# Extension of the MPEG-7 Fourier Feature Descriptor for Face Recognition using PCA


**Naser Zaeri**
*Center for Vision, Speech and Signal Processing*
*University of Surrey*
*Guildford*
*Surrey*
*GU2 7XH*
*UK*
*n.zaeri@surrey.ac.uk*

**Farzin Mokhtarian**
*Center for Vision, Speech and Signal Processing*
*University of Surrey*
*Guildford*
*Surrey*
*GU2 7XH*
*UK*
*f.mokhtarian@surrey.ac.uk*

**Abdallah Cherri**
*Electrical Engineering Dept.*
*Kuwait University*
*P.O. Box 5969*
*Safat 13060*
*Kuwait*
*cherri@eng.kuniv.edu.kw*



## Abstract

*The Principal Component Analysis or the eigenface technique provides a practical solution to the problem of face recognition. Recently, many face descriptors for MPEG-7 have been proposed for face retrieval in video streams. In this paper, a new method for face recognition is presented based on extracting the most discriminant features of the MPEG-7 Fourier Feature Descriptors of the face space, defined by MPEG-7 face recognition technique, through the implementation of the eigenface technique. It will be demonstrated that the proposed method improves the recognition rate and copes better with pose variations under different facial expressions and varying face conditions, as well as illumination variations. In addition, the proposed method achieves substantial savings in the computation time needed by the recognition system.*


## 1. Introduction

Face recognition finds many important applications in many life sectors and in particular in commercial and law enforcement applications. Although humans have the ability to recognize faces in cluttered scene with relative ease, machine recognition is a much more difficult dilemma, where an input face image is compared against models in a library of known faces. Over the last 20 years researchers from different fields have investigated a number of issues related to face recognition. In general, the process of recognizing a face passes through many steps such as segmentation of faces from cluttered scenes, extraction of face features, identification, and matching.

Many different techniques have been proposed for face recognition [1]. Among them are two traditional classes of techniques [2-5] applied to the recognition of digital images of faces. The first class of techniques is based on the computation of a set of geometrical features from the picture of a face, where the overall geometrical configuration of the face is described by a vector of numerical data to represent the position and size of the main facial features: eyes and eyebrows, nose, and mouth. The second class of techniques is based on template matching, where a template is face-like grey scale image found from processing the face as a whole. Furthermore, neural network-based, classical correlation templates, wavelet-based and probabilistic –based algorithms to recognize faces are found in the literature [6-17].

In addition, facial features extraction to find the most appropriate representation of face images is one of the most important components of a face recognition system for identification purposes. The eigenface – based technique is one of the well known approaches to achieve a good representation of a face image [3, 4]. Representing the data in a low-dimensional space has always been a main concern in the last decade.

Since the total number of pixels in a face image is typically large (on the order of several thousands even for small image sizes) and since the face images in such a high-dimensional space are not randomly distributed, therefore, it would be very efficient and beneficial to project them to a lower-dimensional subspace using principle component analysis (PCA) . In PCA the set of all face images is considered as the vector space, where the eigenfaces are the dominant principal component of the covariance matrix of the data. Thus, eigenface approach produces a compact representation - a feature vector with a few elements that can concisely represent a facial image.

Recently, MPEG-7 has inspired a wide range of applications, with face recognition being one of them. MPEG-7 is an ISO Standard developed by the Moving Picture Experts Group (MPEG) over the period 1999-2001 [18]. The MPEG committee is primarily known for the successful development of a series of video compression Standards: MPEG-1, MPEG-2, and MPEG-4. Face descriptors for MPEG-7 have been proposed for face retrieval in video streams.

In this paper, we describe a new approach for face recognition system by extracting the most discriminant features of the MPEG-7 Fourier Feature Descriptors of the face space through the implementation of the eigenface technique.

It will be shown that the new technique:

(i) reduces the size of face space dramatically compared to applying the PCA alone and

(ii) improves the performance of the recognition rate compared to the MPEG-7 FFD vector method.

Note that this technique deals with images of different illumination, facial expressions, and poses.

This work is organized as follows. A brief description of MPEG-7 Fourier Feature descriptor is given in Section 2. The formulation of the proposed technique is presented and discussed in Section 3. In Section 4, results of testing and implementing the new technique on two databases; the Olivetti Research Ltd. (ORL) database and xm2vts, CVSSP – University of Surrey database, are presented. Conclusions are given in Section 5.

