

A New Fast and Efficient HMM-Based Face Recognition System Using a 7-State HMM Along With SVD Coefficients

H. Miari-Naimi* and P. Davari*

Abstract: In this paper, a new Hidden Markov Model (HMM)-based face recognition system is proposed. As a novel point despite of five-state HMM used in pervious researches, we used 7-state HMM to cover more details. Indeed we add two new face regions, eyebrows and chin, to the model. As another novel point, we used a small number of quantized Singular Values Decomposition (SVD) coefficients as features describing blocks of face images. This makes the system very fast. The system has been evaluated on the Olivetti Research Laboratory (ORL) face database. In order to additional reduction in computational complexity and memory consumption the images are resized to 64×64 jpeg format. Before anything, an order-statistic filter is used as a preprocessing operation. Then a top-down sequence of overlapping sub-image blocks is considered. Using quantized SVD coefficients of these blocks, each face is considered as a numerical sequence that can be easily modeled by HMM. The system has been examined on 400 face images of the Olivetti Research Laboratory (ORL) face database. The experiments showed a recognition rate of 99%, using half of the images for training. The system has been evaluated on 64×64 jpeg resized YALE database too. This database contains 165 face images with 231×195 pgm format. Using five training image, we obtained 97.78% recognition rate where for six training images the recognition rate was 100%, a record in the literature. The proposed method is compared with the best researches in the literature. The results show that the proposed method is the fastest one, having approximately 100% recognition rate.

Keywords: Face Recognition, Hidden Markov Model, Singular Value Decompositions.

1 Introduction

Face recognition is the recognizing a special face from a set of different faces. Face has a significant role in human beings communications where, each person along with his/her feelings mainly is distinguished by his/her face image. One can easily find out that one of the main problems in machine-human being interactions is the face recognition problem. A human face is a complex object with features varying over time. So a robust face recognition system must operate under a variety of conditions. Face recognition has been undoubtedly one of the major topics in the image processing and pattern recognition in the last decade due to the new interests in, security, smart environments, video indexing and access control. Existing and future applications of face recognition are many [1]. We divide these applications into two main categories of governmental and commercial uses. Rapid progression

through customs by using face as a live passport in immigration, comparison of surveillance images against an image database of known terrorists and other unwanted people in security/counterterrorism, and Verifying identity of people found unconscious, dead or individuals refusing to identify themselves in hospital are examples of governmental uses.

Withdrawing cash from an automated teller machine (ATM) without cards or pin numbers in banking and access control of home and office in premises access control are some examples of commercial uses of face recognition systems which demonstrate the importance of these systems.

There have been a several face recognition methods, common face recognition methods are Geometrical Feature Matching [2], Eigenfaces method [3,4], Bunch Graph Matching [5], Neural Networks [6,7], Support Vector Machines [8], Elastic Matching [4] and Hidden Markov Models [9]-[15].

Here we briefly reviewed some of these approaches. The first approach, proposed by Kanade [2] in the 70's, were based on geometrical features. Geometrical Feature Matching techniques are based on the extraction

Iranian Journal of Electrical & Electronic Engineering, 2008.
Paper first received 23rd August 2007 and in revised form 2nd March 2008.

* The Authors are with the Department of Electrical and Computer Engineering, Babol University of Technology, Babol, Iran.
E-mail: h_miare@nit.ac.ir, P.Davari@gmail.com.

of a set of geometrical features forming the picture of a face. Their system achieved 75% recognition rate on a database of twenty persons, using two images per person; one for training and the other for test. In summary, geometrical feature matching techniques do not provide a high degree of accuracy and also are rather time-consuming. The other approach, one of the well known face recognition algorithms, is eigenfaces method [3,4]. This method uses the Principal Component Analysis (PCA) to project faces into a low-dimensional space. This method showed to be not very robust versus the variations of face orientation. In [3], the authors reported 90.5% correct recognition on ORL database. In summary, eigenfaces method is a fast, simple and practical method. However, it generally does not provide invariance over changes in scale and lighting conditions. Neural Networks is one of the approaches which have been used in many pattern recognition tasks. The attractiveness of using neural networks could be due to their ability in nonlinear mapping. The paper [7] used Probabilistic Decision-Based Neural Network (PDBNN) for face recognition which had been capable to recognize up to 200 people and could achieve up to 96% correct recognition rate. In general, neural network approaches encounter problems when the number of face classes increases. The last approach is the stochastic modeling of nonstationary vector time series based on Hidden Markov Models (HMM) which has been widely used in recent years. Samaria and Fallside [15] used this method for human face recognition. In [9] a top-down HMM using gray levels as features has been proposed. The recognition rate of this method was 87% on ORL database. A top-down HMM using DCT coefficients as features [12] and a Pseudo 2D HMM using gray levels as features [10] were reported to achieve 84% and 95% recognition rates respectively on ORL database. In recent years face recognition methods using HMM, like Pseudo 2D HMM with DCT coefficients features [13] and with Wavelet coefficients features [14], has been reported to have a recognition rate of 100% on ORL database.

This paper presents a new approach using one dimensional Discrete HMM as classifier and Singular Values Decomposition (SVD) coefficients as features for face recognition. We used a seven-state HMM to model face configuration. Here despite of previous papers, we added two new states to HMM to take into account the chin and eyebrows regions. The proposed approach has been examined on ORL database. To speed up the algorithm and reduce the computational complexity and memory consumption (in hardware implementation), along with using small number of SVD coefficients we resized the 112×92 pgm formatted images of this database to 64×64 jpeg images. Image resizing results in information losing and therefore seems to reduce the recognition rate. But we have gained 99% classification accuracy on ORL database while speeding up the system considerably. We also

have examined the proposed system on YALE face database which contains 165 images of fifteen persons, eleven images per person. We resized YALE database from 231×195 pgm format into 64×64 jpeg face images as well. No other changes like background cutting or cropping the images were preformed. Most of the previous papers have reported no specific number of training images on this database. We report our system results for 1 to 10 training images. Table 3 shows these results.

The rest of the paper is organized as follows. In Section 2, Hidden Markov Models and Singular Value Decomposition are briefly discussed. Section 3.1 describes an order-statistic filter and its role in the proposed system. Observation vectors calculations along with feature extraction process are discussed in Sections 3.2 and 3.3. Section 3.4 describes feature selection method. In Section 3.5 we introduce features quantization and labeling process. Sections 3.6 and 3.7 represent training and recognition procedures and discuss on results gained on ORL and YALE face databases. Thus Section three completely describes the proposed system. Finally in Section 4 conclusions are drawn.

