The CSI Journal on Computer Science & Engineering

The CSI Journal on
Computer Science and Engineering
Vol. 1, No. 2(b), Summer 2003
Pages 15-25
**Regular Paper**

# Face Verification Using Improved One-dimensional Hidden Markov Model

Shaghayegh Naderi* Mohammad Shahram Moin* Nasrollah Moghaddam Charkari**

*Biometrics Research Laboratory, Information Society Group, Iran Telecommunications Research Center
Tehran, Iran
**Electrical Engineering Department, Tarbiat Modarres University
Tehran, Iran

## Abstract

In this paper, we propose an improved version of 1D-HMM for face verification. DCT coefficients of face images are used as observation vectors in HMM states. Three types of modifications have been proposed to improve the overall performance of the classical model: (1) Replacing Baum-Welch algorithm with K-means clustering algorithm, (2) Replacing K-means with adaptive K-means and (3) Adaptive selection of training images amongst the available images in data set. The results show identical computational complexity in verification phase and better verification performance compared with other 1D-HMM methods. The proposed algorithm has been successfully tested on the ORL face image data set, exhibiting an accuracy of 96%. This is almost 10% higher than the identification rate of the classical 1D-HMMs and is comparable with 2D-HMMs accuracy; which basically has much higher processing complexity than 1D-HMM.

**Keywords**: Face verification, 1D Hidden Markov Model, 2D Hidden Markov Model, K-means, Discrete Cosine Transform, ORL database

## 1. Introduction

Face recognition and face verification techniques are categorized into three principal classes: geometric feature methods, template based methods and model based methods. Geometric methods use the angles and distances of the facial components for recognition. Template based methods such as PCA[1] [1,2], LDA[2] [1], Neural networks [3] and DLA[3] [4] use the template matching approaches. A simplified implementation of dynamic link architecture, the so-called EGM[4], is often preferred for locating objects in a scene with a known reference [5]. It has been shown that the template based approaches generally perform better than geometric feature methods [6] and work well for classifying frontal views of faces [7]. However, both methods are very sensitive to the variations in the size of images. This problem can be solved by using templates with different sizes; however this solution increases the computational complexity of the problem. Unlike the template based methods, the model based methods, such as HMM[5] [8] and SVM[6] [9][10], allow for greater flexibility and less sensitivity with respect to different lighting conditions, facial expressions and orientations such as hair style, facial hair and eye wear (glasses / no glasses). This flexibility is a result of using a mathematical model to incorporate information from different instances of faces at different scales and orientations. In model based methods, rather than comparing feature vectors that represent face templates to determine the identity of a face, the face model parameters are used for recognition [11]. The advantage of the HMM-based approach is its ability to handle variations in scale [11], which is a challenging problem for any face detection/ recognition /verification technique, and its computational efficiency; which is comparable to or lower than other approaches. This robustness is due to the statistical nature of the model.

HMMs are a set of statistical models used to characterize the statistical properties of a signal. These models have been widely used for speech and speaker recognition, where data is naturally one-dimensional along the time axis. HMMs are spreading to a variety of applications including digital signal processing and pattern recognition. 1D and 2D-HMM have also been applied for face recognition. Samaria introduced the use of the 1D continuous HMM for face recognition, yielding an identification rate of 84% on ORL[7] face database [12,13]. The highest identification rate of 2D-HMM tested on ORL face database for face recognition is reported in [14,15]. The 1D-HMM has been extended to a pseudo 2D-HMM by Samaria [8]. The results show a recognition rate of more than 95%, but also an important increase of the computational complexity. The equivalent full-connection 2D-HMMs presented by Nefian [16], Kohir [17,18] and Eickeler [19] reach an identification rate of 98-100%, but lead to very high computational complexities, which makes it inappropriate for real-time applications.

We have enhanced the performance of the 1D-HMM for face verification by increasing its accuracy almost up to the identification rate of the 2D-HMM. Moreover, the computational complexity of our proposed model is as low as the computational complexity of the classical 1D-HMM. The high accuracy and the low complexity of the proposed model make it appropriate for real-time applications.

The remainder of this paper is organized as follows: section 2 is dedicated to a general overview of the HMM-based approach in face verification/recognition, section 3 presents the proposed technique, section 4 contains the experimental results and section 5 concludes the paper.

## 2. Human Face Modeling in Classical HMM

The significant facial regions (hair, forehead, eyes, nose, mouth) come in a natural order from top to bottom for frontal face images, even if the images are taken under small rotations and/or rotations in a plane perpendicular to the image plane [14]. A state is assigned to each of these regions in a left to right continuous 1D-HMM. The state structure of the face model and the transition probabilities $a_{ij}$ are shown in Figure 1.

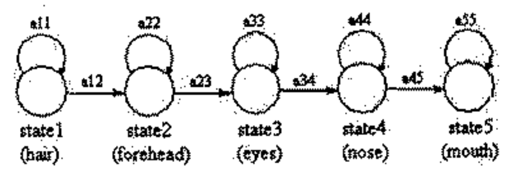



Figure 1. Face modeling using left to right 1D-HMM

The basic elements of a HMM are the number of states (N), state transition probability matrix (A), the state probability matrix (B) and initial distribution ($\pi$). Using a shorthand notation, a HMM is defined as a triplet: $\lambda = \{A, B, \pi\}$.

In face detection applications, a set of face images are used to train a universal HMM for face class, which includes all faces. For face recognition/ verification, each subject in the database is represented by one HMM face model and face images of the same subject are used to train its independent model.

In block extraction phase, the observations sequence is generated using the technique shown in Figure 2, where a window scans an image from top to bottom, to generate blocks with some overlap. After extracting blocks from each image in the training set, the observation vectors, which contain DCT coefficients [11] or pixels gray scales [12], are used to train each of the HMMs.

