Subspace Analysis Using Random Mixture Models

Xiaogang Wang[1] *and Xiaoou Tang*[2,3]

[1]Computer Science and Artificial Intelligence Laboratory
Massachusetts Institute of Technology, Cambridge, MA, 02139
[2]Department of Information Engineering
The Chinese University of Hong Kong, Shatin, Hong Kong
[3]Microsoft Research Asia, Beijing, P. R. China

Abstract

In [1], three popular subspace face recognition methods, PCA, Bayes, and LDA were analyzed under the same framework and an unified subspace analysis was proposed. However, since they are all based on a single Gaussian model, a global linear subspace often fails to deliver good performance on the data set with complex intrapersonal variation. They also have to face the problem caused by high dimensional face feature vector and the difficulty in finding optimal parameters for subspace analysis. In this paper, we develop a random mixture model to improve Bayes and LDA subspace analysis. By clustering the intrapersonal difference, the complex intrapersonal variation manifold is learned by a set of local linear intrapersonal subspaces. To boost the system performance, we construct multiple low dimensional subspaces by randomly sampling on the high dimensional feature vector and randomly selecting the parameters for subspace analysis. The effectiveness of our method is demonstrated by experiments on the AR face database containing 2340 face images.

1. Introduction

Subspace analysis has been widely used in face recognition in recent years. PCA [2], Bayes [3], and LDA [4] are three popular subspace methods and they were unified under the same framework in [1]. As shown in [1], the key task of a successful subspace analysis method is to effectively reduce the intrapersonal difference caused by lighting, pose, and expression changes. The intrapersonal subspace of Bayes and LDA focuses on the intrapersonal variation, thus can effectively reduce it.

Both Bayes and LDA use a single Gaussian distribution to model the intrapersonal variation and work well on a relatively simple data set. However, when a data set contains significant transformation difference caused by large lighting, pose, and expression variations, the intrapersonal variation manifold will become highly non-convex and complex. A global linear subspace based on a single Gaussian model often fails to deliver good performance. In this paper, we improve Bayes and LDA subspace analysis using Gaussian mixture models. The complex intrapersonal variation manifold is decomposed into several clusters with simple distributions, and learnt by multiple local intrapersonal subspaces.

Although several mixture models linear subspace methods [5][6][7] have been proposed in previous work, their clustering procedures are all based on face images or face class centers instead of intrapersonal difference, which is the most significant factor affecting recognition performance. The intrapersonal variation cannot be effectively reduced in their methods. These approaches usually require that the face class has at least one reference image in each cluster and this condition is difficult to meet in real applications. Our method needs only one reference sample for each face class in recognition.

Another problem for previous methods is that face feature vector dimensionality is usually very high compared with the small training set. This often leads to biased and unstable results. In order to overcome this problem, in this paper, we apply the random sampling LDA we developed earlier [8] in the new mixture model to develop a random mixture model. In addition, we further extend the random sampling process to randomly selecting the parameters to boost the system performance.

In summary, our subspace analysis algorithm based on random mixture models can be understood as a framework integrating multiple subspaces, which are constructed in three steps: 1. clustering the intrapersonal differences from the training samples; 2. randomly sampling the high dimensional feature vector; and 3. randomly selecting parameters. It is more stable and more effective on the difficult data set with complex distribution.

2. Subspace Analysis Based on a Single Gaussian Model

We first briefly review several conventional subspace analysis methods, including Bayes, LDA, and null space LDA. They are all based on a single Gaussian Model.

2.1. Bayesian Analysis

In the Bayesian algorithm, the similarity between two images can be measured as the intrapersonal likelihood

Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)
1063-6919/05 $20.00 © 2005 **IEEE**

$P(\Delta \mid \Omega_I)$. Here Δ is the difference between face images and Ω_I is the intrapersonal variation. Principal component analysis (PCA) is applied on the intrapersonal difference set $\{\Delta \mid \Delta \in \Omega_I\}$ to compute the intrapersonal principal subspace F and its complementary subspace $\overline{F}$. It assumes that Ω_I has a Gaussian distribution. $P(\Delta \mid \Omega_I)$ is estimated as the product of two independent marginal Gaussian densities in F and $\overline{F}$,

$$P(\Delta \mid \Omega_I) = \left[\frac{\exp\left(-\frac{1}{2} d_F(\Delta)\right)}{(2\pi)^{M/2} \prod_{i=1}^{M} \lambda_i^{1/2}} \right] \left[\frac{\exp\left(-\varepsilon^2(\Delta)/2\rho\right)}{(2\pi\rho)^{(N-M)/2}} \right]. \quad (1)$$