2 Background

2.1 Hidden Markov Models

HMMs are usually used to model one dimensional data but in the recent years, they have been used in vision: texture segmentation [16], face finding [17], object recognition [18] and face recognition [19]. For a deep study on HMMs the reader may refer to [20]. Every HMM is associated with non-observable (hidden) states and an observable sequence generated by the hidden states individually. The elements of a HMM are as below:

- $N=|S|$ is the number of states in the model, where $S=\{s_1, s_2, \dots, s_N\}$ is the set of all possible states. The state of the model at time t is given by $q_t \in S$.
- $M=|V|$ is the number of the different observation symbols, where $V=\{v_1, v_2, \dots, v_M\}$ is the set of all possible observation symbols v_i (also called the code book of the model). The observation symbol at time t is given by $o_t \in V$.

As the reader is expected to know, observation vector is another concept that frequently is used in the HMM models. Indeed the training and test process of HMM models are performed in the observation vectors space. Each observation vector is a vector of observation symbols of length T . T is defined by user based on the in hand problem.

- $A=\{a_{ij}\}$ is the state transition probability matrix, where:

$$a_{ij} = P[q_{t+1} = s_j | q_t = s_i], \quad 1 \leq i, j \leq N, \quad 0 \leq a_{ij} \leq 1 \quad (1)$$

$$\sum_{j=1}^N a_{ij} = 1, \quad 1 \leq i \leq N \quad (2)$$

- $B = \{b_j(k)\}$ is the observation symbol probability matrix, where:

$$b_j(k) = P[o_t = v_k | q_t = s_j], \quad 1 \leq j \leq N, \quad 1 \leq k \leq M \quad (3)$$

- $\pi = \{\pi_1, \pi_2, \dots, \pi_N\}$ is the initial state distribution, where:

$$\pi_i = P[q_1 = s_i], \quad 1 \leq i \leq N \quad (4)$$

Using shorthand notation HMM is defined as following triple:

$$\lambda = (A, B, \pi) \quad (5)$$

As said before HMMs generally work on sequences of symbols called observation vectors, while an image usually is represented by a simple 2D matrix. In the case of using a one dimensional HMM in face recognition problems, the recognition process is based on a frontal face view where the facial regions like hair, forehead, eyes, nose and mouth come in a natural order from top to bottom. In this paper we divided image faces into seven regions which each is assigned to a state in a left to right one dimensional HMM. Figure 1 shows the mentioned seven face regions.

Figure 2 shows equivalent one-dimensional HMM model for a partitioned image into seven distinct regions like figure 1.

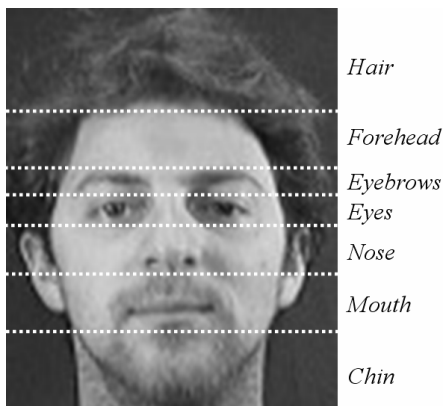


Fig. 1 Seven regions of face coming from top to down in natural order.

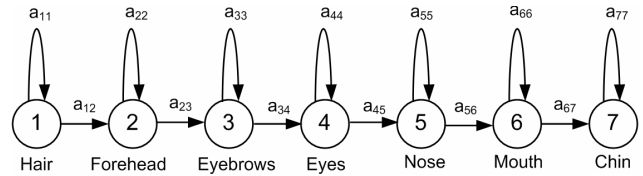


Fig. 2 A one dimensional HMM model with seven states for a face image with seven regions.

The main advantage of the model above is its simple structure and small number of parameters to adjust.

In previous researches, the eyebrows and chin regions weren't considered and the models have five states. Also the regions were featured using gray levels or some transforms like DCT and wavelets. We use SVD coefficients. The next section is a brief description of the SVD.

2.2 Singular Value Decomposition

The Singular Value Decomposition (SVD) has been an important tool in signal processing and statistical data analysis. Singular values of given data matrix contain information about the noise level, the energy, the rank of the matrix, etc. As singular vectors of a matrix are the span bases of the matrix, and orthonormal, they can exhibit some features of the patterns embedded in the signal. SVD provides a new way for extracting algebraic features from an image.

A singular value decomposition of a $m \times n$ matrix X is any function of the form:

$$X = U \Sigma V^T \quad (6)$$

where $U (m \times m)$ and $V (n \times n)$ are orthogonal matrix, and Σ is an $m \times n$ diagonal matrix of singular values with components $\sigma_{ii} = 0, i \neq j$ and $\sigma_{ii} > 0$. Furthermore, it can be shown that there exist non-unique matrices U and V such that $\sigma_1 \geq \sigma_2 \geq \dots \geq 0$. The columns of the orthogonal matrices U and V are called the left and right singular vectors respectively; an important property of U and V is that they are mutually orthogonal [21].

The main theoretical property of SVD relevant to face image recognition is its stability on face image. Singular values represent algebraic properties of an image [22]. So because of these reasons and some experimental results, we find out that SVD is a robust feature extraction technique for face images.

3 The Proposed System

3.1 Filtering

Most of the face recognition systems commonly use preprocessing to improve their performance. In the proposed system as the first step, we use a specific filter which directly affects the speed and recognition rate of the algorithm. Order-statistic filters are nonlinear spatial filters. Their operations are as follows; a sliding window moves from left to right and top to down with steps of

size one pixel, at each situation the centered pixel is replaced by one of pixels of the window based on the type of filter. For example minimum, maximum and median of pixels of the window may replace the centered pixel. A two dimensional order statistic filter, which replaces the centered element of a 3×3 window with the minimum element in the window, was used in the proposed system. It can simply be represented by the following equation.

$$\hat{f}(x,y)=\min_{(s,t)\in S_{xy}} \{g(s,t)\} \quad (7)$$

In this equation, $g(s,t)$ is the grey level of pixel (s,t) and S_{xy} is the mentioned window. Most of the face databases were captured with camera flash. Using the flash frequently caused highlights in the subjects eyes which affected the classification accuracy [23]. According to sentences above, this filter is expected compensate the flash effect (See figure 3). It also reduces salt noise as a result of the min operation.