During the training phase (Figure 3), the training data is first segmented uniformly in N sections[8] from top to bottom, and then, the HMM parameters are initialized. At the next step, the uniform segmentation is replaced by the Viterbi segmentation [20,21] (Figure 4). Once the Viterbi segmentation likelihood becomes smaller than a predefined threshold, iterations stop. The best sequence of states in Viterbi algorithm is computed as shown in Figure 4.

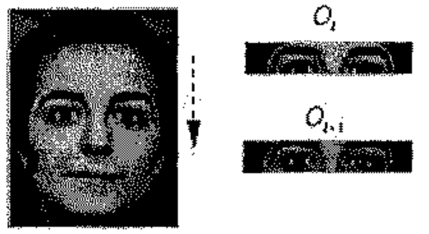



Figure 2. Image scanning in 1D-HMM

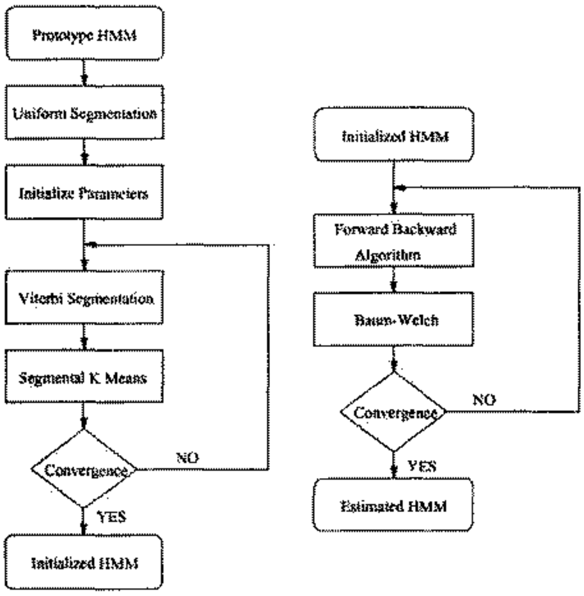



Figure 3. Training phase in classical 1D-HMM [11]

The final parameters of the HMM are obtained using the Baum-Welch recursive procedure [21,23, 24], where the model parameters are re-estimated.

In a face recognition system, during the recognition phase, the probability of the observation sequence for each HMM face model is calculated. Then, a model with the highest likelihood is selected, revealing the identity of the unknown face.

During the face verification phase in face verification applications, the probability of the observation sequence is calculated for the HMM model belonging to the claimed identity, using the Viterbi algorithm. If this probability exceeds a threshold, which is determined during the training phase, the claimed identity is accepted. This threshold, which depends on the required degree of security of environment, is calculated during the training phase.

1. Initialization:

$$\delta_1(i) = \pi_i \cdot b_i(o_1) \qquad 1 \le i \le N \qquad (1)$$

$$\psi_1(i) = 0$$

where $\delta_t(i)$ indicates the probability of occurring observation sequence at time $t$ for state $i$, $N$ is the number of states, $\pi_i$ is the probability of being in state $i$ in time $t$ and $b_i(o_t)$ is the probability of $t$ observation $o_t$ in state $i$.

2. Recursion:

$for \quad 2 \le t \le T \quad , \quad 1 \le j \le N$

$$\delta_t(j) = \max_i[\delta_{t-1} \cdot a_{ij}] \cdot b_j(o_t) \qquad (2)$$

$$\psi_t(j) = \arg\max_i[\delta_{t-1}(i) \cdot a_{ij}] \qquad (3)$$

where $\psi_t(j)$ indicates the previous state's number for the best sequence in time $t$ and state $j$, $T$ is the number of vertical observations and $a_{ij}$ indicates the probability of transitioning from state $i$ to state $j$.

3. Termination:

$$p^* = \max_i[\delta_T(i)] \qquad (4)$$

$$i^*(t) = \arg\max_t[p^*] \qquad (5)$$

where $p^*$ indicates the best previous state with the largest probability for observation s equence.

4. Reconstruction:

$for \quad t = T-1,\ T-2,\ \ldots,\ 1.$

$$i^*(t) = \psi_{t+1}\left(i^*{}_{(t+1)}\right) \qquad (6)$$

The resulting trajectory, $i^*(1), i^*(2), \ldots, i^*(T)$ is the most likely sequence of states for model and observation.

Figure 4. Viterbi algorithm in 1D-HMM

Three categories of HMMs have been used in face recognition/ verification as shown in Figure 5:

1) *1D-HMM:* this ordering suggests the use of a top-down model, where only the transactions between adjacent states is allowed. The recognition rate obtained on ORL dataset is 85% [14, 15].

2) *2D-HMM:* in this model, images are scanned in two directions (horizontal and vertical) [26,28]. The number of states is N*M, where N and M are the number of vertical and horizontal states, respectively. The recognition rate obtained on ORL dataset is about 100%. However, the main problem of 2D-HMM is its deficiency in real time processing, caused by its high computational complexity.

3) *Pseudo 2D-HMM (p2D-HMM):* this is a mixture of 1D-HMM and 2D-HMM. In this model, 1D-HMM is generalized, giving it the appearance of a 2D structure by allowing each state in a 1D-HMM to be a 1D-HMM itself. The model consists of a set of super states with a set of embedded HMMs. Images are scanned in a zigzag form. The p2D-HMM was first proposed by Samaria in [8]. Depending the structure of the model and the image sampling rate, the recognition rate of this model varied between 90% to 95%, on the ORL dataset. Then, Nefian in [15,16] and kohir in [18] increased the recognition rate of 2D-HMM, on the ORL dataset, up to 98% and 100%, respectively. The complexity and the accuracy of the p2D-HMM are both higher than those of 1D-HMM and are close to the 2D-HMM's complexity and accuracy [15]. A comparison between three categories of HMMs in term of complexity is shown in Table 1.

Table 2 shows a brief survey of HMM models used in face recognition and verification systems.

