

Performance Analysis of Robust 2D-PCA and 2D-LDA Feature Descriptors on SVM

Victor Mokaya Mageto¹ | Mukesh Kumar Gupta²

¹Research scholar, Department of computer engineering and technology, Suresh Gyan Vihar University, Jaipur, Rajasthan, India.

²Associate Professor, Department of electrical engineering, Suresh Gyan Vihar University, Jaipur, Rajasthan, India.

To Cite this Article

Victor Mokaya Mageto and Mukesh Kumar Gupta, "Performance Analysis of Robust 2D-PCA and 2D-LDA Feature Descriptors on SVM", *International Journal for Modern Trends in Science and Technology*, 6(8S): 112-116, 2020.

Article Info

Received on 16-July-2020, Revised on 15-August-2020, Accepted on 25-August-2020, Published on 28-August-2020.

ABSTRACT

Facial expressions are the most intuitive way of communicating non-verbal messages. This type of communication provides effective response and feedback from the speaker to listener and vice-versa. In this paper robust macro facial expression recognition techniques are presented. 2D-PCA and 2D-LDA are robust geometric feature descriptors presented in this paper capable of cancelling noise and extracting maximum spatial features from image samples with unstable illumination condition which leads to correct classification of results. Experiments are carried out separately on both feature descriptors using Cohn Kanade (CK+) dataset, MMI dataset, and Japanese Female Facial Expressions (JAFFE) database for analysis. The experimental results are evaluated and their performance compared using Support vector machine (SVM) classifier with three kernels; linear, polynomial and radial basis function (RBF). The results from the proposed methods are also compared with existing novel methods and it is found out that the results from the proposed methods perform significantly well.

KEYWORDS: 2D-LDA, 2D-PCA, Kernel function, SVM, CK+, MMI, JAFFE, macro-facial expression

I. INTRODUCTION

In the recent times research in facial expression recognition has significantly become a dominant research topic in the field of digital image processing. A lot of industrial applications have created a scope for the development of this technology. Among the current application areas include; biometric identification systems, diagnosis of medical conditions, customer satisfaction feedback for digital marketing platforms, entertainment, fashion and Human computer interactions (HCI). The demand for efficient robust systems from the market is the reason behind the consistent research activities in this field. There are currently three techniques for performing facial

expression recognition which includes; geometric methods, appearance methods, and a currently a hybrid method of combining geometric and appearance together has become commonly in use. In the proposed method the geometric method was adopted. This method works by encoding the complex facial landmarks into regions of interest. According to early works by Paul Ekman, a standard coding system is developed for standard labelling of all facial components. This coding system is known as Facial Actions Coding System (FACS). In literature novel methods have been developed in respect to the proposed method. The best novel techniques available in literature include works by[1], in which a computational

efficient probabilistic neural networks (PNNs) with improved kernel linear discriminant analysis (IKLDA) is proposed. This technique was robust to dimensionality reduction and retaining salient information. According to [2][3] strong illumination and lighting conditions are sensitive for effective feature description hence using robust feature representation compensate for the lost spatial features. According to work proposed by [4] application of Min-Max and Z-score normalization schemes along with fusion methods results in improved recognition rates. According to [5] PCA-SIFT is proposed to extract salient features facial components with different posture variations, the results obtained are significant in similarity finding. In works proposed by [6]–[8] a linear dimensionality reduction technique is proposed to integrate neighbor-hood-free discriminative features with adaptive features for robust similarity finding. According to [9] a patch-based neighbor-hood-preserving discriminant analysis algorithm called ONPDA. From perspective of path alignment intra-class relationships effective feature representation is obtained.

II. PROPOSED METHOD

The performance evaluation of the 2D-PCA and 2D-LDA feature descriptors proposed in this paper are implemented according to the following steps. (I) Image pre-processing, (II) feature, and (III) feature classification.

A. Image Pre-processing

This is the initial step in this process, input image frames are exported to the system for initial image pre-processing. This process is suitable for removal of noise, scaling and illumination correction. In this paper the images were scaled to equal pixel size of 256×256 . Colour conversion was performed to convert them from coloured to gray-scale image. The images were labelled according to the emotion they possess since it's a supervised classification experiment. The input images were fed to the next processing channel for feature extraction

B. Feature Extraction

The proposed feature extractors are explained in detail in this subsection

➤ 2D-PCA

Principal Component Analysis(PCA) [11] is widely used as a dimensionality reduction feature

representation. In this paper, 2D-PCA is proposed and its implementation is explained hereafter. Denoised pre-processed data is imported for locating of salient features and regions of interest. PCA is used identify eigen vectors connected to feature components with highest frequency values. Image frames with eigen vector possess covariance matrix which are interpreted as high dimensional vector space for utilization of feature extraction. Eigen faces from each image are used for classification and compression of weights. The 2D-PCA is implemented as follows. Suppose P training samples possess R class (R_1, R_2, R_3, R_R) and the samples are denoted as $Q_i = (i = 1, 2, 3 \dots, L)$. Therefore, the sample size will be stated as D_j ($j = 1, 2, 3 \dots, P$) is $p \times q$. Since the objective is to find the best projection vector such that when D_j is projected to X then the projection feature vector B_j of image D_j is obtained using equation (1).