Figure 4 shows a simple example demonstrating how minimum order-static filter works. Here a min filter using a 3×3 window operates on a 6×6 region of an image.

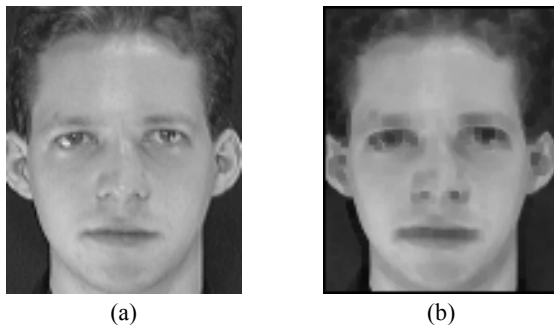


Fig. 3 An example of operation of the order static filter
a)Before filtering, b)After filtering.

190	191	188	187	184	180
193	194	189	188	185	179
194	194	189	188	185	179
188	188	187	184	183	178
179	180	181	178	178	176
182	175	176	181	178	179

a)Before filtering

0	0	0	0	0	0
0	188	187	184	179	0
0	187	184	183	178	0
0	179	178	178	176	0
0	175	175	176	176	0
0	0	0	0	0	0

b) After filtering

Fig. 4 An example of the operation of minimum order-static filter.

It can easily be found out that this filter has a smoothing role and reduces the information of the image

3.2 The Observation Vectors

Since HMMs require a one-dimensional observation sequence and face images are innately two-dimensional, the images should be interpreted as a one dimensional sequence.

The observation sequence is generated by dividing each face image of width W and height H into overlapping blocks of height L and width W . The technique is shown in figure 5. A $L \times W$ window is slid from top to bottom on the image and creates a sequence of overlapping blocks. These successive blocks are the mentioned interpretation.

The number of blocks extracted from each face image is given by:

$$T = \frac{H-L}{L-P} + 1 \quad (8)$$

where P is overlap size of two consecutive blocks. A high percent of overlap between consecutive blocks significantly increases the performance of the system consequently increases the computational complexity.

Our experiments showed that as long as P is large ($P \leq L-1$) and $L \approx H/10$, the recognition rate is not very sensitive to the variations of L .

3.3 Feature Extraction

As mentioned before in order to reduce the computational complexity and memory consumption, we resize ORL face database from 112×92 into 64×64 which results in data losing of images, so to achieve high recognition rate we have to use robust feature extraction method.

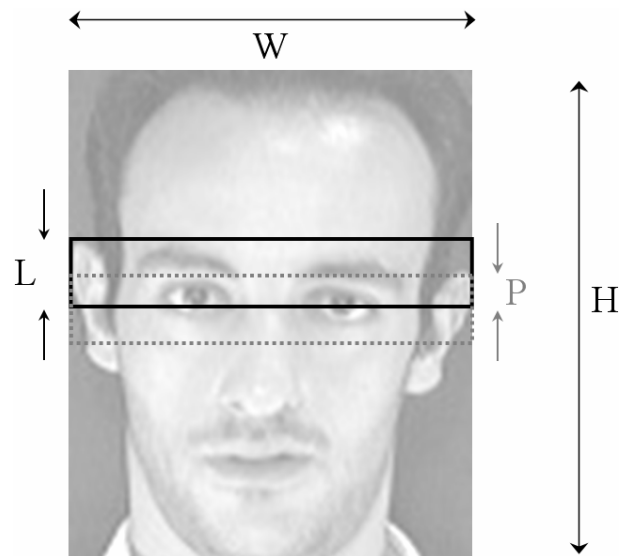


Fig. 5 The sequence of overlapping blocks.

A successful face recognition system depends heavily on the feature extraction method. One major improvement of our system is the use of SVD coefficients as features instead of gray values of the pixels in the sampling windows (blocks). We use a sampling window of five pixels height and 64 pixels width, and an overlap of 80% in vertical direction ($L=5, P=4$). As we mentioned in previous section the observation vector is composed of large number of $L \times W$ ($L=5$ and $W=64$) blocks, each containing 320 pixels. The number of blocks of an 64×64 image is 60. Here one can simply see that how long the observation vectors will be. Using pixels value as features describing blocks, increases the processing time and leads to high computational complexity in training and recognition procedure. In this paper, we compute SVD coefficients of each block and use them as our features.

3.4 Feature Selection

The problem of feature selection is defined as follows: given a set of d features, select a subset of size m that leads to the smallest classification error and smallest computational cost. We select our features from singular values which are the diagonal elements of Σ . An image with 64×64 resolution has a Σ matrix with same dimensions, which means 64 distinct singular values. It has been shown that the energy and information of a signal is mainly conveyed by a few big singular values and their related vectors. Figure 6 shows the singular values of a 64×64 face image. Obviously the first two singular values are very bigger than the other ones and consequently, based on the SVD theory, have more significance.

Figure 7 shows a face image along with its five approximations, using different combinations of the first three singular values and their related vectors. The last approximation (figure 7f) contains all three singular values together and simply shows the large amount of the face information.

To select some of these coefficients as feature, a large number of combinations of these were evaluated in the system described below.

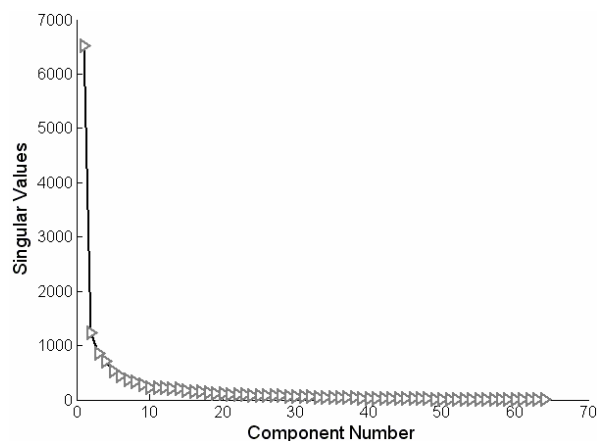


Fig. 6 SVD coefficients of a 64×64 face image.