Table 1. Complexities of HMMs for face recognition

| Testing Complexity | Complexity / Model |
|---|---|
| $N_0{}^2 T_0$ | **Classical 1D-HMM** |
| $\left(\sum_{k=1}^{N_0} N_1{}^{(k)}\right)^2 T_0 T_1$ | **2D-HMM Classical** |
| $\sum_{k=1}^{N_0} (N_1^{(k)})^2 T_1 T_2 + N_0^2 T_0$ | **Pseudo 2D-HMM** |

($N_0$=number of super states, $N_1^{(k)}$=number of states in the $k$'th super state , $T_0$=number of vertical observations, $T_1$=number of horizontal observations in 2D and p2D-HMM only)

As shown in Tables 1 and 2, 2D-HMM and p2D-HMM reached an accuracy of 100% on the ORL database, but their complexity is too high for real-time applications.

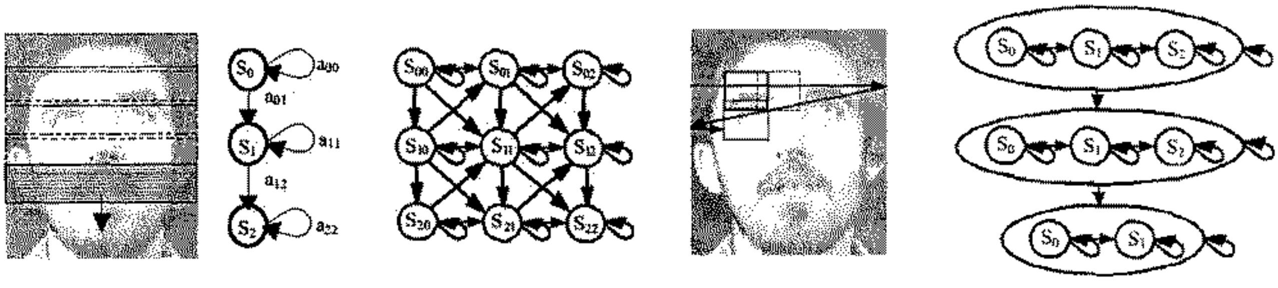
On the other hand, despite its low complexity, which makes it appropriate for fast operations in real-time face verification applications, the identification rate of the classical 1D-HMM is too low to be considered as a strong and reliable model for face verification. As a result, in order to implement an efficient face verification method for real-time systems, we have investigated on improving the 1D-HMMs performance to obtain both the low computational complexity of the 1D-HMM and the high identification rate of the 2D-HMM in a unique system. The detail of the proposed model is described in the following section.

# 3. Improved 1D-HMM

Our main objective is to find an efficient method for real-time face verification applications. Due to its low computational complexity, the 1D-HMM is a good choice for such a systems. However, as it's shown in Table 2, the highest recognition rate of 1D-HMM is 85%. In order to improve the recognition rate of the model, without increasing its computational complexity, we proposed the following modifications in the training algorithm of the 1D-HMM [30]:

1) Replacing the Baum-Welch algorithm with the K-means clustering algorithm;
2) Using adaptive clustering for each state;
3) Selecting the best training images for each subject.

In this section, we will describe the details of these modifications and in section 4, we will present the results



a) 1D-HMM b) 2D-HMM c) p2D-HMM

Figure 5. Three categories of HMM models used in face verification/recognition [25,26,27]

obtained in different experimentations, based on incremental modifications.

## 3.1 Replacing Baum-Welch algorithm with K-means clustering algorithm

Figure 6 shows the graphical representation of this modification in training algorithm with respect to the classical method, shown in Figure 3.

In the proposed method, the main part of the training phase consists of consecutive repetition of K-means[9] and Viterbi algorithms. Thus, the Baum-Welch algorithm is completely eliminated from the training phase. In this approach, unlike the classical method, the K-means algorithm is repeated for each state to estimate the pdf of state and to calculate the parameters of the model, i.e. covariance matrix, averages, A and $\pi$ matrices. The number of clusters is identical for all of the states[10]. The Viterbi algorithm determines the optimum path in each repetition of the whole algorithm. Training phase is continued until the predefined threshold value of the Viterbi probability is reached.

As it will be shown in Section 4, by this modification, we could reduce the error rate of the face verification procedure. It is also important to note that, despite the fact that K-means has been formerly used in speech recognition applications based on HMM, in the field of the face recognition using HMM, Baum-Welch algorithm has been replaced, for the first time in our approach, with K-means clustering algorithm.

## 3.2 Replacing K-means with adaptive K-means

In first improved version of 1D-HMM (section 3.1), the number of clusters in all of the states are taken identical. This might not be the best method to model different states. Indeed, different states of the model, which correspond to different regions of a face are best modeled with different number of clusters, depending on the amount of details in each region. Regions with more details need to be clustered with a higher number of clusters.

In order to determine the best number of clusters for each state, we propose to replace K-means with adaptive K-means. In our proposed method, to make adaptive clustering, the 'S' criterion is used to evaluate the performance of the clustering algorithm. It is defined as the ratio of the entire scattering of clusters to the minimum distance between centers of clusters, as follows:

$$S=\frac{\max\limits_{i=1}^{\max No.cluster} (\text{variance } x_i)}{\min\limits_{i,j\in\{1,\ldots,\max No.cluster\}} (dist(x_i,x_j))}=\frac{\text{scattering of clusters}}{\text{mindist between the cluster centers}} \quad (7)$$

where $x_i$ is the center of cluster $i$.

Figure 7 shows the modified training algorithm with 'S' criterion incorporated in algorithm to make adaptive clustering. The number of clusters varies between 2 and a maximum number, which is set to be 5 in our work. We used the hierarchical clustering for each state with "dendrogram[11]" [31].