## 2. MPEG-7 Fourier Feature Descriptor

The MPEG-7 objective is to describe the content of multimedia data, so that it can be efficiently searched, accessed, transformed or adapted for the use by any device and to support different applications.

The MPEG-7 uses the FFD vector to represent the facial feature of an image by a small single vector derived from two feature vectors; one is a Fourier Spectrum Vector $x_1^f$ , and the other is a Multi-block Fourier Amplitude Vector $x_2^f$ of a normalized face image. The Fourier Spectrum Vector describes the image globally, whereas the Multi-block Fourier Amplitude Vector describes the image locally through dividing the image into blocks. More detailed description of the FFD can be found in [19].

Fig. 1 shows the process of extracting the FFD for a single image. The quantized elements representing the FFD vector, $W_f$ is of size 63.

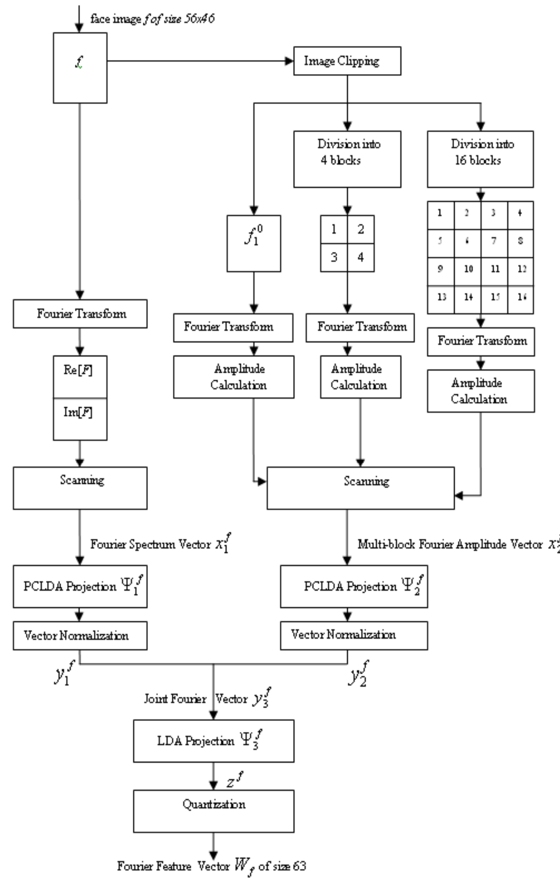


**Figure 1. Schematic diagram for the process of Fourier Feature extraction for a single image in the MPEG-7 algorithm**

## 3. Principal Component Analysis of Fourier Feature Descriptor Vectors

The proposed system is based on applying the principal component analysis to the existing Fourier Feature Descriptor of the MPEG-7 algorithm. The PCA technique, proposed by Turk and Pentland [3, 4], extracts the relevant information in a face image, encodes it as efficiently as possible, captures the variation in a collection of face images, and compares one face encoding with a database of models encoded similarly. The images of faces, being similar in overall configuration, will not be randomly distributed in the huge image space and, consequently, they can be described by a relatively low dimensional subspace.

This process is achieved by finding the principal components of the distribution of faces, or the eigenvectors of the covariance matrix of the set of face images. The eigenvectors are ordered, each one accounting for a different amount of the variation among the face images. These eigenvectors can be thought of as a set of features that together characterize the variation between face images. Each image contributes more or less to each eigenvector, so that the eigenvector is displayed as a sort of ghostly face which is called an eigenface. Example of some of the images used from the ORL database and some of the resultant eigenfaces using the conventional PCA is shown in Figs. 2 and 3, respectively.

**Fig. 2. Example of some of the images used from the ORL.**

**Fig. 3. Some of the eigenvectors using conventional PCA.**

Next, we will apply the PCA to the FFD vectors found for the images using the MPEG-7 algorithm instead of applying it to the images directly. Fig. 4 shows the new proposed system.