$$B_j = D_j X \quad j = 1, 2, 3 \dots, P. \quad (1)$$

To identify the in-between scatter matrix SM_T and in-within scatter matrix SM_A of the projected vectors from the training samples, a linear manifold is projected to the sample images in order to maximize in-between scatter matrix while still minimizing in-within scatter. Lastly, the covariance matrix of the projected feature vector is derived using linear projection as shown in (2) and (3) below.

$$J(X) = \frac{\text{tr}(SM_T)}{\text{tr}(SM_A)} = \frac{X^S M_T X}{X^S M_A X} \quad (2)$$

Where

$$M_T = \sum_{i=1}^R Q_i (\bar{D}_i - \bar{D})^S (\bar{D}_i - \bar{D}) \quad \text{and} \quad (3)$$

$$M_A = \sum_{i=1}^R \sum_{D_k \in R_i} (D_k - \bar{D}_i)^S \times (D_k - \bar{D}_i)$$

The optimal projection X_{opt} is derived at whenever the linear manifold is maximum. Hence $X_{opt} = \text{argmax} J(x)$. The answer to this problem is achieved by generalizing the eigenvalue equation. Given that b is optimal projection to $x_1, x_2 \dots x_b$ axes equivalent to the largest eigen value then the feature of image I can be obtained as in (4) below.

$$y = I[x_1, x_2 \dots, x_b] \quad (4)$$

Moreover, given that I is matrix $m \times n$ and $[x_1, x_2, \dots, x_b]$ is matrix $y \times b$, therefore, matrix $z \times b$ of y is derived to represent I . The derived 2D-PCA features possess the feature properties desired for similarity finding

➤ 2D-LDA

Linear Discriminant Analysis (LDA) [10] is a classical feature extraction and dimensionality reduction technique. LDA objective is to project the input features into a reduced dimensional vector through a feature projection. The initial step is to derive the intensity of separation among the samples using scatter matrix similar to PCA. Given that the initial data is represented as $X \in R^{l \times n}$ in c classes $\pi = \pi_1, \pi_2, \dots, \pi_c$ then the projection will be $W \in R^{l \times c}$. Therefore, the dimensionality reduction of LDA is given by $y_i = W^T x_i$, where $x_i \in R^{l \times 1}$ is a sample from the initial data. Finally, the in-between class and in-within class scatter matrix S_w in equation (5) does the dimensional reduction for matrix W .

$$S_b = \sum_{i=1}^c n_i (M_i - M)(M_i - M)^T, \quad (5)$$

$$S_w = \sum_{i=1}^c \sum_{x_j \in \pi_i} (X_j - M_i)(X_j - M_i)^T$$

Where n_i the class sample size in class π_i , and M is the whole sample size whose average for the whole dataset is given by (6).

$$M_i = \frac{1}{n_i} \sum_{x_j \in \pi_i} X_j \quad (6)$$

Likewise, the transformation of lower dimensional in-between class and in-within class matrix is given by (7) below.

$$\bar{S}_b = W^T S_b W, \quad (7)$$

$$\bar{S}_w = W^T S_w W$$

The above conditions decompose the optimal projection matrix W as given by (8) below.

$$\frac{\max \|\bar{S}_b\|}{W \|\bar{S}_w\|} \quad (8)$$

According to [11], the above function can also be rewritten as shown below in equation (9).

Where function $T_r(\cdot)$ denotes the matrix trace operation. Having formulated the LDA feature extractor, it can be summarized that data representation is the major difference between LDA

and 2D-LDA. To apply the 2D-LDA, a set of transformation matrices are utilized. Given that X and Y are transformation matrices, Z is given by $Z = [z_1, z_2, \dots, z_n]$ as input, where by $Z_i \in P^{m \times n}$. Therefore, the projected in-between class and in-within class scatter matrices are derived as shown in (9).

$$S_b = \sum_{i=1}^c n_i X^T (M_i M) Y Y^T (M_i - M)^T X, \quad (9)$$

$$S_w = \sum_{i=1}^c \sum_{z_j \in \pi_i} U^T (Z_j - M_i) Y Y^T (Z_j - M_i)^T X$$

Hence, the objective function of LDA is attained through 2D scatter matrix and the output of the objective function used to achieve the optimum vector via a set of transformation matrix X, Y is as shown in (10).

$$\frac{\max}{U, V} T_r((\bar{S}_w)^{-1} \bar{S}_b) \quad (10)$$

C. Feature Classification

The linear Support vector Machine (SVM) [12] is a supervised feature classification classifier for object detection. In this paper it was utilized to classify the universal six facial expressions consisting of anger, fear, sad, happy, fear and surprise. Three kernels used to evaluate the performance of the models include; linear, radial basis function and polynomial.