We describe the difference between K-means and adaptive K-means clustering algorithms using the example shown in Figure 8. The data in Figure 8(a) and 8(b) represent both the same distributions, i.e. five different clusters generated using a Gaussian random generator. Figure 8(c) and 8(d) illustrate, in a similar way, three different Gaussian distributions. Samples in Figure 8(a) and 8(c) are clustered using K-means algorithm, with a constant number of clusters (4 in this example), while the data in Figure 8(b) and 8(d) are clustered using adaptive K-means algorithm, with a variable number of clusters. Clearly, in Figure 8(a) and 8(c), samples are clustered into a wrong number of clusters, but adaptive K-means could determine the correct number of clusters for both cases in Figure 8(b) and 8(d), i.e. 5 and 3, respectively.

## 3.3 Adaptive selection of training images

Relatively low recognition rate of the classical 1D-HMM for face recognition is, essentially, caused by the inappropriate selection of the training images and, therefore, lack of important information in these images, needed for an adequate training.

Indeed, in classical algorithm [11][12][22], the first five images of each person, in ORL dataset, are used to train the model. Despite the fact that there are some degrees of randomness in these selected images, they do not necessarily contain uncorrelated information. In addition, there are some important information existing in other images, that are absent in first five images.

To overcome this problem, we propose to select the best training images automatically, amongst the available images

Table 2. A survey of HMMs used for face recognition/verification

| Othman, Aboulnasr [25,26,27] | Eickler, Muller Rigoll,Wallhoff [19] | Kohir , Desai [18,17] | Nefian , Monson , Hayes [14,15,16] | Samaria, Harter[8,3] | Research groups / Models specifications |
|---|---|---|---|---|---|
| Dataset of AT&T Labs., Cambridge | ORL | ORL | ORL | ORL | **Database** |
| Not mentioned | 112 ×92 | 112 ×92 | 112 × 92 | 112 × 92 | **Image Resolution** |
| DCT | DCT | DCT | DCT | Raw image | **Feature(s)** |
| 9 | Not mentioned | 10 | 1D: 39<br>p2D: 6 | 39 | **No. of DCT coefficients** |
| 8×8 | Between 11×14 & 43×53 | 16×16 | 1D : 10×92<br>p2D : 8×10 | Width: Between 1&10 | **Scanning Window Size** |
| 0% | 75% | 75% | 1D : 8 pixel | Between 0 & (width -1) | **Scanning Window Overlap (in %)** |
| 9 | 5 | 5 | 5 | 5 | **Number of Training Images** |
| 2D | p2D | p2D | 1D<br>p2D | 1D<br>p2D | **HMM model** |
| 7×3 | 8×8 | 5 | 5 | 5 | **Number of States** |
| Between 2 and 16 | Between 1 and 3 | Not mentioned | 1D : m>1<br>p2D : 0 | 0 | **Number of Mixtures** |
| Not mentioned | Not mentioned | K-means | Uniform | Uniform | **Initial Segmentation** |
| Modified Viterbi | Baum-Welch | Viterbi | Viterbi,<br>Baum-Welch | Viterbi,<br>Baum-Welch | **Training Algorithm** |
| 100% | 100% | 100% | 1D: 85%<br>p2D: 98% | 1D: 84%<br>p2D: 90-95% | **Recognition Rate** |

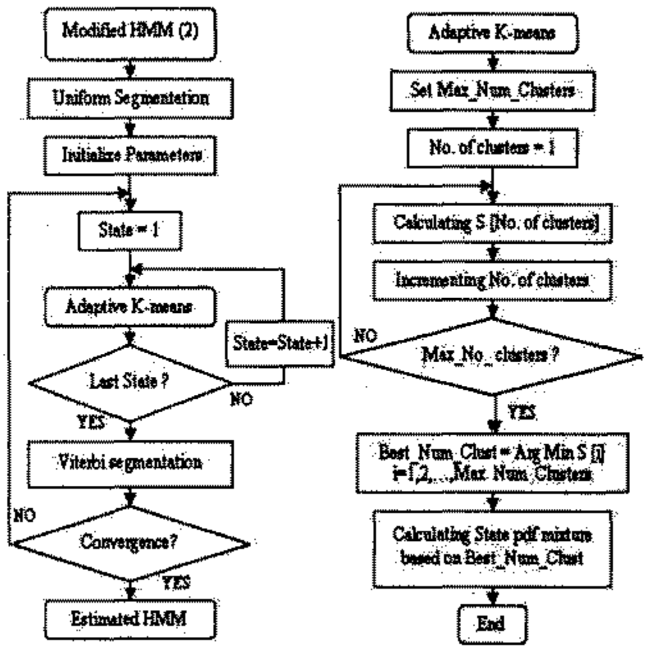



Figure 6. Replacing the Baum-Welch algorithm with K-means in training phase

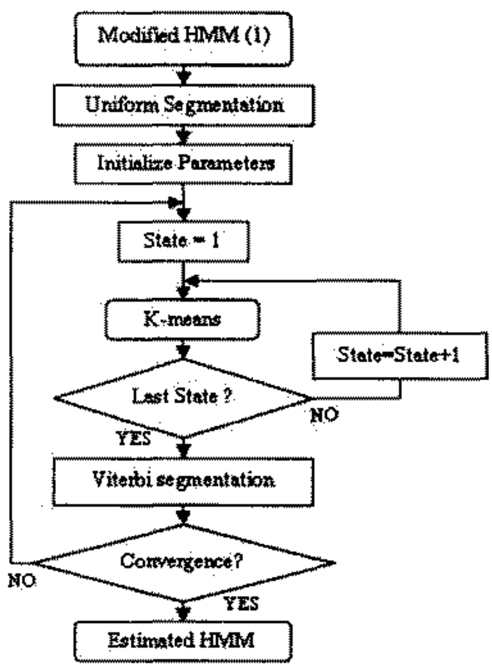



Figure 7. Using Adaptive K-means in training phase

for each subject. In the case of the standard ORL database, the problem consists of choosing five optimal training images out of the total ten images available for each subject. Obviously, the best training set is one that contains the maximum amount of information, i.e. images with different aspects of a subject's face, taken in different conditions. In order to construct such a training set, we first calculate DCT coefficients of the images[12], and then select images with the most distinguishable coefficients for HMM model training. Thus, one can extract the most important information available in subject images, from the obtained training data set [32].