In Eq. (1), $d_F(\Delta)$ is a Mahalanobis distance in F, referred as "distance-in-feature-space" (DIFS),

$$d_F(\Delta) = \sum_{i=1}^{M} \frac{y_i^2}{\lambda_i}, \quad (2)$$

where y_i is the principal component of Δ projected into F and λ_i is the eigenvalue. $\varepsilon^2(\Delta)$ is defined as "distance-from-feature-space" (DFFS), equivalent to PCA residual error in $\overline{F}$.

In the recognition procedure, Eq.(1) is equivalent to evaluating the distance

$$d(\Delta) = d_F(\Delta) + \varepsilon^2(\Delta)/\rho. \quad (3)$$

Both DIFS and DFFS are two distinctive components and thus can be independently used for recognition. If the intrapersonal variation Ω_I has a Gaussian distribution, PCA on $\{\Delta \mid \Delta \in \Omega_I\}$ computes a set of principal axes dominated by the energy of Ω_I. When a face difference Δ is projected into the intrapersonal subspace, its intrapersonal variation is therefore compacted onto a small number of large eigenvectors in F. Since λ_i explicitly describes the energy distribution of intrapersonal variation, intrapersonal variation can be effectively reduced by the inverse weighting of eigenvalues in DIFS. DFFS is also a distinctive component for recognition. It throws away most intrapersonal variation on large eigenvectors.

2.2. LDA

LDA tries to find a set of projecting vectors W maximizing the ratio of determinant of the between-class scatter matrix S_b and the determinant of the within-class scatter matrix S_w,

$$W = \arg\max \left| \frac{W^T S_b W}{W^T S_w W} \right|. \quad (4)$$

W can be computed from the eigenvectors of $S_w^{-1} S_b$ [9]. In face recognition, the training set is usually small compared to the high dimensional feature vector. To avoid the singularity of S_w, most of the LDA methods first reduce the face data dimension by PCA and then apply discriminant analysis in the reduced PCA subspace.

Computing the eigenvectors of $S_w^{-1} S_b$ is equivalent to simultaneous diagonalization of S_w and S_b [1]. First, Compute the eigenvector matrix Φ and eigenvalue matrix Θ of S_w. It has been proved that the subspace spanned by Φ and Θ is essentially the intrapersonal subspace computed in the Bayesian subspace analysis. Next project the class centers onto Φ and normalize it by $\Theta^{-1/2}$. Thus S_b is transformed to $K_b = \Theta^{-1/2} \Phi^T S_b \Phi \Theta^{-1/2}$. This whitening process reduces the intrapersonal variation just like the DIFS in Bayes. After computing the eigenvector matrix Ψ and eigenvalue matrix Λ of K_b, the projection vectors of LDA can be defined as $W = \Phi\Theta^{-1/2}\Psi$. This step further makes the class centers distant.

2.3. Null Space LDA

Chen *et. al.* [10] suggested that the null space of S_w, in which $W^T S_w W = 0$, also contains much discriminative information. It is possible to find some projection vectors W satisfying $W^T S_w W = 0$ and $W^T S_b W \neq 0$, thus the Fisher criteria in Eq. (4) definitely reaches its maximum value. A LDA approach in the null space of S_w was proposed. First, the null space V of S_w is computed as,

$$V^T S_w V = 0. \quad (5)$$

S_b is projected to the null space of S_w,

$$\tilde{S}_b = V^T S_b V. \quad (6)$$

The LDA projection vectors are defined as $W = V\Phi$, where Φ contains the eigenvectors of $\tilde{S}_b$ with the largest eigenvalues.

The null space of S_w is equivalent to the complementary intrapersonal subspace $\overline{F}$ in Bayes. Most of the intrapersonal variation has been removed in this space. Null space LDA further makes the class centers in this space distant.

