Volume 8, No. 3, March – April 2017

ISSN No. 0976-5697

International Journal of Advanced Research in Computer Science

RESEARCH PAPER

Available Online at www.ijarcs.info

Superiority of PCA algorithm for Facial Expression Recognition

Mamta Santosh
Junior Research Fellow: CSE
M.M. College of Engineering
Mullana, India
mloura112@gmail.com

Ashok Kumar
Professor: CSE
M.M. College of Engineering
Mullana, India
dr.ashok@mmumullana.org

Abstract: Facial expression recognition is a rapidly growing area of research in computer vision due to its immense applications such as human-computer interaction, psychology and intelligent robotic systems etc. Various methods have been proposed in past two decades aiming to solve different issues in this aspect. Most of the researches used distinct databases for the evaluation of their techniques, so it is very challenging task to infer which one is superior. In this paper, we have implemented three methods of facial expression recognition named: Principal Component Analysis (PCA), Artificial Neural Network (ANN) and Local Binary Pattern (LBP) and evaluated their performances using JAFFE database.

Keywords: Facial Expression Recognition, Emotion Recognition, JAFFE Dataset, PCA, LBP, ANN.

I. INTRODUCTION

With the advancement in the field of image processing and cameras, human facial expression recognition (FER) has become an important research area. Facial expressions serve as mean for communicating and depicting human emotion and moods. Real-time and automated facial expression help in a vast majority of application such as virtual reality, driver safety, video conferencing, machine-human interaction personality analysis, emotion analysis, health care applications, physiological, image retrieval, video analysis and many more applications. Facial expressions are one key aspect of the affect recognition systems. Different Psychologists have proposed different systems and measures that quantify and describe facial expressions and behaviors. Two of the most popular and widely used facial action coding system (FACS) is developed by Ekman and Friesen and Ekman et al. FACS provides the description of all 33 action units (AUs) that provide possible and visually detectable facial variations. AU or a combination of AU's can describe every human facial expression. Different methods are presented that are used for facial expression recognition and it was found that most of the research is focused on recognition of six basic emotions namely sadness, fear, anger, surprise happiness and disgust, defined by P. Ekman[15].

Fig. 1: Images with expressions from JAFFE Dataset [19]

Some researchers also used neutral expression in their research. In this paper, we focus on recognizing these seven basic facial expressions.The features that are normally used in FER system can be broadly categorized into two basic categories namely appearance-based features and geometry-based features.

Appearance-based feature rely upon the texture of the face created by different expression whereas Geometry-based features basically describe the shape of the different components of the face such as mouth, eyebrows, eyes etc. Some of the geometric features that are extracted to represent the physical shape of facial components include the distance between fiducial points.

In this paper we have examined three different FER algorithms namely Principal Component Analysis (PCA), Local Binary Pattern (LBP) and Artificial Neural Network (ANN). We have compared each method's performance i.e. their error rates.

II. LITERATURE SURVEY

Different appearance based features are utilized by different researchers for the recognition of facial expression. The combination of both appearance features and geometric information are also used by several for the purpose of FER.

One of the most widely used appearance features are texture features and within texture features, Local Binary Pattern (LBP) and its different variants are most commonly used for Facial Expression Recognition. A study conducted in [10] details different variants of LBP used in FER systems. Caifeng et al. used different classifiers such as Support Vector machine, Linear Discriminant Analysis, Linear Programming on LBP extracted features using different databases. R. Zhi [13] proposed Graph-preserving Sparse NMF (GSNMF) algorithm for Facial Expression Recognition. They derived Graph-preserving Sparse NMF from original NMF by exploiting both graph-preserving and sparse properties.

In geometric feature based approach, key facial points are first detected and then these facial points are tracked in a temporal sequence to recognize the expression. Ghimire et al. [7] detected 52 facial key points and modeled these facial key points in the form of points and lines features using multicast

Adaboost. SVM classifier was used for recognition of facial expressions.

A large number of different classification techniques have been used by different researchers for the purpose of facial expression recognition. In [11], authors used Artificial Neural Networks (ANNs) to classify facial expressions.With recent developments in machine learning, people are using deep learning, which integrates both feature extraction and learning procedure within deep networks, is being widely used for FER[6][2]. Heechul et al. compared convolutional neural network and deep neural network on CK+ database. They used Viola-Jones features for face detection and deep neural network for classification. They used OpenCV library for face detection and concluded that convolutional neural network performs better than deep neural network in real time facial expression recognition system [2].