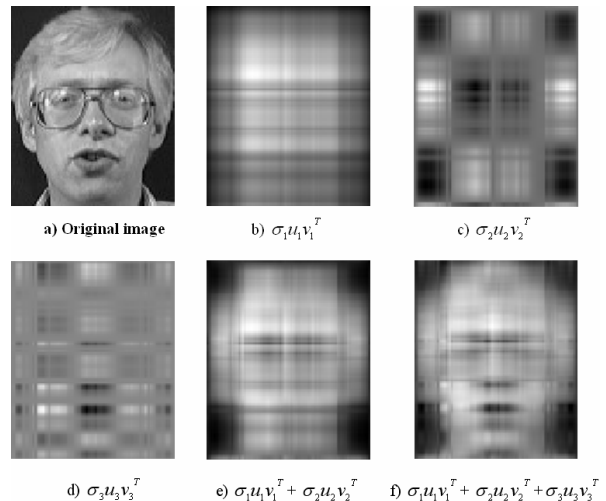


Fig. 7 a) Original image b, c, d, e and f are approximations of the original image.

As expected, the two biggest singular values along with U_{11} have the best classification rate. The classification rates of some of these coefficients are shown in figure 10. As these figures show, the three selected features have the highest recognition rates.

As mentioned before SVD contains three matrixes (U , Σ and V), and based on the above discussion we use two first coefficients of matrix Σ and first coefficient of matrix U as three features (Σ_{11} , Σ_{22} , U_{11}) associating each block. Thus each block of size $320 (=5 \times 64)$ pixels, is represented by 3 values. This significantly decreases the length of observation vectors. This decreases computational complexity and sensitivity to image noise, changes in illumination, shift or rotation.

3.5 Quantization

The SVD coefficients have innately continuous values. These coefficients build the observation vectors. If they are considered in the same continuous type, we will encounter an infinite number of possible observation vectors that can't be modeled by discrete HMM. So we quantize the features described above.

This is ordinarily done by a process of rounding, truncation, or some other irreversible, nonlinear process that all lead to information loss. To show the details of the quantization process, used in the proposed system, consider a vector $X=(x_1, x_2, \dots, x_n)$ with continuous components. Suppose x_i is to be quantized into D_i distinct levels. So the difference between two successive quantized values will be as equation x.

$$\Delta_i = (x_{i\max} - x_{i\min}) / D_i \quad (9)$$

$x_{i\max}$ and $x_{i\min}$ are the maximum and minimum values that x_i gets in all possible observation vectors respectively.

Knowing Δ_i , component x_i will be replaced with its quantized value computed as below:

$$X_{i\text{quantized}} = \left\lfloor \frac{X_i - X_{i\text{min}}}{\Delta_i} \right\rfloor \quad (10)$$

Thus all components of X will be quantized. At last each quantized vector is associated with a label that here is an integer number. So each block of image is mapped to an integer. Considering all blocks of an image, the image is mapped to a sequence of integer numbers that is considered as an observation vector. Thus the images are now so interpreted that we can simply use HMM for classification. In this paper we quantized the first feature (Σ_{11}) into 10, the second feature (Σ_{22}) 7 and the third one (U_{11}) into 18 levels, leaving 1260 possible distinct vectors for each block. In fact these quantization levels have been found based on experimental results. Choosing low values for D_i causes the number of these vectors increases and consequently the space of sequence describing each image be larger and vice versa.

3.6 The Training Process

After representing each face image by observation vectors, they are modeled by a seven -state HMM shown in figure 2. Five images of the same face are used to train the related HMM and the remaining five are used for testing.

A HMM is trained for each person in the database using the Baum-Welch algorithm [20]. At the first step $\lambda=(A,B,\pi)$ is initialized. The initial values for A and π are set due to the left to right structure of the face model corresponding to figure 2.

The initial values for A and π are as follows:

$$\begin{aligned} a_{i,i} &= a_{i,i+1} = 0.5 \quad 1 \leq i \leq 6 \\ a_{7,7} &= 1 \\ \pi_0 &= 1 \end{aligned} \quad (11)$$

Initial estimates of the observation probability matrix B are obtained as following simple Matlab statement:

$$B = \frac{1}{M} \text{Ones}(N,M) \quad (12)$$

where M is the number of all possible observation symbols obtained from quantization procedure and N is the number of states (in the proposed model N equals to 7). After representing the images by observation vectors, the parameters of the HMM are estimated using the Baum-Welch algorithm which finds $\lambda^* = \max_{\lambda} P(O|\lambda)$. In the computed model the probability of the

observation O associated to the learning image is maximized. Figure 8 shows the estimation process related to one learning image. This process is iterated for all training images of a person (class). The iterations stop, when variation of the probability of the observation vector (related to current learning image) in two consecutive iterations is smaller than a specified threshold or the number of iterations reaches to an upper bound. This process is reiterated for the remaining training images. Here the estimated parameters of each training image are used as initial parameters of next training image. The training scheme of one image is shown in figure 8. The estimated HMM of the last training image of a class (person) is considered as its final HMM.

3.7 Face Recognition

After learning process, each class (face) is associated to a HMM. For a K -class classification problem, we find K distinct HMM models. Each test image experiences the block extraction, feature extraction and quantization process as well. Indeed each test image like training images is represented by its own observation vector. Here for an incoming face image, we simply calculate the probability of the observation vector (current test image) given each HMM face model. A face image m is recognized as face d if:

$$P(O^{(m)}|\lambda_d) = \max_n P(O^{(m)}|\lambda_n) \quad (13)$$

The proposed recognition system was tested on the Olivetti Research Laboratory (ORL) face database. The database contains ten different face images per person of 40 people with the resolution of 112×92 pixels. As we mentioned before in order to decrease computational complexity (which affects on training and testing time) and memory consumption, we resized the pgm format images of this database from 112×92 to 64×64 jpeg format images.

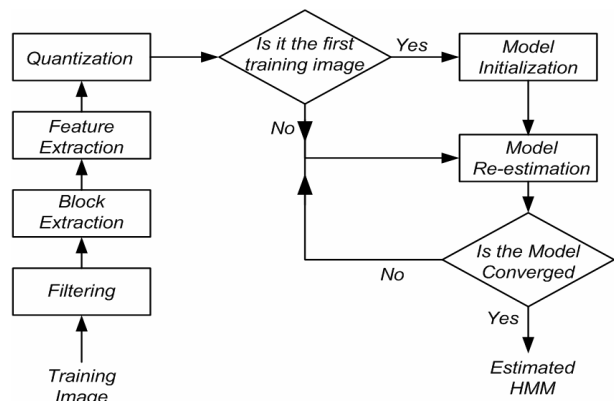


Fig. 8 The training process of a training image.