3. Subspace Analysis Using Random Mixture Models

A common theme for Bayes, LDA, and Null space subspace analysis is to use the intrapersonal subspace to reduce the intrapersonal variation. However, they are all based on a uniform Gaussian model. When the intrapersonal difference in the face data set is very large, the intrapersonal variation manifold will be too complex to be modeled as a single Gaussian distribution. The derived intrapersonal subspace cannot effectively reduce the intrapersonal variation. We project the intrapersonal differences of samples from AR database onto the first two eigenvectors of the intrapersonal subspace and plot them in Figure 1. Apparently it is not a Gaussian distribution. A better choice is to decompose the complex manifold into K simpler clusters and use multiple intrapersonal subspaces to model local regions. Based on this consideration, we propose random Gaussian mixture models to model intrapersonal differences.

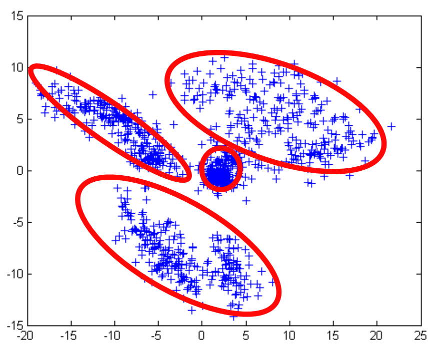

Figure 1. Project the samples of intrapersonal difference from AR database to the first two eigenvectors of the intrapersonal subspace. The complex distribution can be better modeled as several local Gaussian models.

3.1. Bayesian Subspace Analysis Based on Random Mixture Models

In Section 2.1, it is shown that both DIFS and DFFS in Bayes are distinctive components and can be independently used for recognition. Here, we first improve them using the random mixture models respectively, and then integrate them using a fusion rule.

3.1.1. DIFS Random Mixture Models

The proposed algorithm of DIFS mixture models classifier is shown in Figure 2. In the training stage, the intrapersonal difference training sample set is first constructed by computing the difference between the training sample and its class center. They are clustered based on the Mahalanobis distance. For each intrapersonal difference cluster, a local intrapersonal subspace is computed and the DIFS is computed as,

$$d_F^k(\Delta) = \sum_{i=1}^{N} (y_i^k)^2 / \lambda_i^k \tag{7}$$

where y_i^k is the principal component of the face difference Δ projecting to the kth local intrapersonal subspace and λ_i^k is the intrapersonal eigenvalue.

When a probe face image is input, we compute its difference Δ with the reference face class center. Compute DIFS of Δ in each local intrapersonal subspace and choose the minimum one as its distance measure. The corresponding local intrapersonal subspace is chosen as the best one to model Δ. As we can see from Figure 1, this local intrapersonal subspace should estimate a more accurate DIFS than using the global intrapersonal subspace.

Since the face feature vector dimension is very high while the training set size is relatively small, some intrapersonal eigenvalues are close to zero and overfitting will happen using Eq. (7) for clustering and recognition. This problem becomes even more serious in the DIFS mixture model, since each cluster only contains part of the training samples. We adopt the random subspaces to solve this problem [8]. Some low dimensional random subspaces are generated by random sampling on the original high dimensional face feature vector. A DIFS mixture model classifier is constructed from each random subspace. Since it is based on a low dimensional feature vector, it is more stable. The multiple classifiers can cover all the discriminative information in the face space.

Another key problem for the mixture model classifier is how to choose the proper cluster number. In order to avoid seriously deteriorating the recognition performance by incorrectly choosing a cluster number, we randomly choose the cluster number k from a proper range ($k_{\min} < k < k_{\max}$) in each random subspace and combine the multiple classifiers to boost the system performance. The final algorithm of the DIFS random mixture models is shown in Figure 3.

3.1.2. DFFS Random Mixture Models

Similar to the procedure described in Figure 2, DFFS can also be extended to the mixture models. The only difference is that DFFS instead of DIFS is used for clustering and recognition in the training step 5 and recognition step 2. The face difference Δ is projected into the local intrapersonal subspace spanned by the r largest eigenvectors $E^k = \left[\vec{u}_1^k, \ldots, \vec{u}_r^k\right]$ and the reconstruction error (DFFS) is computed as

$$\varepsilon_k^2(\Delta) = \left\| (\Delta - \bar{\alpha}_k) - E^k (E^k)^T (\Delta - \bar{\alpha}_k) \right\|^2 . \tag{8}$$

Different from DIFS, DFFS does not suffer from the high dimensionality of feature vector. In fact, it extracts the discriminative information encoded on the intrapersonal eigenvectors with zero eigenvalues. So we do not need to use random subspaces to reduce the feature dimension. However, DFFS mixture models classifier

DIFS Mixture Models Classifier

Input: A set of training samples $\{\vec{x}_i\}_{i=1}^{M}$; the number (K) of local intrapersonal subspaces.

