

1-1-2014

Face recognition across pose variation and the 3S problem

JAMAL HUSSAIN SHAH

MUHAMMAD SHARIF

MUDASSAR RAZA

AISHA AZEEM

Follow this and additional works at: <https://journals.tubitak.gov.tr/elektrik>



Part of the [Computer Engineering Commons](#), [Computer Sciences Commons](#), and the [Electrical and Computer Engineering Commons](#)

Recommended Citation

SHAH, JAMAL HUSSAIN; SHARIF, MUHAMMAD; RAZA, MUDASSAR; and AZEEM, AISHA (2014) "Face recognition across pose variation and the 3S problem," *Turkish Journal of Electrical Engineering and Computer Sciences*: Vol. 22: No. 6, Article 3. <https://doi.org/10.3906/elk-1108-70>
Available at: <https://journals.tubitak.gov.tr/elektrik/vol22/iss6/3>

This Article is brought to you for free and open access by TÜBİTAK Academic Journals. It has been accepted for inclusion in Turkish Journal of Electrical Engineering and Computer Sciences by an authorized editor of TÜBİTAK Academic Journals. For more information, please contact academic.publications@tubitak.gov.tr.

Face recognition across pose variation and the 3S problem

Jamal Hussain SHAH, Muhammad SHARIF*, Mudassar RAZA, Aisha AZEEM

Department of Computer Sciences, COMSATS Institute of Information Technology, Wah Cantt., Pakistan

Received: 23.08.2011 • Accepted: 11.06.2012 • Published Online: 07.11.2014 • Printed: 28.11.2014

Abstract: Most face recognition methods are based on linear and nonlinear subspace features extraction and classification tasks. These classification methods are used for global and local facial features for person identification. Both local and global features play different roles for recognition and classification. There are a number of face recognition methods that have been proposed up until now, and they produce good results. However, when small sample size (3S) and pose variation problems are taken into consideration, face recognition becomes more complex and does not produce good results. In this paper, 3S and pose variation problems are dealt with. First, linear discriminate analysis (LDA) is considered to minimize the singularity problem that arises when small samples of individuals are available. In the next step, the proposed framework utilizes global and local facial features and constructs a combined subspace using an enhanced LDA method that is discussed later in the sections.

Key words: Global and local features, small sample size problem, combined subspace

1. Introduction

Face recognition is an imperative phenomenon in still and video images due to its variety of functions, like in identity verification, human-computer interaction, and video surveillance. It has gained much attention within the past 2 decades, which has resulted in a large number of algorithms handling face recognition tasks [1]. Although it has an extensive range of applications, some issues that arise in the zone of face recognition are illumination variation, viewing directions, facial expression changes, aging, and occlusion due to hair, glasses, or makeup.

Face recognition seems to be a very challenging task under different variations of images, e.g., head and pose. Like in the other problems of face recognition as listed earlier, pose variation is tremendously important in many applications. For nonfrontal faces, at least 2 normalized facial features are required, but in a situation where the faces are in-depth rotated, geometrical normalization is practically impossible. The complexity becomes a major issue when other problems of face recognition like illumination [2], occlusion [3], pose variation [4], and expressions [5] are present on nonfrontal faces.

Categorically, pose variation algorithms are grouped into 3 main approaches, which are 1) the invariant feature extraction-based approach 2) the multiview-based, and 3) the 3D range image-based technique. The first approach learns the subspace that symbolizes the amount of variations. On the other hand, model-based methods exploit a 2D deformable model in order to capture the pose variations. Finally, multiview-based approaches accumulate multiple images of different poses referred to as real multiview images.

The most common approaches based on invariant features are appearance- and geometric-based, to handle

*Correspondence: muhammadsharifmalik@yahoo.com

the pose variation problem in face recognition. In appearance-based approaches, the image pixels are represented in a high-dimensional vector space and are converted into a low-dimensional feature space. Most of the algorithms used in this class are principal component analysis (PCA) [6,7] and Fisher linear discriminant/linear discriminant analysis [8,9] algorithms. Many algorithms have been proposed by different authors with slight variations to handle the same problem [10]. A second class of invariant feature-based approaches is geometric model algorithms, which create a model to estimate the variation. In this class, commonly used algorithms are the elastic bunch graph matching method [11], active shape model [12], and active appearance model [13]. According to their nature, these methods are effective in face recognition under pose verification. Many authors adopt these methods to recognize the face as presented in [14–16].

The second category of face recognition under pose variation is multiview-based and it synthesizes new images from multiple poses. The basic idea behind synthesized images is to apply the synthesized algorithm using the previous information of a given pose. As the multiview-based approaches use the previous information of the face images, it is considered more practical and efficient for face recognition tasks under varying viewpoint problems. Different authors have proposed many algorithms related to multiview-based statistical methods, as discussed in [17,18]. The most commonly used method to handle pose is the statistical model, which is obtained for each subject of class at every possible angle [19,20]. An interrelated concept to such a method is to exploit a number of samples of each subject under varying poses and build a general statistical model that interposes to unseen views.