The five images of each person were used for the training task; the remaining five images were used for testing the system. The recognition rate is 99%, which corresponds to two misclassified images in whole database. Figure 9 shows ten example faces from ORL face database where the crossed images represent incorrect classifications. Table 1 represents a comparison among different face recognition techniques and the proposed system on the ORL face database. It is important to notice that all different face recognition techniques in Table 1 use 112×92 resolution of ORL face database where we use 64×64 image size.

As Table 1 shows the proposed system has a recognition rate of 99%, a high rate that can simply be approximated to 100%. The significance of this result is that such a high recognition rate is achieved using only three features per block along with image resizing.

As mentioned before we evaluated the classification ability of each feature individually. Indeed the classification process was repeated using each of these features individually. Figure 10 shows the obtained results for some important features. Here the recognition rate of the system was calculated for different values of the number of symbols. The number of symbol is simply changed by changing the number of quantization levels. Obviously the three features used in the system have the best classification rates. As mentioned before we quantized the first feature (Σ_{11}) into ten, the second feature (Σ_{22}) into seven and the third feature (U_{11}) into eighteen levels and obtained 99% recognition rate with 1260 symbols. To illustrate the relation between number of symbols and recognition rate we varied the number of symbols (changing the number of quantization levels of these three features together) from 8 to 1800. The recognition rate is illustrated in figure 11.

Increasing the number of symbols to achieve greater recognition rate leads to more time consumption for training and testing procedure. To prevent this event we can use low number of symbols. For example as we can see in figure 11 our system achieve 80%, 94.5% and 97% accuracy respectively with 24, 182 and 630 symbols. Again we notice here that because we resized our images database into low dimension (64×64), our system operates very fast with 1260 symbols and 99% recognition rate.

We also resized face images to 32×32 resolution and obtain 92.5% recognition rate. Table 2 shows the comparison results on 64×64 and 32×32 face images.

Our system was implemented in Matlab 7.1 and tested on a machine with CPU Pentium IV 2.8 GHz with 512 Mb Ram and 1 Mb system cache.

Table 3 shows a comparison of the different face recognition techniques on the ORL face database which reported their computational cost.

As we can see from Table 3 our system has a recognition rate of 99% and a low computational cost. Besides these advantages, the system has low memory space consumption because of resizing the face images.

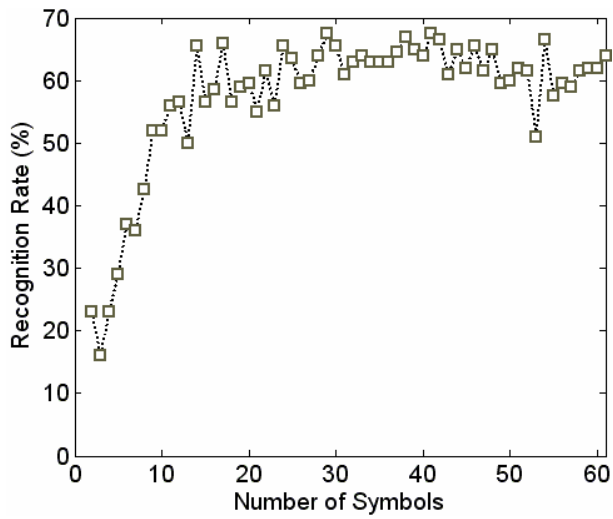


Fig. 9 Four examples of ORL face database. Crossed images represent incorrect classifications.

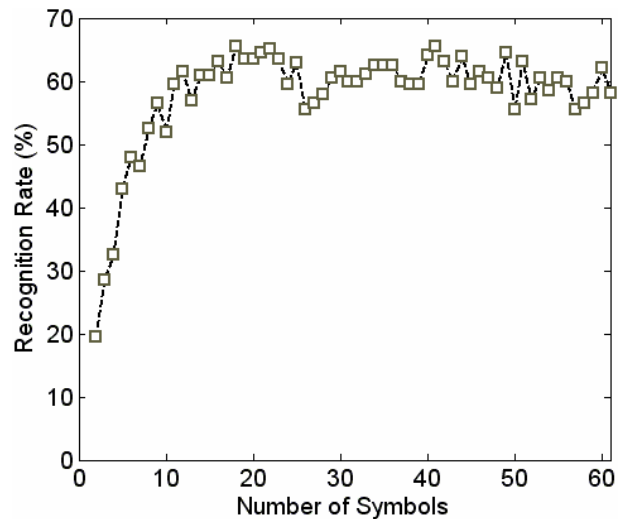
Table 1 Comparative results on ORL database. “1D HMM + SVD” represents the proposed method which tested on images size of ORL database where the other methods use images size.

Method	Error	Ref.	Year
Top-down HMM + gray tone features	13%	[9]	1994
Eigenface	9.5%	[3]	1994
Pseudo 2D HMM + gray tone features	5%	[10,24]	1994
Elastic matching	20.0%	[4]	1997
PDNN	4%	[7]	1997
Continuous n-tuple classifier	2.7%	[25]	1997
Top-down HMM + DCT coef.	16%	[11]	1998
Point-matching and correlation	16%	[26]	1998
Ergodic HMM + DCT coef.	0.5%	[12]	1998
Pseudo 2D HMM + DCT coef.	0%	[13]	1999
SVM + PCA coef.	3%	[8]	2001
Independent Component Analysis	15%	[27]	2002
Gabor filters + rank correlation	8.5%	[28]	2002
Pseudo 2D HMM + Wavelet	0%	[14]	2003
Markov Random Fields	13%	[29]	2004
1DHMM + Wavelet	0%	[30]	2003
1D HMM+SVD (The proposed system)	1%	This paper	2007

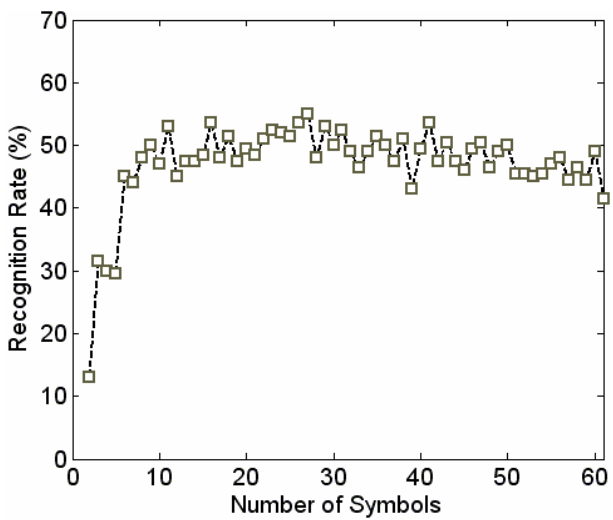
Finally we tested our system on YALE face database. The Yale face database contains 165 images of fifteen subjects. There are eleven images per subject with different facial expressions or lightings. Figure 12 shows the 11 images of one subject.