Training:

Step 1: Compute the face class centers $\{\vec{m}_l\}_{l=1}^{L}$ of the training samples.

Step 2: Compute the intrapersonal difference between the training sample and its class center, $\Delta = \vec{x}_j - \vec{m}_l$. Construct the intrapersonal difference training sample set $\{\Delta_i\}_{i=1}^{M}$.

Step 3: Randomly choose the initial cluster assignment $\ell(\Delta) \in \{1,\ldots,K\}$ for each intrapersonal difference sample Δ in the training set.

Step 4: Compute the eigenvectors $U^k = \left[\vec{u}_1^k,\ldots,\vec{u}_N^k\right]$, eigenvalues $\Lambda^k = \left[\lambda_1^k,\ldots,\lambda_N^k\right]$, and cluster center $\vec{\alpha}_k$ for each cluster $E_k = \{\Delta \mid \ell(\Delta) = k\}$.

Step 5: Project the training intrapersonal difference sample Δ into each local intrapersonal subspace $y_i = (\vec{u}_i^k)^T(\Delta - \vec{\alpha}_k)$, and compute the DIFS Mahalanobis distance $d_F^k(\Delta) = \sum_{i=1}^{N} (y_i^k)^2 / \lambda_i^k$.

Step 6: Assign Δ to the cluster with the minimum DIFS distance.

Step 7: Stop if no training example has changed cluster, otherwise, return to step 4.

Recognition:

Step 1: When a probe face image $\vec{x}$ is input, compute its difference with each face class center, $\Delta_l = \vec{x} - \vec{m}_l$.

Step 2: Project Δ_l into each local intrapersonal subspace and compute the DIFS as $d_F^k(\Delta_l)$.

Step 3: Choose the minimum DIFS in the local intrapersonal subspace as the distance of probe image $\vec{x}$ to class center $\vec{m}_l$, $d(\Delta_l) = \min_k \{d(\Delta_l^k)\}$.

Step 4: Recognize the face class with the minimum distance to the probe face image $\vec{x}$, $\ell(\vec{x}) = \arg\min_l (d(\Delta_l))$.

Figure 2. Algorithm of DIFS mixture models classifier.

also needs several parameters to be set. Besides the local intrapersonal subspace number K, the intrapersonal eigenvector numbers (r_1 and r_2) to compute the reconstruction error (DFFS) in clustering and recognition are also required as input. Different parameters and cluster initialization will lead to different classifier. We also construct multiple classifiers by randomly selecting parameters from proper range and randomly setting the cluster initialization, and combine them using a fusion rule.

3.1.3. Integrating DIFS and DFFS

DIFS and DFFS extract discriminative information from two complementary subspaces. So they are complementary to each other. Using a fusion rule, we further combine the classifiers constructed by DIFS random mixture models and DFFS random mixture models to cover more discriminative information in the face space.

3.2. LDA and Null Space LDA Random Mixture Models

Section 2 has shown that DIFS and DFFS can be viewed as intermediate steps of LDA and null space LDA. So LDA and null space LDA can also be extended to the random mixture models. The algorithms of LDA and null space LDA mixture models classifiers are shown in Figure 4 and Figure 5. Based on the Bayes mixture model, they further project the class centers into the local intrapersonal subspace and its null space, and seek the principal directions to make the class centers further separated in each local subspace. They are extended to random mixture models using random subspaces and randomly selecting parameters in the same way as shown in Figure 3.

3.3. Discussion

Our subspace analysis using random mixture models can be understood as a multiple classifier integration framework. Since the face data has a very complex

DIFS Random Mixture Models Classifier

Input: A set of training sample $\{\vec{x}_i\}_{i=1}^{M}$; the number of random subspaces H; the dimension of random subspaces N; the range of the local intrapersonal subspace number in each random subspace ($k_{\min}$, $k_{\max}$).

Step 1: Generate the feature indexes $(S_j)_{j=1}^{H}$ to construct random subspaces with dimension N by random sampling on the high dimensional face feature vector. $\{\vec{x}_i(S_j)\}_{i=1}^{M}$ is the new feature vector set of training samples projected to the random subspace S_j .

