis with ISLSR classification technique (CI) is analysed with Cepstral-LDCA (CL) and the results are obtained. The proposed CI method is 2.5% better in terms of accuracy than the conventional CL combination

REFERENCES

[1]. Davood Gharavian, Mehdi Bejani Mansour Sheikhan,"Audio-visual emotion recognition using FCBF feature selection method and particle swarm optimization for fuzzy ARTMAP neural networks",Multimedia Tools and Applications, vol. 76, no.2, pp. 2331-2352,January 2017.

[2]. M. Shamim Hossain,Ghulam Muhammad,"Audio-visual emotion recognition using multi-directional regression and Ridgelet transform,"Journal on Multimodal User Interfaces, vol.10, no.4, pp.325–333,December 2016.

[3]. S. Mariooryad and C. Busso, "Facial Expression Recognition in the Presence of Speech Using Blind Lexical Compensation,"IEEE Transactions on Affective Computing, vol. 7, no. 4, pp. 346-359, Oct.-Dec. 1 2016.

[4]. F. Shaw and B. J. Theobald, "Expressive Modulation of Neutral Visual Speech," IEEE Multimedia, vol. 23, no. 4, pp. 68-78, Oct.-Dec. 2016.

[5]. J. Yan, W. Zheng, Q. Xu, G. Lu, H. Li and B. Wang, "Sparse Kernel Reduced-Rank Regression for Bimodal Emotion Recognition From Facial Expression and Speech," IEEE Transactions on Multimedia, vol. 18, no. 7, pp. 1319-1329, July 2016

[6]. Wenming Zheng, Minghai Xin, Xiaolan Wang, and Bei Wang, "A Novel Speech Emotion Recognition Method via Incomplete Sparse Least Square Regression," IEEE SIGNAL PROCESSING LETTERS, VOL. 21, NO. 5, MAY 2014.

[7]. G. Liu, Z. Lin, S. Yan, J. Sun, Y. Yu, and Y. Ma, "Robust recovery of subspace structures by low-rank representation," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 35, no. 1, pp. 172–184, 2013

[8]. Z. Qin and d. Goldfarb, "Structured sparsity via alternating direction methods," *J. Mach. Learn. Res.*, vol. 13, pp. 1333–1367, 2012

[9]. J. Yang, L. Zhang, Y. Xu, and J.-Y. Yang, "Beyond sparsity: The role of -optimizer in pattern classification," *Patt. Recognit.*, vol. 45, pp. 1104–1118, 2012

[10]. W.H. Yang, D. Q. Dai, Two-Dimensional Maximum Margin Feature Extraction for Face Recognition, IEEE Trans.Sys.Man.Cybern.vol-39, no-4, pp. 1002-1012, Aug-2009

[11]. Kiran P Gaikwad, V.M. Wadhai, Prasad S Halgaonkar, Santosh Kumar, "Design and Implementation of Robust 2D Face Recognition System for Illumination Variations," International Journal of Computer Applications, Vol-10, No-3, November 2010.

[12]. I T. Jolliffe, "Principal Component Analysis," New York: Springer-Verlag, 1986.

[13]. M. A. Turk and A. P. Pentland, "Eigenfaces for recognition," J. Cogn. Neurosci., vol. 3, no. 1, pp. 71–86, 1991.

[14]. P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection," IEEE Trans. Pattern Anal. Mach. Intell., vol. 19, no. 7, pp. 711 720, Jul. 1997

[15]. S. J. Raudys and A. K. Jain, "Small sample size effects in statistical pattern recognition: Recommendations for practitioners," IEEE Trans.Pattern Anal. Mach. Intell., vol. 13, no. 3, pp. 252–264, Mar. 1991.

[16]. W. J. Krzanowski, P. Jonathan, W. V. McCarthy,and M. R. Thomas, "Discriminant analysis with singular covariance matrices: Methods and applications to spectroscopic data," Appl. Stat., vol. 44, no. 1, pp. 101–115, 1995.