Zhang et al. [12] proposed NN-based and multi-class Support Vector Machine (SVM) based classifiers. The Multi-class SVM with the radial basis function kernel enabled the robot to outperform the NN-based emotion recognizer. To train NN based classifier 380 training examples were used. 477 test examples were used to test the SVM classifiers. They evaluated that neural network based facial emotion recognizer performed with 76% accuracy for the detection of the six emotions.

Yeshudas et al. [1] proposed modified Local Binary Pattern algorithm for feature extraction and used neural network for classification. They used Viola-Jones algorithm for face detection on Japanese and Taiwanian database and implemented the system using matlab 2014a. Proposed modified LBP algorithm performed better for high size images.

III. METHODOLOGIES

A. *Pre Processing*

The first step in the preprocessing stage is the Face Detection. In this paper, Viola-Jones algorithm is used to detect faces from images. The Viola-jones object detection method was suggested by Paul Viola and Michael Jones in 2001[14] and is reported as real time object detection method. Although it can detect variety of objects, still it is most widely used for face detection. Viola-Jones requires full view frontal upright faces [8]. The method reads an input image with a window. It checks for facial featues, when enough features are discovered, then this window is accounted to be a face [9].

For diverse size faces, the window is scaled and the procedure is repeated. To decrease the number of features each window is passed through the levels. Early levels include less features and are easier to pass. More and more features are added as level increases. At each level, the value is calculated for the features, if the calculated value passes the threshold, the level is passed and this window is recognized as a face. Due to different image sizes, this method becomes time consuming in training. But once trained, it can detect the faces in real time. The stages in Viola-Jones face detection method are - Integral image, AdaBoost classifier learning and cascading classifier. The detected face is cropped from the rest of the image and is used for further processing.

To reduce processing time all the images are resized to 32x32. Local shadowing and highlights Gamma correction followed by histogram equalization are applied to eliminate illumination effect

(a) (b)

Fig. 2: (a) Original Image (b) Face detected cropped image

.

(a) (b)

Fig. 3: (a) Gamma Corrected Image (b) Histogram Equalized image.

B. *Principal Component Analysis with Euclidean Distance*

Principal component analysis is one of the most important method used for feature extraction and pattern recognition. PCA is a common statistical method which uses a holistic approach to find patterns in high dimensional data. PCA breaks down facial images into small sets of characteristic feature images. These feature images are called Eigenfaces and are used to represent both existing and new faces [4]. The 2-Dimensional face image matrices must be transformed into a 1-Dimensional vector. The 1-Dimensional vector can be either row or column vector. [5]

PCA Method Steps are as Follows [4][16]:

- Consider data D having M x N dimensions where M is the number of instances and N is the number of features. To calculate PCA first the mean of data is calculated.
- the mean is subtracted from each feature vector to center the data.
- Calculate the Covariance Matrix as shown in the equation below:

$$\text{Covariance} = \sum_{n=1}^{M} \text{Sub}(\text{n})\, Sub^{T}(\text{n})$$

- Calculate Eigenvectors and their corresponding Eigenvalues.
- Sort the Eigenvectors in the descending order according to their corresponding eigenvalues.
- Each of the mean centered data is then projected into eigenspace.

Once the features are calculated, the next step is to measure the distance between images. Euclidean distance is used for distance measurement between images. Euclidean Distance is defined as the straight line distance between two points, which examines the root of square differences between the coordinates of a pair of objects [16].

Euclidean Distance can be calculated using the equation below:

$$EuclideanDistance(X,Y) = \sqrt{\sum_{n=1}^{\text{No of Images}} (X_n - Y_n)^2}$$

Where n represents the total number of features in the feature vector.

C. Local Binary Pattern (LBP) and Euclidean Distance

Local Binary Features (LBP) was first described in 1994 and is widely used for feature extraction and classification in computer vision. LBP divides the window to be examined into cells. Each pixel in a cell is compared to each of its 8 neighbors (on its left-top, left-middle, left-bottom, right-top, etc.) subtracting the center pixel value. Where the neighbor pixel's value is less than the center pixel's value, it is encoded as 0 the others with 1;

By concatenating all these binary codes in a clockwise direction, a binary number is calculated. This binary number is converted to its corresponding decimal value which is used for labeling. These binary numbers are referred to as Local Binary Patterns.