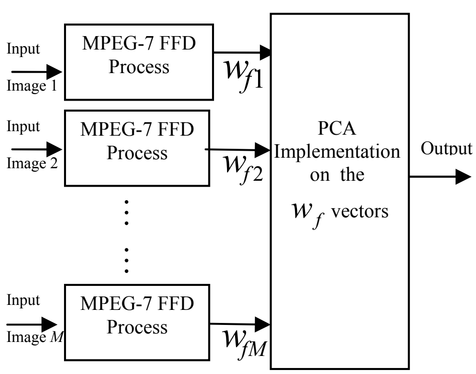


**Figure 4. Block diagram for the proposed system.**

An $N \times N$ face image is rearranged as a vector of dimension $N^2$ so that a typical image of size 256×256 becomes a vector of dimension 65,536. Accordingly, in the conventional PCA, the covariance matrix $C_s$ of the recognition system has a size of $N^2 \times N^2$. Next, one needs to determine the $N^2$ eigenvectors of this huge covariance matrix $C_s$, which proves to be a very complicated task. Turk and Pentland suggested to find the eigenvectors of $M \times M$ matrix (where $M<<N$, is the number of images used to construct the covariance matrix) and then made appropriate transformation or calculation to get the eigenvectors of the matrix $C_s$. However, in general, the value of $M$ will rarely be less than 63 which is the size of FFD vector of the MPEG-7 face recognition code as shown in Fig. 1.

Next, the PCA technique is applied to the FFD's, which were obtained from the MPEG-7 algorithm. For a set of FFD vectors $W_{f1}, W_{f2}, W_{f3}, \ldots, W_{fM}$, the average FFD vector of the set is defined by

$$m = \frac{1}{M}\sum_{n=1}^{M} W_{fi} \qquad (1)$$

Each $W_f$ vector differs from the average by the vector $\Phi_i = W_{fi} - m$. This set of vectors is subject to PCA. The covariance matrix of the data space is given by

$$C_s = AA^T \qquad (2)$$

where $A = [\Phi_1 \ \Phi_2 \ldots\ldots \Phi_M]$. Eq. (2) can be rewritten as

$$C_s = \sum_{i=1}^{M} (W_{fi} - m)(W_{fi} - m)^T \qquad (3)$$

The PCA is applied to this matrix and the eigenvectors with the corresponding highest eigenvalues are chosen to represent the input database. The PCA finds the vectors that best account for the distribution of these vectors within the entire FFD space.

As a result, the covariance matrix $C_s$ will be of size 63×63. This is much smaller than the size $M \times M$ of the covariance matrix that is obtained using the conventional PCA. In practice, a value of $M'$ (smaller than the number of images $M$ and the FFD vector size 63) is sufficient for identification, since accurate reconstruction of the image is not required. Consequently, these significant eigenvectors span an

$M^{'}$ - dimensional subspace of the original image space $N^2$ . The $M^{'}$ significant eigenvectors are chosen as those with the largest associated eigenvalues.

Now, the FFD vector ($W_f$) is transformed into its eigenvector components by the following weight equation,

$$\omega_k = u_k^T (W_f - m) \quad (4)$$

where $u_i$ is the $i^{th}$ eigenvector and $k = 1, ...., M^{'}$. The weights form a vector $\Omega^T = [\omega_1, \omega_2, ....., \omega_{M'}]$ that describes the contribution of each eigenvector in representing the input FFD vector, and consequently the $\Omega^T$ represents the corresponding (and the original) face image.

It is worth mentioning that finding the eigenvectors and eigenvalues of the covariance matrix of the proposed method has become a much easier task. This can be easily seen by noting the amount of vector size reduction from $N \times N$ (the image size) which is usually a vector of length not less than ~1000 elements (an image of size 32×32 will give a vector of size 1024) to a vector of length 63 elements. Consequently, there is no need to treat the covariance matrix in any special way (as suggested by Turk and Pentland) to facilitate the task of finding the eigenvectors.

The simplest method for determining which face class provides the best description of an input face image is to find the face class $k$ that minimizes the Euclidian distance

$$\varepsilon_k = \|(\Omega - \Omega_k)\|^2 \quad (5)$$

where $\Omega_k$ is a vector describing the $k^{th}$ FFD of a face class. These classes are calculated by averaging the results of the eigenvector representation over a small number of FFD vectors (as few as one) of each individual. A face is classified as belonging to class $k$ if the corresponding $\varepsilon_k$ is the minimum among all other $\varepsilon_k$'s.

# 4. Experiments

Without any pre-processing steps and in order to test the performance of the proposed technique, many different experiments were performed. We carried out experiments on two independent and different databases, one is the ORL and the other is the xm2vts database. Both sets include a number of images for each person, with variation in poses, expression and the lighting. The ORL set includes 400 images of 40 different individuals with each individual represented by 10 images. For the xm2vts set, we have arbitrarily chosen 400 images for 100 different individuals with each individual represented by 4 different images.