III. EXPERIMENTS AND RESULTS

The experiments were conducted separately on three benchmark datasets namely: extended Cohn Kanade dataset (CK+) [13], Japanese female facial expression (JAFFE) [14], and MMI [15] datasets. CK+ dataset consists of image frames obtained from subjects of different age, gender, ethnic background and continent. Peak expressions were utilized and 225 image frames were used. JAFFE dataset consists of Japanese female models and the image frames were captured in environments with different lighting conditions. The input sample size collected for these experiments was 213 images. MMI dataset consists of subjects of different gender and age. The dataset is made up of video frames captured in sequential order beginning with neutral expression and rising to peak expressions then falling back to original expression. The input sample size extracted from

these video frames was 225 images. The sample cross-section of MMI, CK+ and JAFFE datasets is shown below in Fig.1 below.



Fig.1 Sample cross section of six expression frames from MMI, CK+ and JAFFE dataset

A. Experiments on 2D-PCA and 2D-LDA

The sample size of all the datasets were divided into two sections; training and testing sample. The training sample consisted of nine-folds while the testing sample one-fold. A cross validation strategy was applied to the training samples to split and shuffle them for best training outcomes. On successful training the algorithm was used to test one-fold of the entire dataset. The results of the performance were evaluated using three kernels to analysis their performance accuracy score. The results of the experiment on 2D-PCA and 2D-LDA are shown below in **Table.1**, **Table.2**, and **Table 3**.

Table.1 Results of six expressions using 2D-PCA on CK+, MMI, JAFFE datasets

Kernel function	Ck+	MMI	JAFFE
RBF	92.30	75.47	54.05
Linear	87.17	66.04	86.48
Polynomial	87.17	71.69	97.29
Average accuracy	88.88%	71.06%	79.27%

Table. 2 Results of six expressions using 2D-LDA on CK+, MMI, JAFFE datasets

Kernel function	Ck+	MMI	JAFFE
RBF	89.74	73.58	81.08
Linear	84.61	69.81	81.63
Polynomial	79.48	75.47	78.37

Average accuracy	78.24%	72.95%	80.36%
-------------------------	---------------	---------------	---------------

Table. 3 Comparison results of 2D-PCA and 2d-LDA on six expressions using ck+, MM, JAFFE

Kernel function	2D-PCA			2D-LDA		
	CK+	MMI	JAFFE	CK+	MMI	JAFFE
RBF	92.3	75.5	54.1	89.7	73.6	81.1
Linear	87.2	66.0	86.5	84.6	69.8	81.6
Polynomial	87.2	71.6	97.3	79.5	75.5	78.4

datasets.

From the above results it is observed that the datasets perform differently according to the technique used. In Table.1 result from CK+ database recorded the highest recognition accuracy as they performed best on 2D-PCA method with accuracy score of 88.88%. Similarly, in Table.2 performance of JAFFE dataset performed slightly fairly on 2D-LDA technique with 80.36% accuracy. Finally, in Table.3 it was observed that the MMI dataset achieved poorly on both techniques with accuracy score of 71.69% on 2D-PCA and 72.95% on 2D-LDA respectively.

IV. DISCUSSION

From the results of the experiments shown above, it is noticed that this technique performs fairly well on the proposed techniques. Moreover, nowadays research in this field has shifted from geometric to machine learning models like neural networks. To account for the performance of this model, it is important to remember that that both techniques are useful for dimensionality reduction hence salient and spatial feature data might have been eroded as a result this process. MMI dataset is extracted from video frames captured from different environments which have significant noise levels. Hence the denoising process done using discrete wavelet transform could not result achieve the required balance needed. Finally, MMI dataset was weakly annotated for the data which resulted in misclassification of class labels. However, using a robust classifier like convolution neural networks [16] through transfer learning is a possible solution to improve the results. In this paper two techniques have been introduced to perform facial expression recognition on benchmark datasets, it is observed that the techniques perform fairly well. The efficiency of the model will be improved in future by using a robust and effective feature classifier. Additionally, a robust feature classifier will be amalgamated with

the proposed technique to achieve excellent performance results.

V. CONCLUSION

In this paper two techniques have been introduced to perform facial expression recognition on benchmark datasets, it is observed that the techniques perform fairly well. The efficiency of the model will be improved in future by using a robust and effective feature classifier. Additionally, a robust feature classifier will be amalgamated with the proposed technique to achieve excellent performance results.