The third most appropriate method for face recognition under varying pose is to learn the statistical information between the frontal and nonfrontal faces instead of the 3D field oriented information. It is constructed using the numeric or geometric information of multiple samples of individuals and is used to compare with the probe image by applying re-rendering at any given pose. Additionally, the same procedure is applied on the appropriate pose of an image. Many authors use the 3D model strategy to recognize the face under pose variation, and recently, a generic elastic model (GEM)-based system [21,22] was introduced to handle the same problem. The more advanced approach is to use the 3D measurements to align the probe and gallery images, which eradicate the pose used for face recognition. All of the methods performed are efficient and produce accurate results but still the statistical model requires particular methods and user collaboration. Moreover, it only works in a situation when each subject has multiple samples to construct the statistical model. The limitations across these approaches indicate that they are not suitable for subjects having a single image. The goal of each technique mentioned above is to recognize faces under varying pose variation. Even though some techniques provide the required results, they do not offer a more accurate domino effect.

In view of the fact of being a part of face recognition, feature extraction plays a pivotal role in recognizing faces within the least processing time. In feature extraction, selective features are processed in order to make it feasible to recognize intrinsic and extrinsic facial variations under a reasonable computational cost. Basically, face recognition algorithms are categorized as local- and global-based approaches. In global-based face representation, the global features are embodied in every part of the face image that contains pixel level information, called holistic data, which is sometimes irrelevant during face recognition.

In this paper, the main emphasis is on the pose variation and small sample size (3S) problem. In the training process, a number of individuals per sample are expected to give better results. Unfortunately, in a real environment, only single training samples are available.

2. Proposed work

There are many problems in face recognition, as discussed in the previous section along with some solutions provided by different authors. Though it has an extensive range of applications, face recognition deals with some serious issues and in particular, the pose variation problem and 3S problem. A 3S problem results when the number of input samples is less than its dimensions. The same problem arises when we use template-based approaches. Other than the template-based approaches, linear and nonlinear approaches also exist for face recognition. Some linear-based approaches are PCA, independent component analysis, and linear discriminate analysis (LDA) and so on. In contrast to other methods, LDA is a more powerful feature extraction and recognition technique, but the major problem in LDA is the 3S due to the singular matrix. In this paper, a solution for the 3S problem in LDA is given. Later, the proposed framework, face recognition across pose variation and the 3S problem (FAPSP), is used to handle the pose variation problem. The discussion and process of this novel technique is given in the later sections.

2.1. How a 3S problem arises

Among the discriminant analysis algorithms, LDA is considered to be a dominant batch classification subspace classifier. This classifier has the ability to project a high dimensional pattern onto a low dimensional space, while satisfying the Fisher's criterion as [8]:

$$J_{LDA}(w) = \frac{|w^T S_b w|}{|w^T S_w w|}. \quad (1)$$

The projection matrix has a generalized form that satisfies $S_w^{-1} S_b$. However, the within-class scatter matrix is singular, and so its inverse is not possible, and hence the computation of the leading eigenvalues cannot be achieved. The 3S problem is the main reason for the singularity constraint of the traditional LDA, for which various authors have proposed solutions over the past several years.

2.2. Singular to nonsingular S_w

It has been discussed in the previous section that the within-class scatter matrix becomes singular when the individual training samples are minimum. In this section, we performed an operation on the diagonal value of the S_w matrix to overcome the singularity problem.

Given a within-class scatter matrix S_w , an eigenvalue λ and its associated eigenvector V that are, by definition, a pair obeying the relation

$$S_w V = \lambda V. \quad (2)$$

It can be tested as a matrix Γ that is generated by the subtraction of λI (the combination of λ and identity matrix I) and S_w . Their mathematical notations are depicted as:

$$\Gamma = S_w - \lambda I, \quad (3)$$

but

$$\det(\Gamma) = 0. \quad (4)$$

In order to calculate the diagonal matrix, we determine the λ_N eigenvalues using Eq. (2) as follows:

