



doi: <https://doi.org/10.20546/ijcrar.2018.605.004>

## Principal Component Analysis and K-Nearest Neighbor Classifier for Facial Expression Recognition System

**Firas Husham Almukhtar\***

*Department of Computer Technical Engineering, Imam Ja'afar Al-Sadiq University, Kirkuk - Iraq*

*\*Corresponding author*

### Abstract

In the last decades, facial expressions recognition systems become one of the most interesting research challenges. Many techniques proposed to achieve these challenges. In this paper, Principal Component Analysis (PCA) is adopted to take the facial features out of the input image, so as these features will be used for the K-Nearest Neighbor (K-NN) classifier to classify it into its nearest categories. The proposed method is tested based on two types of databases (JAFFE (Japanese Female Facial Expressions) and KDEF (The Karolinska Directed Emotional Faces)). Each of these databases has many categories of facial expressions images (6 basic face expressions and one neutral expression) labelled from No.1 to No.7 such that each number represents one class category. The experimental result shows the robustly and feasibility of our suggested system during the identification and the classification step to assign the new tested image to its suitable class label.

### Article Info

Accepted: 20 April 2018  
Available Online: 20 May 2018

### Keywords

Facial Expressions,  
Recognition system,  
Principal Component Analysis,  
K-NN.

### Introduction

Body language is an important aspect in a nonverbal human communications. It consists of Facial expressions as well as the postures of all the rest of the human body. Facial expressions or emotions play a significant role in our daily life to give the impressions of either ("Neutral", "Happy", "Surprise", "Angry", "Sadness", "Fear" and "Digest"). These expressions also work in the identification of the people in addition to their voice tone (Mathur and Pandey, 2017).

Automatic system to recognize facial expressions is a need nowadays to check the face expressions according to the basic seven impressions mentioned above. Many researchers and academies give a great interest in the

field of Facial expressions with different methods. All of these methods are concentrating on how to extract the facial features such that these features then submitted to the later step for classification. The role of these features is so important to the rate of classification (Ashir and Akdemir, 2018).

One of the most popular and famous methods for extracting facial features are "Eigenface" technique. Eigenface used by an algorithm called Principle Component Analysis (PCA) that transforms the facial image and reduces its dimensionality into a low dimensional space. PCA converting the image matrix into a reduced single orthogonal vector with some special values such that this vector called the Eigenface. Applying PCA to more than one image as a training set,

lead to produce a linear matrix that has some data variances used for the later step (Pumlumchiak and Vittayakorn, 2017).

The second step of the automatic recognition system is the classification step. There are many methods used as classifiers to classify the input image according to the matrix of data variances result from the PCA step. One of these methods or algorithms is the Nearest Neighbor (NN). In the NN method, one set of samples are handled such that the samples can be classified by minimizing the distance between the samples, after that these samples can be represented in different classes. In the NN method, it classifies the testing samples according to prediction. It predicts the best linear representation for each class (Mustafa Zuhaer Nayef Al-Dabagh *et al.*, 2018).

The structure of this paper consists of many sections. The first section includes the introduction; Section II explains the design of the facial expression recognition system, involving Principal Component Analysis method and the K-Nearest Neighbor method. Section III reviews the results of the proposed system performance by applying the set of training images with the sample that we want to investigate. Section IV includes work conclusions.

**System design**

Three stages performed in our proposed system; these are preprocessing, features extraction and classification. The first step (Preprocessing): median filter and histogram equalization are performed on the face image to make it look better especially according to the contrast and the brightness levels. The second step concerned with the application of feature extraction method using the PCA. This method can take unique features from the image, then send these features to the third step (K-Nearest Neighbor) to give the suitable classification of the face image. Figure 1 describes the steps of the proposed system as follows:-

**Pre-processing**

The initial step performed in our recognition system, was the preparation of the input image for further processing by applying the median filter, just to clarify the input image against any dust or noise. This step is followed by the histogram equalization that improves the illumination of the image by distributing the colors in a normalized way to guarantee that the features of the facial image will

be extracted in an easy and effective way (Firas AL-Mukhtar and Mustafa Zuhaer Nayef AL-Dabagh, 2017).

**Feature extraction using PCA**

A mechanism that decreases the size of the valuable interconnected datasets into a small set of unconnected linear data. This mechanism called Principal component analysis (PCA). The process of this mechanism can be illustrated as a vertical linear transformation of data converted into a new form such that the extreme change can be noticed on the first occurrence (called the prime principal component), while the second extreme change happens on the second occurrence, etc. (Meher and Maben, 2014). Mathematically PCA algorithms can be constructed using the following formulas:-

Let us suppose that X represents the set of facial pictures to be trained:

$$X_1, X_2, X_3, X_4, \dots, X_n \quad (1)$$

Then the average facial image from all other images is obtained by:

$$A = \frac{1}{n}(X_1 + X_2 + \dots + X_n) \quad (2)$$

Each image differs from the average one by:-

$$D = |X_1 - A| \dots \dots \dots |X_n - A| \quad (3)$$

Now calculate the covariance matrix, Co as a matrix:

$$Co = \frac{1}{n-1} DD^T \quad (4)$$

This matrix has morthogonal vectors; each represents some features of the facial images called Eigenfaces. Each vector contributes in the original set by a weight vector V, such that:

$$V_m = u_m^T D_m \quad (5)$$

Where u represent the features of the facial images orthogonal vectors.