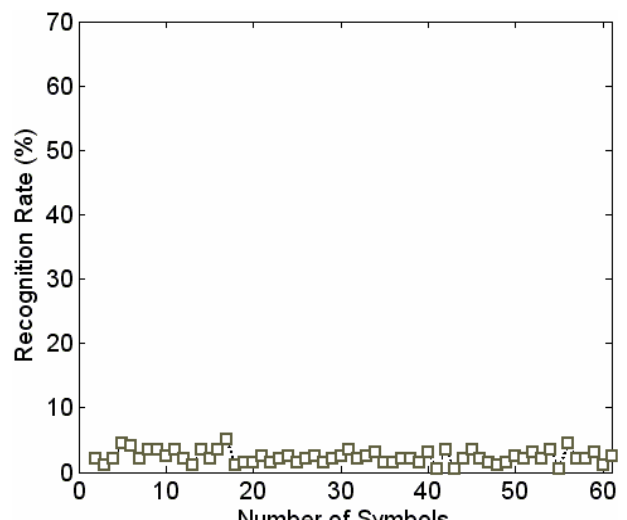
(a)



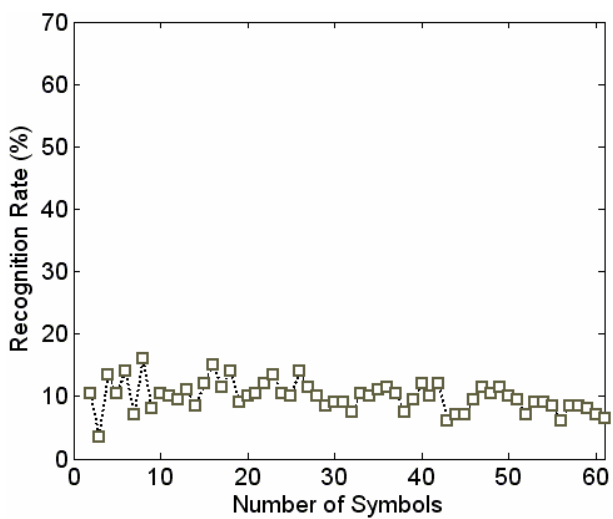
(b)



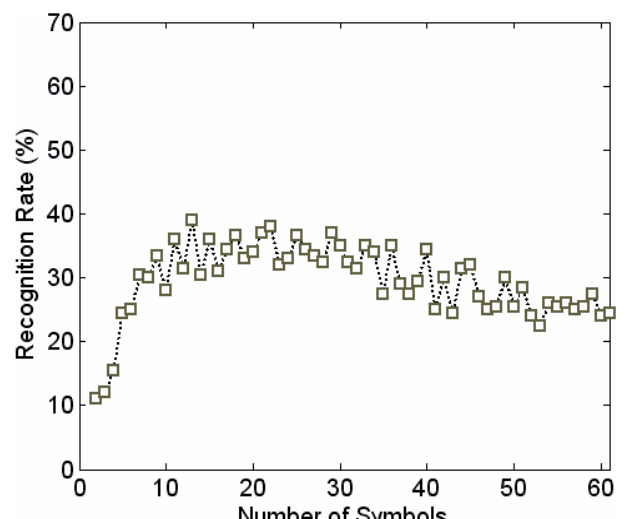
(c)



(d)



(e)



(f)

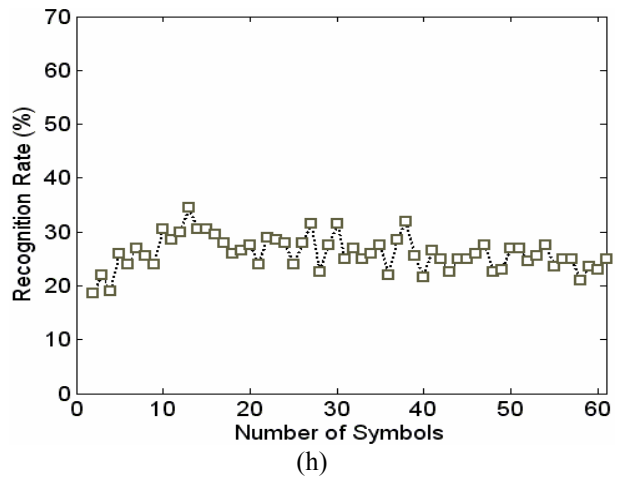
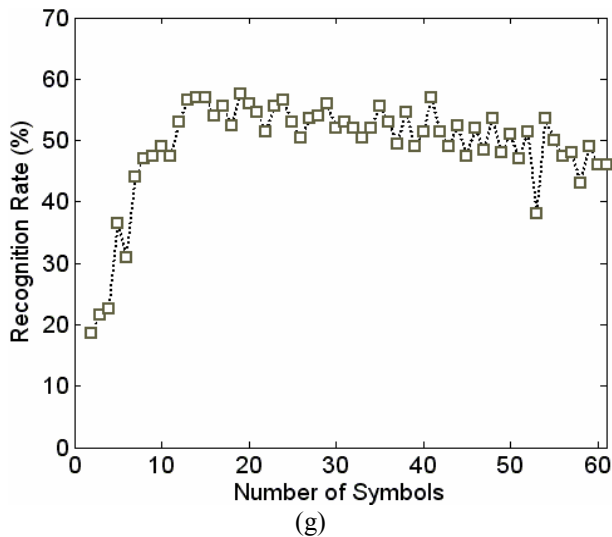


Fig. 10 a) $U(1,1)$ classification ability. Maximum value is 67.5% for 28 symbols b) $\Sigma(1,1)$ classification ability. Maximum value is 65.5% for 17 symbols c) $\Sigma(2, 2)$ classification ability. Maximum value is 55% for 26 symbols d) $V(1,1)$ classification ability. Maximum value is 5% for 16 symbols e) $V(1,2)$ classification ability. Maximum value is 16 for 7 symbols f) $V(2,2)$ classification ability. Maximum value is 39% for 12 symbols g) $U(2,1)$ classification ability. Maximum value is 57.5% for 18 symbols h) $U(2,2)$ classification ability. Maximum value is 34.5% for 12 symbols.