REFERENCES

- [1] A. Ouyang, Y. Liu, S. Pei, X. Peng, M. He, and Q. Wang, "A hybrid improved kernel LDA and PNN algorithm for efficient face recognition," *Neurocomputing*, vol. 393, pp. 214–222, 2020, doi: 10.1016/j.neucom.2019.01.117.
- [2] M. U. Haq, A. Shahzad, Z. Mahmood, A. A. Shah, N. Muhammad, and T. Akram, "Boosting the face recognition performance of ensemble based LDA for pose, non-uniform illuminations, and low-resolution images," *KSIIT Transactions on Internet and Information Systems*, vol. 13, no. 6, pp. 3144–3164, 2019, doi: 10.3837/tiis.2019.06.021.
- [3] Z. B. Lahaw, D. Essaidani, and H. Seddik, "Robust Face Recognition Approaches Using PCA, ICA, LDA Based on DWT, and SVM Algorithms," *2018 41st International Conference on Telecommunications and Signal Processing, TSP 2018*, pp. 413–417, 2018, doi: 10.1109/TSP.2018.8441452.
- [4] M. Asadur Rahman, M. Faisal Hossain, M. Hossain, and R. Ahmmed, "Employing PCA and t-statistical approach for feature extraction and classification of emotion from multichannel EEG signal," *Egyptian Informatics Journal*, vol. 21, no. 1, pp. 23–35, 2020, doi: 10.1016/j.eij.2019.10.002.
- [5] A. Chater and A. Lasfar, "New approach to the identification of the easy expression recognition system by robust techniques (SIFT, PCA-SIFT, ASIFT and SURF)," *Telkomnika (Telecommunication Computing Electronics and Control)*, vol. 18, no. 2, pp. 695–704, 2020, doi: 10.12928/TELKOMNIKA.V18I2.13726.
- [6] G. F. Bin Mostafa, "Revisiting the Performance of PCA Versus FDA Versus Simple Projection for Image Recognition," *SSRN Electronic Journal*, pp. 1–16, 2020, doi: 10.2139/ssrn.3606738.
- [7] Z. Shi, D. Wu, J. Huang, Y.-K. Wang, and C.-T. Lin, "Supervised Discriminative Sparse PCA with Adaptive Neighbors for Dimensionality Reduction," 2020, [Online]. Available: <http://arxiv.org/abs/2001.03103>.
- [8] M. S. Khan, A. Ware, and A. Khan, "Face Acknowledgment using Principle Component Analysis (PCA) of Eigenfaces," *International Journal of Innovative Technology and Exploring Engineering*, vol. 9, no. 7, pp. 1196–1200, 2020, doi: 10.35940/ijitee.e2861.059720.
- [9] L. Hu and W. Zhang, "Orthogonal neighborhood preserving discriminant analysis with patch embedding for face recognition," *Pattern Recognition*, vol. 106, p. 107450, 2020, doi: 10.1016/j.patcog.2020.107450.
- [10] Q. Wang, Z. Qin, F. Nie, and Y. Yuan, "Convolutional 2D LDA for nonlinear dimensionality reduction," *IJCAI International Joint Conference on Artificial Intelligence*, pp. 2929–2935, 2017, doi: 10.24963/ijcai.2017/408.
- [11] W. Hu, W. Shen, H. Zhou, and D. Kong, "Matrix Linear Discriminant Analysis," *Technometrics*, 2020, doi: 10.1080/00401706.2019.1610069.
- [12] P. W. Wang and C. J. Lin, "Support vector machines," in *Data Classification: Algorithms and Applications*, 2014.
- [13] P. Lucey, J. F. Cohn, T. Kanade, J. Saragih, Z. Ambadar, and I. Matthews, "The extended Cohn-Kanade dataset (CK+): A complete dataset for action unit and emotion-specified expression," in *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops, CVPRW 2010*, 2010, doi: 10.1109/CVPRW.2010.5543262.
- [14] M. Lyons, S. Akamatsu, M. Kamachi, and J. Gyoba, "Coding facial expressions with Gabor wavelets," in *Proceedings - 3rd IEEE International Conference on Automatic Face and Gesture Recognition, FG 1998*, 1998, doi: 10.1109/AFGR.1998.670949.
- [15] M. Pantic, M. Valstar, R. Rademaker, and L. Maat, "Web-based database for facial expression analysis," in *IEEE International Conference on Multimedia and Expo, ICME 2005*, 2005, doi: 10.1109/ICME.2005.1521424.
- [16] J. Teuwen and N. Moriakov, "Convolutional neural networks," in *Handbook of Medical Image Computing and Computer Assisted Intervention*, 2019.