[17]. S. Raudys and R. P. W. Duin, "Expected classification error of the Fisher linear classifier with pseudo-inverse covariance matrix," Pattern Recognit. Lett., vol. 19, no. 5/6, pp. 385–392, Apr. 1998.

[18]. L. Chen, H. Liao, M. Ko, J. Lin, and G. Yu, "A new LDA-based face recognition system which can solve the small sample size problem," Pattern Recognit., vol. 33, no. 10, pp. 1713–1726, Oct. 2000.

[19]. H. Yu and J. Yang, "A direct LDA algorithm for high-dimensional data - With application to face recognition," Pattern Recognit., vol. 34, no. 10, pp. 2067–2070, Oct. 2001.

[20]. D. Q. Dai and P. C. Yuen, "Regularized discriminant analysis and its application to face recognition," Pattern Recognit., vol. 36, no. 3, pp. 845–847, Mar. 2003.

[21]. X. S. Zhuang and D. Q. Dai, "Inverse Fisher discriminate criteria for small sample size problem and its application to face recognition," Pattern Recognit., vol. 38, no. 11, pp. 2192–2194, Nov. 2005.

[22]. H. Li, T. Jiang, and K. Zhang, “Efficient and robust feature extraction by maximum margin criterion,” IEEE Trans. Neural Netw., vol. 17, no. 1, pp. 157– 165, Jan. 2006

[23]. M. Kyperountas, A. Tefas, and I. Pitas, “Weighted piecewise LDA for solving the small sample size problem in face verification,” IEEE Trans. Neural Netw., vol. 18, no. 2, pp. 506–519, Mar. 2007.

[24]. P. Howland, J. Wang, and H. Park, “Solving the small sample size problem in face recognition using generalized discriminant analysis,” Pattern Recognit., vol. 39, no. 2, pp. 277–287, Feb. 2006.

[25]. J. Yang, D. Zhang, A. F. Frangi, and J. Y. Yang, “Two-dimensional PCA: A new approach to appearance-based face representation and recognition,” IEEE Trans. Pattern Anal.Mach. Intell., vol. 26, no. 1, pp. 131–137, Jan. 2004.

[26]. M. Li and B. Yuan, “2D-LDA: A statistical linear discriminant analysis for image matrix,” Pattern Recognit. Lett., vol. 26, no. 5, pp. 527–532, Apr. 2005.

[27]. D. Zhang and Z. H. Zhou, “(2D)2PCA: Two directional two-dimensional PCA for efficient face representation and recognition,” Neuro-computing, vol. 69, no. 1–3, pp. 224–231, Dec. 2005.

[28]. S. Noushath, G. Hemantha Kumar, and P.Shivakumara, “(2D)2LDA: An efficient approach for face recognition,” Pattern Recognit., vol. 39, no. 7, pp. 1396–1400, Jul. 2006.

[29]. X. S. Zhuang and D. Q. Dai, “Improved discriminate analysis for high-dimensional data and its application to face recognition,” Pattern Recognit., vol. 40, no. 5, pp. 1570–1578, May 2007.

[30]. D. Q. Dai and P. C. Yuen, “Face recognition by regularized discriminant analysis,” IEEE Trans. Syst., Man, Cybern. B, Cybern., vol. 37, no. 4, pp. 1080–1085, Aug. 2007.

[31]. M. E. Ayadi, M. S. Kamel, and F. Karray, “Survey on speech emotion recognition: Features, classification schemes, and databases,” *Patt. Recognit.*, vol. 44, pp. 572–587, 2011

[32]. C. M. Bishop*, Pattern Recognition, and Machine Learning*. Berlin, Germany: Springer, 2006

[33]. Gaikwad Kiran P., C. M. Sheela Rani , Mahajan S. B., Sanjeevikumar P. "Dimensionality Reduction of Facial Features to recognize Emotion State,” Ist Springer International Conference (ETAEERE-2016), Sikkim

[34]. Maged M.M.Fahmy, "Palmprint recognition based on Mel frequency Cepstral coefficients feature extraction", Ain Shams Engineering Journal, vol.1, no.1, pp.39-47, September 2010.