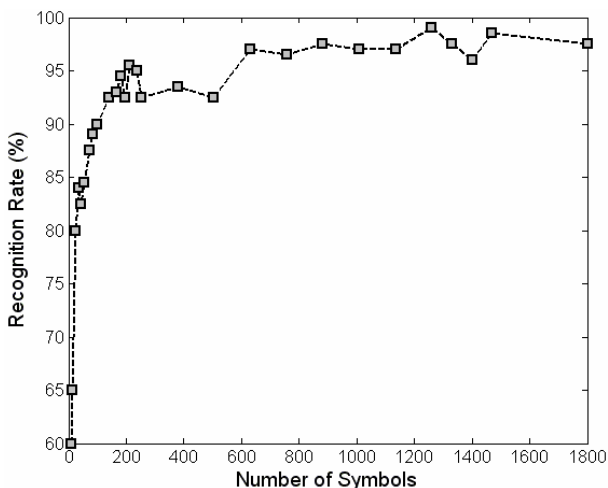


Fig. 11 Showing the relation between number of symbols and recognition rate. Maximum value is 99% and encounter on 1260 symbols.

Table 2. Comparing results of different image size of ORL database.

Image size	64×64	32×32
# of train image(s)	5	5
Training time per image(second)	0.63	0.46
Recognition time per image(second)	0.28	0.15
# of symbols	1260	960
Recognition (%)	99	92.5

Table 3. Comparative computational costs and recognition results of some of the other methods as reported by the respective authors on ORL face database.

Method	% Recognition	Training time per image	Recognition time per image
PDBNN [13]	96	20min	≤ 0.1 sec.
n-tuple [25]	86	0.9 sec.	0.025 sec.
1DHMM + Wavelet [30]	100	1.13 sec.	0.3 sec.
Pseudo-2D HMM [24]	95	n/a	240 sec.
DCT-HMM [12]	99.5	23.5 sec.	3.5 sec.
1D HMM + SVD (The proposed system)	99	0.63 sec.	0.28 sec.

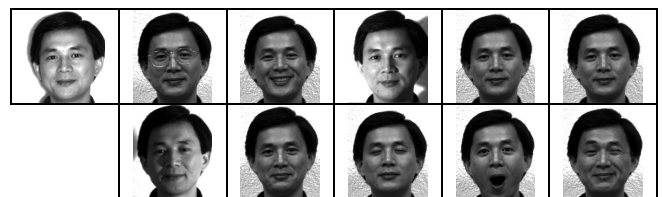


Fig. 12 A class of the YALE face database.

We resized YALE database from 231×195 into 64×64 jpeg face images. No other changes like background cutting or cropping the images were performed. In most

of the papers which tested their systems on this database no specific number of train images were reported, so we obtained our system results on 1 image to 10 images for training using 960 symbols, which means we quantized our first feature (Σ_{11}) into 8, second feature (Σ_{11}) 6 and our third feature (U_{11}) into 20 levels. Table 4 shows Comparative results on this database.

The proposed system has the best recognition rate when the number of training images is greater than five. Here, though the LDA has good performance, but for ORL database the proposed system had very higher performance than LDA.

Again we notice that all different face recognition techniques in Table 4 use 231×231 resolution of Yale face database or in [15] they cut out the face images background, where we use 64×64 image size, leaving a high speed system. Table 5 shows experimentation of the proposed system on the Yale database.

Table 4. Experiments on YALE face database
(Our accuracy obtained on 64×64 resolution face images).

# of train image(s)	MRF [29]	LDA [29]	PCA [29]	NN [29]	ID HMM + SVD (this paper)
1	81.6 %	N/A	60. %	68.31 %	78 %
2	93.11 %	91.23 %	75.2 %	80.49 %	82.22 %
3	95.2%	98.2 %	79. %	83.5%	90.83 %
4	95.9 %	99.41 %	79.8 %	84.0 %	94.29 %
5	96.11 %	99.69 %	81.1 %	83.51 %	97.78 %
6	96.67 %	99.87 %	81.1 %	82.8 %	100 %
7	98.67 %	99.97 %	81.9 %	82.63 %	100 %
8	97.33 %	100 %	81.2 %	81.6 %	100 %
9	97.33 %	100 %	81.7 %	82.2 %	100 %
10	99.33 %	100 %	81.7 %	83.07 %	100 %

Table 5. Result of experimentation of our proposed method on Yale database.

Image size	64×64	64×64
# of train image(s)	5	6
Train time per image(second)	0.54	0.53
Recognition time per image(second)	0.15	0.15
# of symbols	960	960
Recognition (%)	97.78	100

4 Conclusion and Future Works

A fast and efficient system was presented. Images of each face were converted to a sequence of blocks. Each block was featured by a few number of its SVD parameters. Each class has been associated to hidden Markov model as its classifier. The evaluations and comparisons were performed on the two well known face image data bases; ORL and YALE. In both data base, approximately having a recognition rate of 100%, the system was very fast. This was achieved by resizing the images to smaller size and using a small number of features describing blocks.

In the future, we will focus on the use of larger and more complicated databases to test the system. For these complicated databases it is simply expected that all the previous methods will not repeat such efficiency reported in the papers. We will try to improve the feature extraction and the modeling of the faces. The use of 2-D HMM or more complicated models may improve the system performance.