Step 2: Randomly select the local intrapersonal subspace number $k_j(k_{\min} \le k_j \le k_{\max})$ in the random subspace S_j .

Step 3: Construct DIFS mixture models classifier $C_j(\{\vec{x}_i(S_j)\}_{i=1}^{M}, k_j)$ in random subspace S_j .

Step 4: Combine the H DIFS mixture models classifiers $\{C_j\}_{j=1}^{H}$ using a fusion rule.

Figure 3. Algorithm of DIFS random mixture models.

distribution in a very high dimensional space, it is difficult to be modeled using the conventional single classifier framework. Instead of constructing a very complicated single classifier, we propose to use multiple simple classifiers to model face difference subspaces and local regions of the face space. Multiple classifiers are constructed by clustering the training samples, randomly sampling the feature vector and randomly setting the classifier parameters. Some proper classifiers are selected from training sample clusters. The multiple classifiers constructed from random subspaces and random parameters are combined. Here we use majority voting as a fusion rule. More advanced multiple classifiers combination methods can be found in [11]. They can be incorporated into our framework to further improve the system performance.

4. Experiments

We conduct experiments on a data set from AR database. It contains 90 subjects (subjects with a full set of photos) and each subject has 26 face images taken in two sessions. For each session, there are 13 face images. Face images in this data set have very significant intrapersonal variations. Some examples are shown in Figure 6. The 1170 face images taken in the first sessions are used for training set to compute the intrapersonal subspace. When testing the 90 class centers of 1170 face images in the first session are used as reference, and the 1170 face images taken in the second session are used as probe. In preprocessing, all the images are normalized for scaling, translation, and rotation, such that the eye centers are in fixed positions. A rectangle mask is used to remove the background and most of the hair region.

Recognition accuracies of linear subspace methods based on the uniform model are shown in Table 1. Because of the significant intrapersonal variation, the manifold as shown in Figure 1 cannot be modeled using a single Gaussian model. Recognition accuracies of these linear subspace methods are extremely low, with around 50% accuracy.

As shown in Table 2 and Table 3, our random mixture models framework significantly improves the performance of these subspace analysis methods on the data set with complex intrapersonal variations. In Figure 7, 8, 9, 10, we plot the accuracy of each random mixture models classifier and the result of multiple classifiers combination. For DIFS and LDA, the recognition rate of the mixture models classifier in each random subspace is low. This is because much information has been discarded in each random subspace. But the recognition performance is significantly improved after combining the multiple classifiers, since the multiple random subspaces can cover almost all the discriminative information in the face space after combination. For DFFS and null space LDA, each single mixture models classifier has already significantly outperformed the classifier based on uniform Gaussian model. However, when using different parameters and cluster initialization, the accuracy has some variation. Using the random generation and integration framework, the system performance is stabilized and is further improved in accuracy. We test the mixture LDA method proposed in [7] by grouping the face classes into four clusters. It fails to improve the performance with only 45% accuracy, because the complex distribution of this data set is caused by large intrapersonal variations, which still exist in each cluster of that method. Several other methods require the face class has at least one reference image in each cluster. They cannot be used in this case since we only use one reference image for each face class.

5. Conclusion

In this paper, we proposed a random mixture models framework for subspace analysis. It focuses on the intrapersonal variation and uses multiple local subspaces to model the complex face data manifold. Different local subspaces are generated by clustering the intrapersonal difference training samples, randomly sampling the feature vector and randomly setting the parameters. Experiments show that our framework significantly boosts the face recognition system and outperforms conventional uniform model subspace methods on the face data set with complex distribution.

Acknowledgment

This work was fully supported by grants from Hong Kong RGC and a Hong Kong RGC/NSFC joint project N_CUHK409/03. This work was performed while both authors were at the Chinese University of Hong Kong.

References

[1] X. Wang and X. Tang, "A Unified Framework for Subspace Face Recognition," IEEE Trans. PAMI, Vol. 26, No. 9, pp. 1222-1228, Sept. 2004.

[2] M. Turk and A. Pentland, "Face Recognition Using Eigenfaces," *in Proceedings of CVPR*, pp. 586-591, Hawaii, June, 1991.

[3] B. Moghaddam, T. Jebara, and A. Pentland, "Bayesian Face Recognition," Pattern Recognition, Vol. 333, pp. 1771-1782, 2000.