In other words, given a pixel position (xc, yc), LBP is defined as an ordered set of binary comparisons of pixel intensities between the central pixel and its surrounding pixels. The resulting decimal label value of the 8-bit word can be expressed as follows [3]:

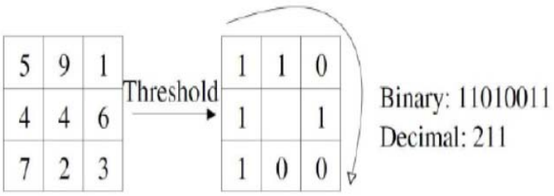


Fig. 4: An example of the basic LBP operator

Where lc corresponds to the grey value of the centre pixel (xc, yc), ln to the grey values of the 8 surrounding pixels, and functions (k) is defined as [18]:

$$s(k) = \begin{cases} 1 & if\ k \geq 0 \\ 0 & if\ k < 0 \end{cases}$$

Many extensions of LBP have been proposed in literature such as OCLBP, tLBP, dLBP, mLBP etc.

Here again Euclidian distance was used as a classifier.

D. Artificial Neural Network

Artificial Neural Networks are one of the most widely used classifiers in the field of pattern recognition. An Artificial Neural Network is mathematical model which mimics a biological brain. A neural network consists of an interconnected group of artificial neurons. These neurons are connected to each other through links called axons. ANN changes its structure during a learning phase and adjusts its synaptic weights. Neural networks are used to model complex relationships between inputs and outputs or to find patterns in data [17].

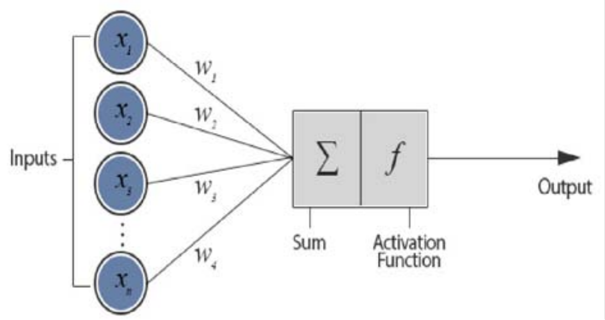


Fig. 5: An Artificial neural Network

A neural network is made up of a set of nodes that are connected through a set of links. The nodes represent the neurons of the network and the connections of data flow between neurons are represented by the links. Connections are quantified by weights, which are adjusted during training to maximize the accuracy of the system. During training phase, a subset of dataset with corresponding labels is provided to the network. Each training instance is represented by a feature vector with all instances having same number of features (called an input vector). Every input feature vector must be connected to a output class which is represented by output vector.

Firstly, the gray scale image is converted to binary image by thresholding the image. The simplest thresholding methods replaces each pixel value in an image with a 0 i.e turn it to a black pixel if the image intensity is less than some fixed constant threshold T or 1 i.e. a white pixel if the image intensity value is greater than that of constant T.

Where g(x,y) is the resultant image and f(x,y) is the input image. Then the binary image is converted to 1024 length 1-D vector to use as an input to neural network

IV. EVALUATION

Japanese Female Facial Expression Database (JAFFE) [19] is used for implementation of techniques. It was created and planned Michael Lyons, Jiro Gyoba, and Miyuki Kamachi. It consists of a total of 213 images approximately equally divided into seven categories i.e. anger, sadness, surprise, neutral, fear and disgust. All the expressions are performed by ten different female Japanese posers. All images have a size of 256 x 256 pixels.