Fig.1 System steps

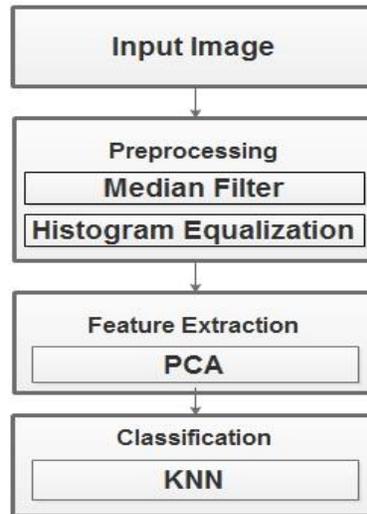


Fig.2 Shows a sample of JAFFE databases



Fig.3 Samples of KDEF databases



**Recognition using K-nearest Neighbor**

This method work to classify the given features among a set of features obtained from training database examples. Implementing K-nearest Neighbor method is by calculating the distance between the training vectors obtained from the step above with the given facial image

input just to show the similarity. This method is implemented using the Euclidian rule that decides how far is the influence of the input facial picture according to the training data influence, such as:-

$$E_m = \sqrt{(V - V_m)} \quad (6)$$

Depending on the above equation, the minimum difference is then chosen and the input facial image is given the label of nearest neighbor class (Kumar and Ali Basha, 2014).

### Experiments and Results

We worked with MATLAB version 16 to implement the recognition system. To evaluate the performance of this system, two good example databases experimented; these are JAFFE (<http://www.kasrl.org/jaffe.html>) (Japanese Female Facial Expressions) and KDEF (Goeleven *et al.*, 2008) (The Karolinska Directed Emotional Faces). These two databases work as the samples set and as the practicing set. These databases were free of charge and are available on the Internet for the academic and research purpose. The JAFFE database includes 63 images of seven different facial poses done by nine Asian woman models. These images were taken with a white homogeneous background and in different lighting conditions. The size of each image was (256 x256) in dimensions, in the class of gray color shades (Fig. 2 and 3).

The KDEF database contains 239 colored pictures for seven face expressions (one neutral pose and six different expressions) cast by 17 female models and 17 male models. These images are also taken with different homogeneous background and in different lighting conditions. The size of each image was (256 x 256).

In this paper, a system of facial expressions recognition based on Principal Component Analysis with K-nearest neighbor (K-NN) classifier is proposed. To evaluate the system performance, Extensive experiments conducted on a database called JAFFE and KDEF. The practical outcomes have proven the good results to work in diverse conditions such as different lighting exposure, facial details, and facial expressions. The results show that the system has a high recognition rate with the accuracy of up to 92.036% on JAFFE databases and 94.1% on KDEF Database. In the future work, we can use a method like SVM (Support Vector Machine) to enhance the process of classification. Another method to enhance the lightening of the image may be performed so to improve the accuracy rate.

### References

- Ashir, A. M., and B. Akdemir, "Facial expression recognition with an optimized radial basis kernel," 2018 6th International Symposium on Digital Forensic and Security (ISDFS), Antalya, Turkey, 2018, pp. 1-6.
- Firas AL-Mukhtar and Mustafa Zuhaer Nayef AL-Dabagh. "Real-Time Face Recognition System Using KPCA, LBP and Support Vector Machine ", International Journal of Advanced Engineering Research and Science, Vol. 4, No.2, 2017, pp. 184-189.
- Goeleven, E., De Raedt, R., Leyman, L. and Verschuere, B., 2008. The Karolinska directed emotional faces: a validation study. *Cognition and Emotion*, 22(6), pp. 1094-1118.
- <http://www.kasrl.org/jaffe.html>
- Kumar, V., and A. S. Ali Basha, "Facial expression recognition using wavelet and k-nearest neighbor," Second International Conference on Current Trends in Engineering and Technology - ICCTET 2014, Coimbatore, 2014, pp. 48-52.
- Mathur, J., and U. S. Pandey, "Facial expression recognition using wavelet-based Support Vector Machine," 2017 Recent Developments in Control, Automation & Power Engineering (RDCAPE), NOIDA, India, 2017, pp. 275-279.
- Meher, S.S., and P. Maben, "Face recognition and facial expression identification using PCA," 2014 IEEE International Advance Computing Conference (IACC), Gurgaon, 2014, pp. 1093-1098.
- Mustafa Zuhaer Nayef Al-Dabagh, Mustafa H. Mohammed Alhabib, and Firas H. AL-Mukhtar, "Face Recognition System Based on Kernel Discriminant Analysis, K-Nearest Neighbor and Support Vector Machine," International Journal of Research and Engineering Vol.5, No.3, 2018, pp. 335-338.
- Pumlumchiak, T., and S. Vittayakorn, "Facial expression recognition using local Gabor filters and PCA plus LDA," 2017 9th International Conference on Information Technology and Electrical Engineering (ICITEE), Phuket, 2017, pp. 1-6.

#### How to cite this article:

Firas Husham AlMukhtar. 2018. Principal Component Analysis and K-Nearest Neighbor Classifier for Facial Expression Recognition System. *Int.J.Curr.Res.Aca.Rev.* 6(5), 17-20.

doi: <https://doi.org/10.20546/ijcrar.2018.605.004>