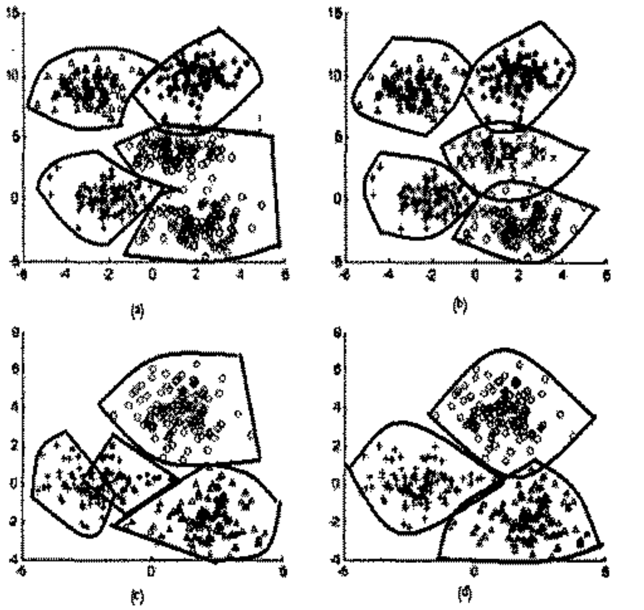


Figure 8. Results of K-means ((a) and (c)) and adaptive K-means, ((b) and (d)) clustering algorithms.

The algorithm used to determine the best training images is as follows:

1. Choose one training image randomly;

2. For all remaining images, calculate the minimum Euclidean distance between each image and the selected training image(s),

$$D_{i,j}^2 = \sum_{n=1}^{m} (d_i(n) - d_j(n))^2 \qquad (8)$$

where $m$ is the length of image vector, i.e. 'No. of rows' × 'No. of columns'.

3. Select the image with the largest minimum distance as the next training image:

$$\textit{Index of the next training image} = \text{Arg Max}_i\,(\text{Min}\,(D_{i,j})) \qquad (9)$$

4. If there are still training images to choose, go to step 2.

5. End.

By applying this algorithm to the image data set, one can expect that the last image contains information, which was not present in the previous face images (Figure 9). Also, one can see that the face images in Figure 9.b are the most different images out of 10 available images. The different steps for training image selection in Figure 9 are shown in Figure 10.

# 4. Experimental Results

The initial parameters used in our 1D-HMM face verification system are:

- Number of states: 5
- Face database: ORL
- Number of training images: 5
- Image dimension: 112×92
- Number of DCT coefficients in each block: 39
- Overlap between consecutive blocks during the feature extraction phase: 75%
- Initial segmentation: uniform

According to [14], [22] and [29], these values are the most appropriate values for 1D-HMM in face recognition.

The experiments have been conducted on the ORL database. The ORL database consists of a total of 400 gray scale images of a total of 40 people, 10 samples per person (Figure 11). This database supports the illumination, pose and expression changes between images of the same subject. All images of ORL have been used in our experiments: five images of each person in training phase and the remaining in test phase.

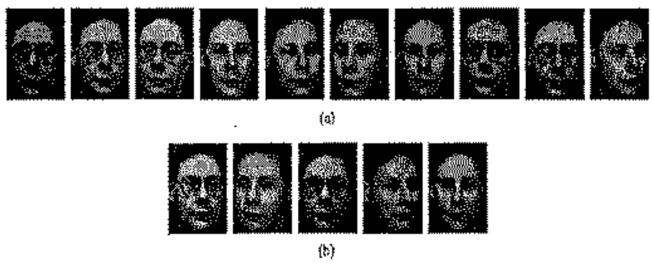


Figure 9. Images of a subject taken from ORL database. a) 10 available images for training and test phases. First five images are used for training according to the classical 1D-HMM. b) Training images selected using adaptive method

In order to compare our results with the results obtained from the classical HMMs, we first calculated a decision threshold to obtain a False Rejection Rate (FRR), similar to one obtained in the previous works, that is 15%, and then, with this threshold, we calculated the False Acceptance Rate (FAR) of different models. Table 3 shows the results of incremental improvements of the classical 1D-HMM.

We considered different techniques for selecting training images. In the classic HMM methods, the first five images of each subject in the ORL data set have been used for training the model and the next five images for evaluating the performance of the model. However, in order to make our results statistically significant, we selected the training images randomly and repeated our tests 10 times for each subject. Then, we calculated the average of FARs and FRRs for different decision thresholds.

| Selected image(s) |  10 |  9 |  8 |  7 |  6 |  5 |  4 |  3 |  2 |  1 | Input Image / Step |
|---|---|---|---|---|---|---|---|---|---|---|---|
|  2 | Random selection of the first image | | | | | | | | | | 1 |
|  2,6 | 5.61 | 4.16 | 5.99 | 5.03 | [6.69] | 4.60 | 4.43 | 6.34 | 0 | 5.90 | 2 |
|  2,6,1 | 5.61 | 4.16 | 5.14 | 5.03 | 0 | 4.60 | 4.43 | 4.95 | 0 | [5.90] | 3 |
|  2,6,1,10 | [5.61] | 4.16 | 4.36 | 5.03 | 0 | 4.60 | 4.43 | 4.58 | 0 | 0 | 4 |
|  2,6,1,10,7 | 0 | 4.16 | 4.36 | [5.03] | 0 | 3.95 | 4.43 | 4.58 | 0 | 0 | 5 |

Figure 10. An example of training image selection steps based on our method. The values in this table show the minimum distance between each image and previously selected images in each step

In the first stage of our experiments, the K-means clustering algorithm replaced the Baum-Welch algorithm. The number of clusters has been considered identical for all HMM states (4 clusters/state) and, through several stages of test, the identification rate of 87% were obtained, i.e. an improvement of 2% in FAR, compared to the classical 1D-HMM's identification rate.