References

- [1] Nes A. and Bo. K. "Face Recognition," *Image Processing Specialization Project*, Norwegian University of Science and Technology, November 2002.
- [2] Kanade T. "Picture Processing by Computer Complex and Recognition of Human Faces," *Technical report*, Dept. Information Science, Kyoto Univ., 1973.
- [3] Turk M. and Pentland A., "Eigenfaces for Recognition," *J. Cognitive Neuroscience*, Vol. 3, No. 1, pp. 71-86, 1991.
- [4] Zhang J., Yan Y., and Lades M. Face recognition: eigenface, *Proceedings of the IEEE*, Vol. 85, No. 9, elastic matching and neural nets. September 1997.
- [5] Wiskott L., Fellous J.-M., Krüger N., and Vonderr malsburg C.. "Face Recognition by Elastic Bunch Graph Matching" *IEEE Transaction on Pattern Analysis and Machine Intelligence*, Vol 19, No 7, pp. 775-779, July 1997.

- [6] Lawrence S., Giles C.L., Tsoi A. C., and Back A.D., "Face Recognition: A Convolutional Neural-Network Approach," *IEEE Trans. Neural Networks*, Vol. 8, pp. 98-113, 1997.
- [7] Lin S., Kung S., and Lin L.. "Face Recognition/Detection by Probabilistic Decision-Based Neural Network" *IEEE Trans. Neural Networks*, Vol 8, No 1 pp. 114–131, January 1997.
- [8] Guo G., Li S. Z., , and Kapluk C.. "Face recognition by support vector machines," *Image and Vision Computing*, Vol. 19, No. 9-10, pp. 631–638, 2001.
- [9] Samaria F. and Harter A.. "Parameterization of a Stochastic Model For Human Face Identification," *In Proceedings of IEEE Workshop on Applications of Computer Vision*, Sarasota, Florida, December 1994.
- [10] Samaria F.. "Face recognition using Hidden Markov Models," *PhD thesis*, Engineering Department, Cambridge University, October 1994.
- [11] Nefian A. V. and Hayes M. H.. "Hidden Markov Models for Face Recognition," *In Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 2721–2724, Seattle, May 1998.
- [12] Kohir V. V. and Desai U. B., "Face recognition using DCTHMM approach," *In Workshop on Advances in Facial Image Analysis and Recognition Technology (AFIART)*, Freiburg, Germany, June 1998.
- [13] Eickeler S., Mller S., and Rigoll G., "Recognition of jpeg compressed face images based on statistical methods," *Image and Vision Computing*, Vol. 18, No. 3, pp. 279–287, 2000.
- [14] Bicego M., Castellani U., and Murino V., "Using Hidden Markov Models and Wavelets for face recognition," *In Proceedings IEEE International Conference on Image Analysis and Processing (ICIAP)*, 0-7698-1948-2, 2003.
- [15] Samaria F. and Fallside F., "Face Identification and Feature Extraction Using Hidden Markov Models," *Image Processing: Theory and Application*, G. Vernazza, ed., Elsevier, 1993.
- [16] Yang He and Amalan Kundu, "Unsupervised Texture Segmentation Using Multi-channel Decomposition and Hidden Markov Models," *IEEE Image Processing*, Vol. 4, No. 5, pp. 603-619, May 1995.
- [17] Rajagopalan A.N., Sunil Kumar K., Jayashree Kar-lekar, Manivasakan, Milind Patil, U.B. Desai, P.G. Poonacha, S. Choudhuri, "Finding Faces in Photographs," *ICCV98*, pp. 640-645, Jan. 1998.
- [18] Joachim Hornegger, Heinrich Niemann, Pulus D. and Schlottke G., "Object Recognition using hidden Markov Models," *Proc. of Pattern Recognition in Practice*, pp. 37-44, 1994.
- [19] Samaria F.S. and Young S., "HMM-based Architecture for Face Identification," *Image and Vision Computing*, Vol. 12, No. 8, pp. 537-543, Oct. 1994.
- [20] Rabiner L. R., "A Tutorial on Hidden Markov Modeles and Selected Application in Speech Recognition," *Proceeding Of the IEEE*, Vol. 77, No. 2, pp. 257-258, 1989.
- [21] Hong Z., "Algebraic feature extraction of image for recognition," *Pattern Recognition*, Vol. 24, No. 4, pp. 211-219, 1991.
- [22] Klema V. C. and Laub A. J., "The Singular Value Decomposition: Its Computation and Some Applications," *IEEE Transactions on Automatic Control*, Vol. 25, No. 2, pp. 0018–9286, April 1980.
- [23] Haralick, R. M. and Shapiro, L. G., *Computer and Robot Vision*, Volume I. Addison-Wesley, 1992.
- [24] Samaria F. S. and Young S., "HMM-based Architecture for Face Identification," *Image and Vision Computing*, Vol. 12, No. 8, pp. 537-543, Oct. 1994.
- [25] Lucas S. M., "Face recognition with the continuous n-tuple classifier," *In Proceedings of British Machine Vision Conference*, September 1997.
- [26] Lam K. M. and Yan H., "An analytic-to-holistic approach for face recognition on a single frontal view," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 20, No. 7, pp. 673–686, 1998.
- [27] Yuen P. C. and Lai J. H., "Face representation using independent component analysis," *Pattern Recognition*, Vol. 35, No. 6, pp. 1247–1257, 2002.
- [28] Ayinde O. and Yang Y., "Face recognition approach based on rank correlation of gabor-filtered images," *Pattern Recognition*, Vol. 35, No. 6, pp. 1275–1289, 2002.
- [29] Huang R., Pavlovic V. and Metaxas D. N., "A Hybrid Face Recognition Method using Markov Random Fields," *IEEE, 0-7695-2128-2*, 2004.
- [30] Le H. S. and Li H., "Simple 1D Discrete Hidden Markov Models for Face Recognition," *Springer Link Computer Science*, Vol. 2849, pp. 41-49, September 2003.



Hossein Miar-Naimi received the B.Sc. from Sharif University Of Technology in 1994 and M.Sc. from Tarbiat Modares University in 1996 and Ph.D. from Iran University of Science and Technology in 2002 respectively. Since 2003 He has been member of Electrical and Electronics Engineering Faculty of Babol University of Technology.

His research interests are analog CMOS design, Image processing and Evolutional Algorithms.



Pooya Davari received the B.Sc. degree in electronic engineering from Babol University of Technology, Babol, Iran, in 2004. He is currently an M.Sc. student in electronic engineering in Babol University of Technology, Babol, Iran. His research interests include image processing, active noise control systems, intelligent signal

processing and machine vision.