The following steps were carried out on both sets:

(i) The MPEG-7 algorithm was applied to these sets, and the FFD for each one of the images was calculated.

(ii) The eigenvectors and eigenvalues of the (63×63) covariance matrix were calculated, the $M'$ eigenvectors corresponding to the highest associated eigenvalues is chosen. $M'$=10 was selected in the experiments.

(iii) For each known individual, the class vector $\Omega_k$ was calculated by averaging the pattern (weight) vectors $\Omega$ calculated from the original FFD vector of each individual.

(iv) For each new face image to be identified: its pattern vector $\Omega$, the distances $\varepsilon_k$ to each known class was calculated. The class vector $\Omega_k$ that has the minimum distance $\varepsilon_k$ will represent this input face.

As a first experiment, 400 images from the xm2vts database are taken at different sessions (different days). This experiment tests the robustness of the proposed system under the variation in *time* conditions of the images. Different timing means different hair style, different clothes and different "modes". Fig. 5 shows examples of the xm2vts database used for this experiment.

The recognition rate was ~82% for this experiment, while under the same conditions the MPEG-7 face recognition method achieved ~79 %.

**Figure 5. Examples from the xm2vts database used in the evaluation of the system.**

The other experiment was to test the proposed technique under different circumstances. The ORL face database is used in this experiment. This database include images with different poses, different illumination, different expression (open or closed eyes, smiling or non smiling), different facial details (glasses or no glasses), and some of them were taken at different times. Examples of the ORL database used are shown in Fig.6.

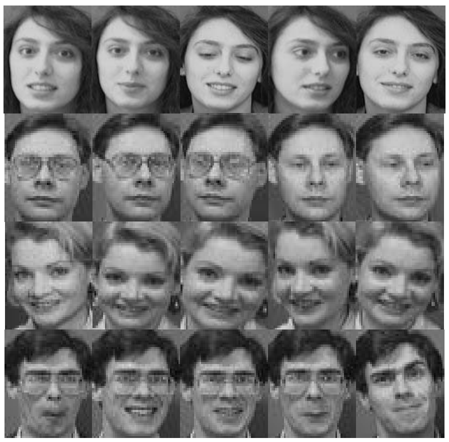

**Figure 6. Examples from the ORL database used in evaluation of the system.**

The proposed technique achieved ~85 % correct classification, while under the same conditions the MPEG-7 face recognition method achieved ~80 % correct classifications. It is of great importance to stress the fact that the new technique is building and designing its decisions based on "representative vectors", $W_f$ , and not the whole image database, i.e., the new technique *further compresses the data space which already had been compressed by MPEG-7.* A summary of the results of the recognition test on both databases is given in Table 1.

**Table 1. Summary of the results.**

| | MPEG-7 | New method |
|---|---|---|
| ORL | 80% | 85% |
| xm2vts | 79% | 82% |

Table 2 demonstrates the amount of time saving in the mathematical operations to find the eigenvectors and eigenvalues of the covariance matrix of the data space for both the conventional PCA and the new method.

**Table 2. Summary of calculations and consumed time for a certain example.**

| | Conventional PCA | Conversion Suggested by Turk and Pentland | New method |
|---|---|---|---|
| Number of multiplication | $N^2 \times N^2 \times M$ | $M \times M \times N^2$ | $63 \times 63 \times M$ |
| Number of summation | $N^2 \times N^2 \times (M-1)$ | $M \times M \times (N^2-1)$ | $63 \times 63 \times (M-1)$ |
| $\tau$ =time | Exceeds Machine Capabilities* | 37 sec. | 0.01 sec. |

* Using Pentium 4 – 2.8MHz – 512k RAM

where $\tau$ = is the time taken by the machine to find the eigenvectors and eigenvalues for the corresponding data space, $N\times N$ = image size, and $M$ = # images used to construct the data space. For example, if we have an image with size 32×32 and if the number of images used is 100 images, then the number of multiplications and summation needed to construct the covariance matrix using the conventional PCA is of order $O(10^8)$. The corresponding number using the conversion suggested by Turk and Pentland is of order $O(10^7)$. On the other hand, the number is of order $O(10^5)$ for the new method.