Adding the S criterion to make adaptive clustering improved the identification rate up to 6%, i.e. an accuracy of 91%. Figure 12(a) shows the average FARs and FRRs obtained with different threshold values.

Finally, by adaptive selection of the training images, the identification rate increased to 96%, as shown in Figure 12(b), i.e. more than 10% decrease in FAR compared to the classical 1D-HMM, considering the fact that the computational complexity is not increased.

Table 4 shows a comparison between our method and the other HMM based methods for face verification. The processing order of our method is $O(N_0^2 T_0)$, which makes it an adequate choice for real time applications.

Table 5 shows the average training and test times for each subject, as well as the identification rates of the classical and the proposed models, obtained on a personal computer with Pentium IV processor (CPU speed = 1.6 GHz and RAM = 512 MB). As it is shown in this table, comparing with the classical model, the training time of the proposed model increased slightly (almost 2 sec), while the test time remained unchanged and the identification rate increased significantly.

Figure 13 shows some of the face images segmented in five regions using our approach. As one can see, the regions does not change (or have a small change) for different images of the same subject, even in different conditions. The crossed image represents incorrect classification, while the rest are examples of correct classification.

Table 3. Incremental improvements of the proposed method in comparison with the classical 1D-HMM

| **Results / Steps of Improvement** | **Verification Accuracy** |
|---|---|
| **Replacing Baum-Welch with K-means** | An increase of 2%* (up to 87%) |
| **Replacing K-means with adaptive K-means** | An increase of 6%* (up to 91%) |
| **Adaptive selection of training images** | An increase of 11%* (up to 96%) |

* In comparison with the identification rate of the classical 1D-H

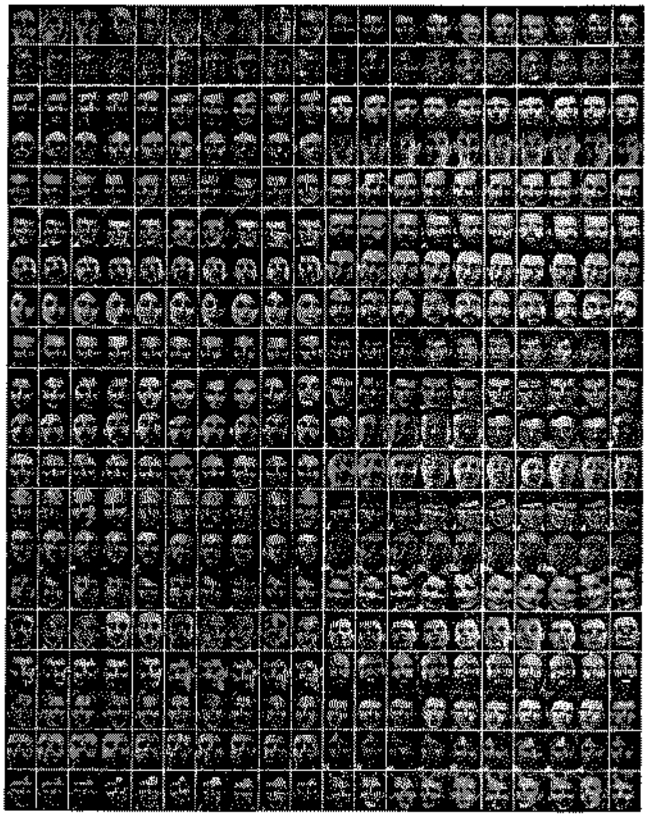


Figure 11. ORL database

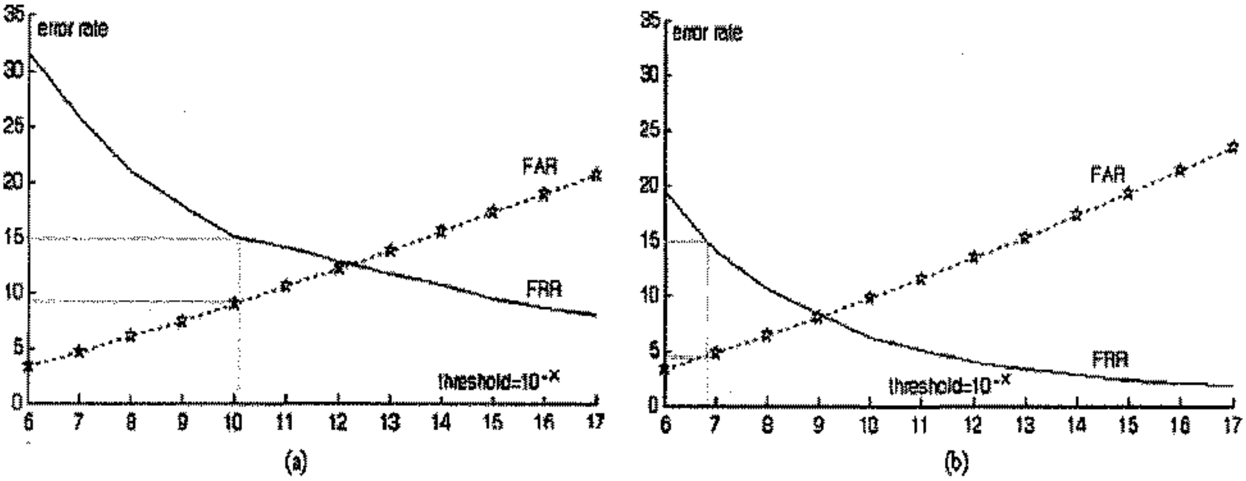



Figure 12. FARs and FRRs obtained with adaptive K-means and (a) random and (b) adaptive selection of training imag

Table 4. Identification rates and complexities of HMMs for face recognition and verification

| Identification rate (FRR=15%) | Test Phase Complexity | |
|---|---|---|
| 85% | $N_0{}^2 T_0$ | **Classical 1D-HMM [15]** |
| 100% | $\sum_{k=1}^{N_0} (N_1^{(k)})^2 T_1 T_2 + N_0^2 T_0$ | **Pseudo 2D-HMM [15]** |
| 96% | $N_0{}^2 T_0$ | **Improved 1D-HMM** |