[4] P. N. Belhumeur, J. Hespanda, and D. Kiregeman, "Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection," *IEEE Trans. on PAMI,* Vol. 19, No. 7, pp. 711-720, July 1997.

[5] A. Pentlabd, B. Moghaddam, and T. Starner, "View-Based and Modular Eigenspaces for Face Recognition," in *Proceedings of CVPR,* 1994.

[6] E. Demir, L. Akarun, and E. Alpaydin, "Two-Stage Approach for Pose Invariant Face Recognition," in *Proceedings of ICASSP*, Vol.4, pp. 2342-2344, 2000.

[7] J. Lu and K. N. Plataniotis, "Boosting face recognition on a large-scale database," in *Proceedings of ICIP*, Vol. 2, pp. 109-112, 2002.

[8] X. Wang and X. Tang, "Random Sampling LDA for Face Recognition," *in Proc. of CVPR*, Vol. 2, pp. 259-265, 2004.

[9] K. Fukunnaga, "Introduction to Statistical Pattern Recognition," Academic Press, second edition, 1991.

[10] L. Chen, H. Liao, M. Ko, J. Lin, and G. Yu, "A New LDA-Based Face Recognition System Which Can Solve the Small Sample Size Problem," *Pattern Recognition*, Vol. 33, pp. 1713-1726, Oct. 2000.

[11] J. Kittler and F. Roli, (Eds): Multiple Classifier Systems.

LDA Mixture Models Classifier

Input: A set of training samples $\{\vec{x}_i\}_{i=1}^{M}$; the number of local intrapersonal subspaces *K*.

Training:

Step 1: Cluster the intrapersonal difference training samples and compute the local intrapersonal subspaces, (U^k, Λ^k) using DIFS.

Step 2: Project the class centers $\vec{m}_l$ into each intrapersonal subspace $\vec{\mu}_l^k = (\Lambda^k)^{-1/2}(U^k)^T(\vec{m}_l - \vec{\alpha}_k)$. Apply PCA on $\{\vec{\mu}_l^k\}$ and compute the eigenvectors Ψ^k.

Step 3: The LDA projection vectors in the *k*th local intrapersonal subspace is $W^k = U^k(\Lambda^k)^{-1/2}\Psi^k$.

Recognition:

Step 1: Compute the difference Δ between the probe face and reference face class center.

Step 2: Choose the local intrapersonal subspace *k* with the minimum DIFS to Δ.

Step 3: Project Δ to W^k and compute the distance for recognition.

Figure 4. Algorithm of LDA mixture models.

Null Space LDA Mixture Models Classifier

Input: A set of training samples $\{\vec{x}_i\}_{i=1}^{M}$; the number of local intrapersonal subspace *K*; the eigenvector number *r* to compute the reconstruction error.

Training:

Step 1: Cluster the intrapersonal difference training samples and compute the local intrapersonal subspaces (U^k, Λ^k) using DFFS.

Step 2: Project the class centers $\vec{m}_l$ to the null spaces of the local intrapersonal subspaces

$$\vec{\beta}_l^k = (\vec{m}_l - \vec{\alpha}_k) - U^k(U^k)^T(\vec{m}_l - \vec{\alpha}_k).$$

Apply PCA on $\{\vec{\beta}_l^k\}$ and compute the eigenvectors Ψ^k.

Step 3: The null space LDA projection vectors for the *k*th local intrapersonal subspace is

$$W^k = \left(I - U^k * (U^k)^T\right)\Psi^k$$

Recognition:

Step 1: Compute the difference Δ between the probe face and reference class center.

Step 2: Choose the local intrapersonal subspace *k* with the minimum DFFS to Δ.

Step 3: Project Δ to W^k and compute the distance for recognition.

Figure 5. Algorithm of null space LDA mixture models.