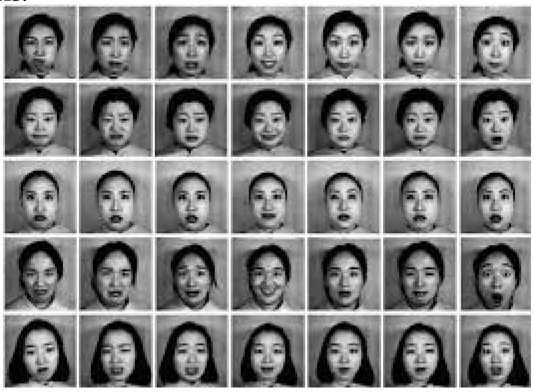
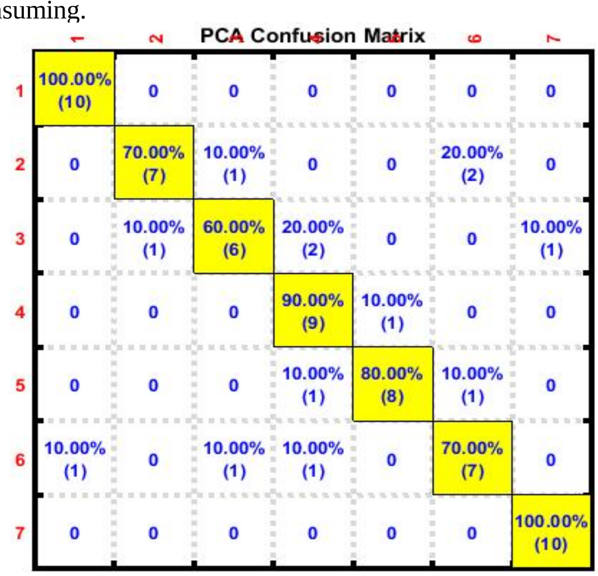
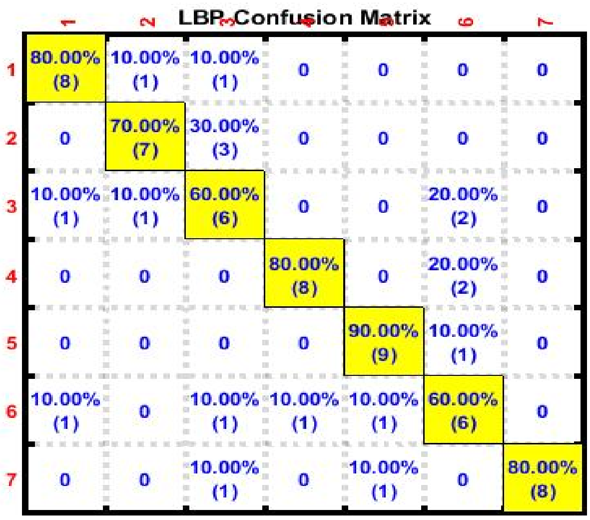


Fig. 5: JAFFE dataset samples

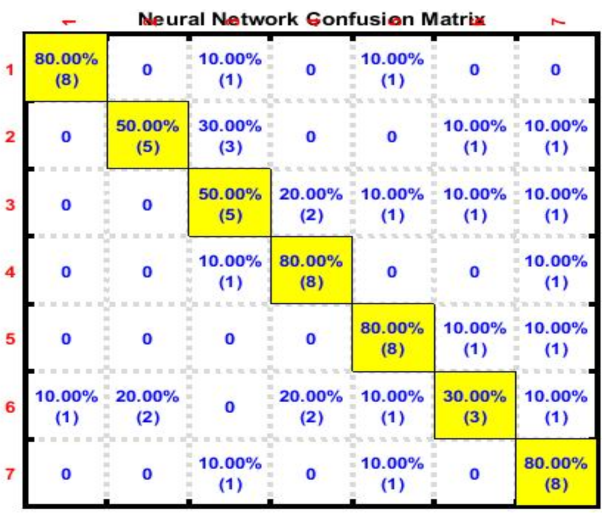
In our project we used 70 images i.e 10 images from each class for testing and all the other images for training purpose. Figure 6 shows the confusion matrix for each algorithm. . The accuracy was 81.42%, 74.28%, and 64.28% for PCA with Euclidean classifier, LBP with Euclidean classifier, and neural networks respectively. PCA with Euclidean distance gives the best accuracy. Not only is it fast but also the least time consuming.



(a) PCA with Euclidean classifier



(b) LBP with Euclidean classifier



(c)Artificial neural networks

Fig 6: Confusion matrices for each for the different classifiers used

V. Conclusion

In this project paper, we have explained three different techniques for facial expression recognition: using Principal Component Analysis and Euclidean distance, Local Binary Pattern and Euclidean distance and Artificial Neural Network. We have evaluated each classifier we developed against the JAFFE expression database. We have compared the results and found PCA with Euclidean distance go give the best accuracy.

VI. Acknowledgement

We are thankful to Michael J. Lyons, Shigeru Akamatsu, Miyuki Kamachi & Jiro Gyoba for providing us the JAFFE Database.

We are thankful to University Grand Commission for providing Junior Research Fellowship for carrying out research work.