($N_0$=number of super states, $N_1^{(k)}$=number of states in the $k$'th super state ,
$T_0$=number of vertical observations, $T_1$=number of horizontal observations in 2D and p2D-HMM only)

Table 5. Average time and accuracy of classical and proposed models

| | Training time (off-line) | Test Time (on-line) | Identification Rate |
|---|---|---|---|
| Classical 1D-HMM | 6.7 Sec | 0.25 Sec | 85% |
| Improved 1D-HMM | 9.8 Sec | 0.25 Sec | 96% |

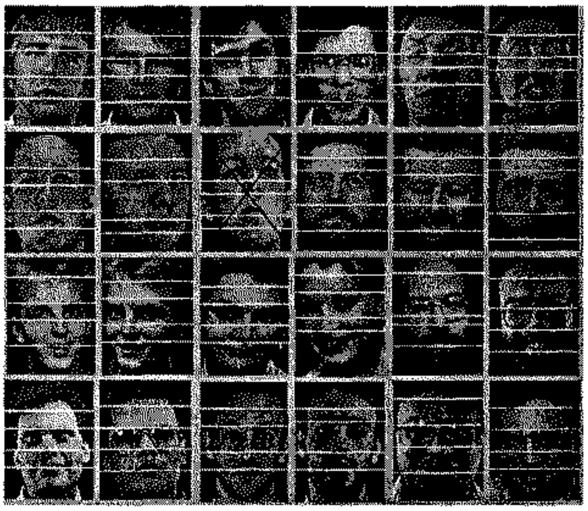


Figure 13. Face segmentation results, using our method

## 5. Conclusions

Classical 1D-HMM model used in face identification obtained an identification rate of 85%. By applying the proposed method in this paper, we improved the model accuracy significantly, reaching an identification rate of %96. 1D-HMM performance improvement were obtained by replacing the Baum-Welch algorithm with the K-means clustering algorithm, by adaptive selection of the number of clusters for each state and finally by finding the best training images for each subject.

The interesting point to note is that, the identification rate of the proposed model increased, without increasing the model complexity in verification phase (Table 4). In fact, only the training time increased slightly, mainly due to the adaptive selection of training images. Since the training phase is generally done in off-line mode, such a small increase in training time is acceptable. Consequently, we believe that the proposed model can be used successfully in accurate real-time face verification applications. The future work consists of the evaluation of the proposed model on a real-time application, using a large image dataset.

## References

[1] A. M. Martinez and A. C. Kak, "PCA versus LDA," *IEEE Trans. Pattern Anal. Mach. Intell.* Vol. 23, No. 2, pp. 228 –233, 2001.

[2] M. - H. Yang, "Face recognition using kernel methods," *Neural Information Processing Systems ( NIPS) , Vancouver*, 2001.

[3] M. Joo Er, S. Wu, J. Lu and H. L. Toh, "Face Recognition With Radial Basis Function (RBF) Neural Networks," *IEEE Transaction on Neural Networks*, Vol. 13, No. 3,pp. 697-710, may 2002.

[4] M. Lades, J. C. Vorbruggen, J. Buhmann, J. Lange, C. V. D. Malsburg, R. P. Wurtz, and W. Konen, "Distortion Invariant Object Recognition in the Dynamic Link Architecture," *IEEE Trans. on Computers*, Vol. 42, No. 3, pp. 300- 311, Mar. 1993.

[5] J. Zhang, Y. Yan, and M. Lades, "Face Recognition: Eigenface, Elastic Matching and Neural Nets," *Proceedings of the IEEE*, Vol. 85, No. 9, pp. 1423-1435, Sep. 1997.

[6] R. Brunelli and T. Poggio, "Face recognition: Features versus templates," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 15, No. 10, pp. 1042–1052, 1993.

[7] B. Heisele, P. Ho, J. Wu and T. Poggio, "Face recognition: component based versus global approaches," *Computer Vision and Image Understanding 91*, pp.6-21, 2003.

[8] F. Samaria, "Face Recognition Using Hidden Markov Models," *Ph.D. thesis*, University of Cambridge, 1994.

[9] B. Heisele, P. Ho, and T. Poggio, "Face recognition with support vector machines: global versus component-based approach," *In Proc. 8th International Conference on Computer Vision, Vancouver*, Vol. 2, pp. 688– 694, 2001.

[10] K. T. Jonsson, "Robust Correlation and Support Vector Machines for Face Identification," *PhD thesis, School of Electronic Engineering, Information Technology and Mathematics, University of Surrey, UK*, March 2000.

[11] A. Nefian , "A Hidden Markov Model-Based Approach for Face Detection and Recognition," *Ph.D. thesis, Georgia Institute of Technology*, August 1999.

[12] F. Samaria, "Face segmentation for identification using Hidden Markov Models," *in British Machine Vision Conference*, 1993.

[13] F. Samaria and S. Young, "HMM based architecture for face identification," *Image and Computer Vision*, Vol. 12, October 1994.

[14] A.V. Nefian and M. H. Hayes, "Hidden Markov models for face recognition," *IEEE International Conference on Acoustics, Speech and Signal Processing*, Vol. 5, pp. 2721 -2724, 1998.

[15] A. V. Nefian and M.H. Hayes, "An embedded HMM-based approach for face detection and recognition," *IEEE International Conference on Acoustics, Speech, and Signal Processing*, Vol. 6 , pp. 3553 -3556, 1994.

[16] A. V. Nefian and Monson H. Hayes III, "Face Recognition using An Embedded HMM," *IEEE Conference on Audio and Video-Based Biometric Person Authentication, Washington DC*, pp. 19-24, March 1999.

[17] V. V. Kohir, and U.B. Desai, "Face Recognition Using a DCT-HMM approach," *Fourth IEEE Workshop on Applications of Computer Vision, WACV 98*, pp. 226 – 231, 1998.