## 5. Conclusion

Although the conventional PCA has demonstrated a good performance in face recognition tasks, it takes a significant amount of time in calculating the eigenvectors of the covariance matrix; a problem that we have overcome in our new technique. In this paper, a new method is proposed for face recognition by modifying the MPEG-7 Fourier Feature Descriptor. The system takes the advantages of the PCA technique and applies it to the FFD.

An important feature of the proposed technique is the dimensionality reduction. The first dimensionality reduction was achieved by projecting the image into its Fourier Feature Descriptor vectors, and the second dimensionality reduction is accomplished through implementing the PCA on the FFD vectors. As a result, the process of finding the principal eigenvectors that defines the face recognition system is simplified, as the calculations related to the eigenvectors of the covariance matrix are reduced. The proposed technique keeps the most discriminant projection direction embedded in the FFD subspace as PCA is applied to MPEG-7 FFD vectors. Moreover, the new system provides better recognition rate compared to the MPEG-7 technique.

Results of applying the new system on two independent databases have been demonstrated. The new method has provided a platform for developing new face recognition algorithms.

## References


[1] R. Chellappa, C. L. Wilson, and N. Sirohey "Human and machine recognition of faces, a survey", *Proc. IEEE* 83, 705-740 (1995).

[2] B. S. Manjunath, R. Chellappa, and C. v. d. Malsburg, "A feature based approach to face recognition", *in Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (IEEE Computer Society, Los Alamitos, Calif., 1992), pp. 373-378.

[3] Turk and Pentland, "Eigenfaces for Recognition", *Journal of Cognitive Neuroscience*, March 1991.

[4] Turk and Pentland, "Face Recognition using Eigenfaces", *in Proc. IEEE Computer Society Conf. Computer Vision and Pattern Recognition*, Maui, Hawaii, 1991.

[5] K. Etemad and R. Chellappa "Discriminant analysis for recognition of human face images", *J. Opt. Soc. Am*. A/Vol. 14, No. 8 pp. 1724-1733 - August 1997.

[6] R. Brunelli and T. Poggio," Face Recognition: Features versus Templates", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 15, No. 10, Oct. 1993.

[7] Alan L. YuilIe, David S. Cohen and Peter W. Hallinan," Feature extraction from faces using deformable templates", *IEEE* 1989.

[8] H. A. Rowley, S. Baluja, and T.Kanade,"Neural Network-Based Face Detection", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 20, No. 1, Jan. 1998.

[9] Lin Hong and Anil Jain, "Integrating Faces and Fingerprints for Personal Identification", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 20, No. 12, Dec. 1998.

[10] K. C. Yow, R. Cipolla, "Feature-based human face detection", *Image and Vision Computing*, 15(1997) 713-735.

[11] B. Moghaddam and A. Pentland, "Probabilistic Matching for Face Recognition", *IEEE*, 1998.

[12] B. Moghaddam and A. Pentlend, "Probabilistic visual learning for object representation", *IEEE Transaction on Pattern Analysis And Machine Intelligence*, 19(7): 696-710, July 1997.

[13] M. J. Lyons, J. Budynek, and S. Akamatsu, "Automatic Classification of Single Facial Images", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 21, No. 12, Dec. 1999.

[14] K. K. Sung and T. Poggio, "Example-Based Learning for View-Based Human Face Detection", *IEEE Transaction on Pattern Analysis And Machine Intelligence*. Vol. 20, NO. 1, Jan. 1998.

[15] P. E. Pearson, "Developments in Model-Based Video Coding", *Proceeding of the IEEE*, Vol.83, No.6, June 1995.

[16] J. Cai, A. Goshtasby," Detecting human faces in color images", *Image and Vision Computing* 18(1999) 63-75.

[17] J. Zhang, Y. Yan and M. Lades, "Face Recognition: Eigenface, Elastic Matching, and Neural Nets", *Proceedings of the IEEE,* Vol. 85, No. 9, September 1997.

[18] Farzin Mokhtarian and Miroslaw Bober, *Curvature Scale Space: Theory, Applications, and MPEG-7 Standardization*, Kluwer Academic Publishers, Netherlands, 2003.

[19] MPEG-7 Face recognition technique, International Organization for Standarization, ISO/IEC JTC1/SC29/WG11, Coding of Moving Pictures and Audio, NPEG03/N6188, December 2003.