VII. References

[1] Y. Muttu and H. G. Virani, "Effective face detection, feature extraction & neural network based approaches for facial expression recognition," 2015 International Conference on Information Processing (ICIP), Pune, 2015, pp. 102-107.

[2] H. Jung et al., "Development of deep learning-based facial expression recognition system," 2015 21st Korea-Japan Joint Workshop on Frontiers of Computer Vision (FCV), Mokpo, 2015, pp. 1-4.

[3] Ruaa Adeeb Abdul Munem Al-falluji. DME Detection using LBP Features. International Journal of Computer Applications 148(8):44-48, August 2016.

[4] Nawaf Hazim Barnouti, Sinan Sameer Mahmood Al-Dabbagh, Wael Esam Matti and Mustafa Abdul Sahib Naser, "Face Detection and Recognition Using Viola-Jones with PCA-LDA and Square Euclidean Distance" International Journal of Advanced Computer Science and Applications(IJACSA), 7(5), 2016.

[5] Barnouti, Nawaf Hazim. 2016. Face Recognition Using Eigenface Implemented On DSP Processor, International Journal of Engineering Research and General Science, 4(2), pp.107-113.

[6] P. Liu, S. Han, Z. Meng, Y. Tong, "Facial expression recognition via a boosted deep belief network," in Proc. IEEE Conf. on CVPR, pp. 1805-1812, 23-28 June 2014.

[7] D. Ghimire, J. Lee, "Geometric feature-based facial expression recognition in image sequences using multi class AdaBoost and support vector machines," Sensors, vol. 13, pp. 7714-7734, 2013.

[8] Cha, Z. and Zhengyou, Z., 2010. A survey of recent advances in face detection. Microsoft Research, Microsoft Corporation.

[9] Hefenbrock, D., Oberg, J., Thanh, N.T.N., Kastner, R. and Baden, S.B., 2010, May. Accelerating viola-jones face detection to fpga-level using gpus. In 2010 18th IEEE Annual International Symposium on Field-Programmable Custom Computing Machines (pp. 11-18). IEEE.

[10] C. Shan, S. Gong, and P.W. McOwan, "Facial expression recognition based on local binary patterns: A comprehensive study," Image and Vision Computing, vol. 27, pp. 803-816 , 2009.

[11] O. Rudovic, M. Pantic, and I Patras, “Coupled Gaussian processes for pose-invariant facial expression recognition,” IEEE Trans. Pattern Anal. and Mach. Intell., vol. 35, no. 6, pp. 1357-1369, June 2013.

[12] Li Zhang; Hossain, A.; Ming Jiang, "Intelligent Facial Action and emotion recognition for humanoid robots," in Neural Networks (IJCNN), 2014 International Joint Conference on , vol., no., pp.739-746, 6-11 July 2014

[13] Ruicong Zhi, M. Flierl, Qiuqi Ruan and B. Kleijn, "Facial expression recognition based on graph-preserving sparse non-negative matrix factorization," 2009 16th IEEE International Conference on Image Processing (ICIP), Cairo, 2009, pp. 3293-3296.

[14] P. Viola and M. J. Jones, “Rapid Object Detection using a Boosted Cascade of Simple Features,” in Proceedings of the IEEE Computer Society International Conference on Computer Vision and Pattern Recognition, vol. 1, pp. 511-518, 2001.

[15] P. Ekman, “The argument and evidence about universals in facial expressions of emotions,” Handbook of Social Psychophysiology, Wiley, Chichester, pp. 143-164, 89.

[16] M. Schels, and F. Schwenker, “A multiple classifier system approach for facial expressions in image sequence utilizing GMM Supervectors,” in Proc. of the 2010 10th Int. Conf. on Pattern Recog., pp. 4251-4254, August 2010.

[17] LiMin Fu, Neural Networks in Computer Intelligence, Ed-McGraw-Hill. 1994.

[18] Ojala, T., M. Pietikainen, and T. Mäenpää. "Multiresolution Gray Scale and Rotation Invariant Texture Classification With Local Binary Patterns." IEEE Transactions on Pattern Analysis and Machine Intelligence. Vol. 24, Issue 7, July 2002, pp. 971-987.

[19] Michael J. Lyons, Shigeru Akamatsu, Miyuki Kamachi, Jiro Gyoba.Coding Facial Expressions with Gabor Wavelets, 3rd IEEE International Conference on Automatic Face and Gesture Recognition, pp. 200-205 (1998)