Table 1. Recognition accuracies of subspace methods based on uniform model (%)

PCA	DIFS	DFFS	DIFS + DFFS	LDA	Null space LDA
22.14	52.91	51.88	52.14	50.09	55.64

Table2. Recognition accuracies of Bayes random mixture models classifiers (%)

DIFS	DFFS	Integrate DIFS and DFFS
74.87	77.18	78.12

Table 3. Recognition accuracies of LDA and Null space LDA (N-LDA) random mixture models classifiers (%)

LDA	N-LDA	Integrate LDA and N-LDA
73.16	77.09	77.56

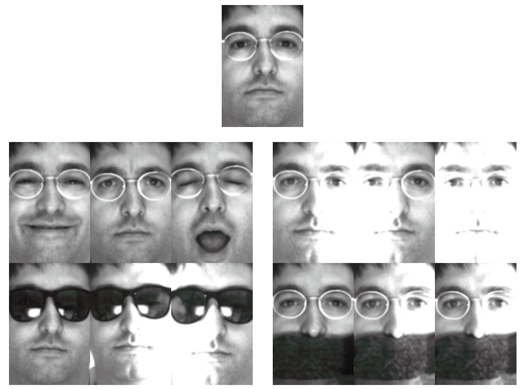

Figure 6. Face image examples taken in the same session for one subject in AR database.

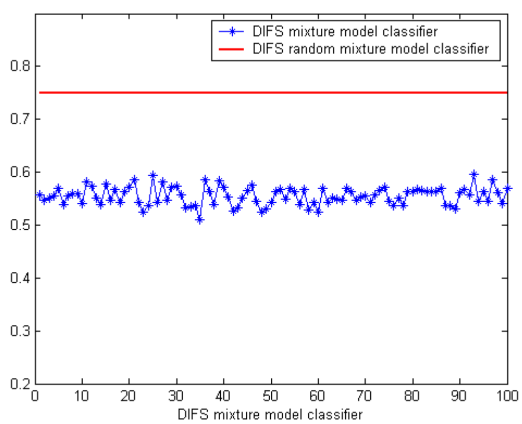

Figure 7. Recognition accuracies of DIFS mixture models classifiers and result of combining multiple classifiers under a random model.

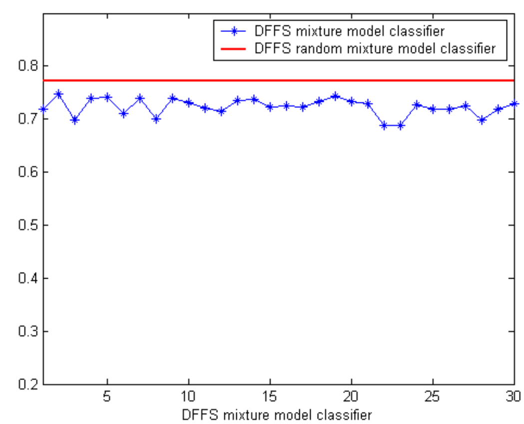

Figure 8. Recognition accuracies of DFFS mixture models classifiers and result of combining multiple classifiers under a random model.

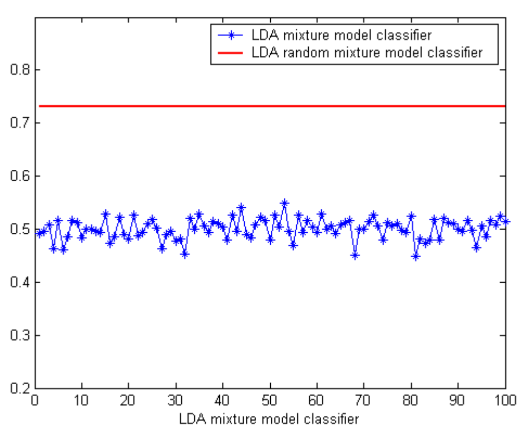

Figure 9. Recognition accuracies of LDA mixture models classifiers and result of combining multiple classifiers under a random model

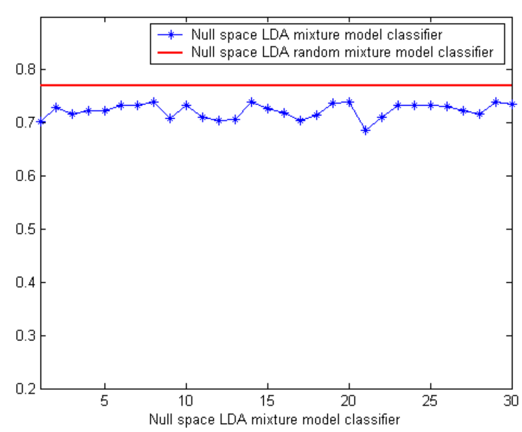

Figure 10. Recognition accuracies of null space LDA mixture models classifiers and result of combining multiple classifiers under a random model.