[18] V. V. Kohir, and U.B. Desai , "A Transform Domain Face recognition approach," *IEEE Region 10*

*Conference*, *TENCON 99*, Vol. 1 , pp. 104 –107, 1999.

[19] S. Eickeler, S. Muller and G. Rigoll, "Recognition of JPEG Compressed Face Images based on Statistical Methods," *Image and Vision Computing*, Vol. 18, No. 4, pp. 279-287, 2000.

[20] A. J. Viterbi. "Error bounds for convolutional codes and asymptotically optimum decoding algorithm," *IEEE Transactions on Information Theory*, Vol. 13, pp. 260-269, 1967.

[21] L. R. Rabiner, "A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition," *Proceeding of the IEEE*, Vol. 77, No. 2, pp. 257-286, February 1989.

[22] F. S. Samaria, and A.C. Harter, "Parameterisation of a stochastic model for human face identification, " *Second IEEE Workshop on Applications of Computer Vision*, pp. 138 –142, 1994.

[23] A. V. Nefian and M.H. Hayes, "Maximum Likelihood Training of the Embedded HMM for Face Detection and Recognition," *International Conference on Image Processing*, Vol. 1, pp. 33 –36, 2000.

[24] J. C. Stiller and G.Radons, "Online estimation of Hidden Markov Models," *IEEE Signal Processing letters*, Vol. 6, No. 8, pp. 213-215, Aug 1999.

[25] H. Othman and T. Aboulnasr, "A Tied Mixture 2-D HMM Face Recognition System," *International Conference on Pattern Recognition, ICPR 2002.*

[26] H.Othman, and T.Aboulnasr, , "A simplified second-order HMM with application to face recognition," *IEEE International Symposium on Circuits and Systems*

[27] V. V. Kohir, and U.B. Desai , "A transform domain face recognition approach," *IEEE Region 10 Conference* , *TENCON 99*, Vol. 1, pp. 104 –107, 1999.

[28] S. Naderi. "Face Verification with HMM," *M.Sc. Thesis, Electrical Engineering Department, Tarbiat Modarres University*, Tehran, Iran, February 2003.

[29] N. Moghaddam, M. S. Moin and S. Naderi, "Optimizing 1D-HMM for Face Verification," *Proceedings of the Second Iranian Conferene On Machine Vision, Image Processing (MVIP 2003)*,Tehran, Iran, pp. 127-133, February 2003.

[30] M. S. Moin, N. Moghaddam and S. Naderi, "1D-HMM for Face Verification: Optimizing the Model by Improving the Algorithm and Intelligent Selection of Training Set," *Proceedings of the Eighth Annual International CSI Computer Conference (CSICC'2003)*, Mashhad, Iran, pp. 365-369, February 2003.

[1] Principal Component Analysis

[2] Linear Discriminant Analysis

[3] Dynamic Link Approach

[4] Elastic Graph Matching

[5] Hidden Markov Model

[6] Support Vector Machine

[7] Available from http:// www. uk. research. att. com/ facedatabase. html

[8] N is fixed to 5 according to [22]

[9] Details of the k-means clustering algorithm is available in [21, 28]

[10] In our experimentations, the number of clusters is defined equal to 4, which corresponds to the best experimental results obtained with different number of clusters between 2 and 5.

[11] Dendrogram is a binary tree that shows the structure of the clusters and in addition provides the similarity measure between clusters (the vertical axis). The best number of clusters corresponds to the smallest value of S.

[12] DCT coefficients have been used for compatibility with classical approach presented in [11]. It has been shown that DCT coefficients are the best features for face verification/recognition using 1D-HMM.

**Shaghayegh Naderi** received the B.Sc. degree in computer engineering from Tarbiat Moallem University, Tehran, Iran, in 2000, and the M. Sc. degree in computer engineering from the Tarbiat Modarres University, Tehran, Iran, in 2002. Since November 2002, she works as a member of Biometrics Research Group in Iran Telecommunications Research Center(ITRC), Tehran, Iran. Her research interests include computer vision, face detection and recognition, statistical modeling for pattern recognition and image analysis. She has published 5 scientific papers.

**Email:** naderi@itrc.ac.ir

**Mohammad Shahram Moin** received the B.Sc. degree in electrical engineering from Amirkabir University of Technology, Tehran, Iran, in 1988; the M.Sc. degree in electrical engineering from Tehran University, Tehran, Iran, in 1991; and the Ph.D. degree in electrical engineering from École Polytechnique de Montréal, Montréal, Canada, in 2000. From 2000 to 2001, he worked as a Research Associate in Scribens Laboratory, École Polytechnique de Montréal, Montréal, Canada and has been involved in the research on dynamic handwritten document processing. Currently, he is assistant professor and project manager in biometric systems at Iran Telecommunications Research Center, Tehran, Iran.

Dr. Moin is lecturer of artificial neural networks, digital signal processing, computer architecture and advanced programming courses at the undergraduate and postgraduate levels. He has published about 16 papers. His current research interests include neural networks, pattern recognition, image processing, digital signal processing and their application in face recognition, signature verification, speaker recognition, iris recognition and multimodal biometric systems.

**Email:** moin@itrc.ac.ir

**Nasrollah Moghaddam charkari** received the B.Sc. degree in computer science from Shaheed Beheshti University, Tehran, Iran, in 1988 and the M.Sc. and Ph.D. degrees in computer engineering from Yamanashi University, Yamanashi, Japan, in 1986 and 1992, respectively. He is currently an assistant professor of computer engineering and the head of computer engineering department at Tarbiat Modares University, Tehran, Iran. Dr. Moghaddam's research interests include robot vision, image analysis and mining and artificial intelligence. He has graduated more than 15 students since 1995 and published more than 45 scientific papers. He is a member of IEEE, INEC and CSI.

**Email:** Charkari@modares.ac.ir