Consider,

$$\Psi = [\lambda_1, \lambda_2, \lambda_3, \dots, \lambda_N], \quad (5)$$

where Ψ shows the N eigenvalue, the mean ϖ of Ψ is computed in Eq. (6) as:

$$\varpi = \frac{1}{N} \sum_{i=1}^N \Psi_i. \quad (6)$$

The next step is to sort the Ψ eigenvalues in descending order and place them in a diagonal matrix D (S_w) (which is equal to the size of S_w). The diagonal matrix is later divided by its mean ϖ , which can be written in a form of Δ as:

$$\Delta = \frac{D(S_w)}{\varpi}. \quad (7)$$

Now add Δ to the original S_w , which can be mathematically depicted as:

$$S_w^{New} = S_w + \Delta. \quad (8)$$

To find the actual within-class scatter matrix S_w^{New} , perform the following steps. Consider that $d(S_w)$ is the diagonal entry of S_w and $S_w(\emptyset)$ is the null matrix whose size is equal to the actual S_w :

$$d(S_w) = [a_{11}, a_{22}, a_{33}, \dots, a_{NN}]. \quad (9)$$

Now, add the null matrix $S_w(\emptyset)$ with $d(S_w)$ represented as:

$$S_w^\varphi = S_w(\varphi) + d(S_w). \quad (10)$$

This can be represented in the form of a matrix as:

$$S_w^\varphi = \begin{bmatrix} 0 & 0 & \cdot & \cdot & 0 \\ 0 & 0 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot & 0 \end{bmatrix} + [a_{11}, a_{22}, a_{33}, \dots, a_{NN}]. \quad (11)$$

Subtract S_w^φ from S_w^{New} to find the actual within-class scatter matrix that minimizes the singularity problem in S_w . Hence, the modified S_w^{New} is:

$$S_w^{New} = S_w^{New} - S_w^\varphi. \quad (12)$$

From Eqs. (12) and (1), the modified form of Eq. (1) is given as:

$$J_{LDA}^{New}(w) = \frac{|w^T S_b w|}{|w^T S_w^{New} w|}. \quad (13)$$

The experimental results on the modified LDA $J_{LDA}^{New}(w)$ in Eq. (13) reveal that the 3S problem of the individually trained images is minimized (near 0). The result can be verified by calculating the determinant of S_w^{New} , which shows that it is no longer singular $\det(S_w^{New}) \neq 0$, as compared to Eq. (4). Another advantage of the proposed S_w^{New} is that there is no effect on the within-class scatter matrix S_w , and it minimizes the distance of the within-class, and merely maximizes the distance of the between-class scatter matrix after projection into the subspace. The investigation of $J_{LDA}^{New}(w)$ on different test benches reveals that it increases the recognition result with the least time complexity.

2.3. Proposed framework

This section details the proposed face recognition technique. Facial features are extracted from interperson frames from different datasets under varying pose variations. The datasets are taken from the Facial Recognition Technology (FERET) and Carnegie Mellon University Pose, Illumination, and Expression (CMU-PIE) databases, which are later discussed in Section 3. The next section portrays the methodology of how to generate interperson frames from a morphological operation. The local and global facial features are extracted in the combined subspace and are then used for recognition purposes. The general framework of entire proposed system is depicted in Figure 1.

2.3.1. Generate interperson frame

In this section, the mesh-morphing technique [23] is used to generate the interpersonal frame. For interpolation, we use the biharmonic spline interpolation, developed by Sandwell [24], rather than bicubic interpolation. First, get the 2 images, I_s and I_d , as the source and destination images. Note that the points on the mesh are marked manually. Now, use an $N \times 2$ matrix to save the mesh points, where N is the number of points associated with the source and destination images.

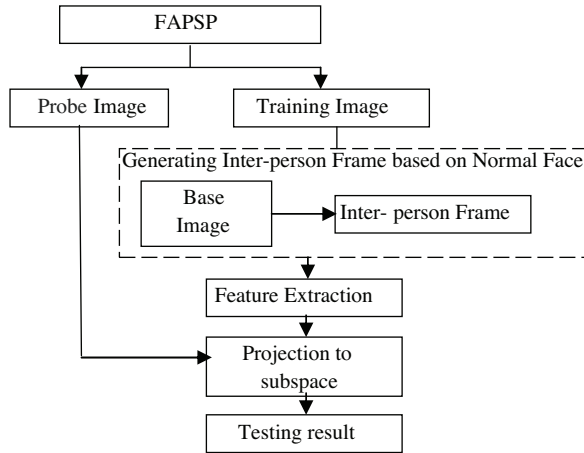


Figure 1. Overview of the face recognition system.

Similarly, $I_I(x,y)$ is an intermediate frame of the same size as the source and destination. Apply the associated weights with M_s and M_d , where M_s and M_d are the corresponding mesh of the source and destination images. Calculate the weights with the following equations:

$$W = np - 1/N - 1, \quad (14)$$

$$W1 = (1 - W), \quad (15)$$

where W and $W1$ are the calculated weights and np and N are the number of intermediate frames and number of mesh points, respectively. From the calculated weights, find the points:

$$T_a = M_x(x, y) * W, \quad (16)$$

$$T_b = M_y(x, y) * W1, \quad (17)$$

where T_a and T_b are the associated points of the $M \times N$ source and destination images. Now, use Eqs. (1) to (4) to combine the images to generate the intermediate frames using the following equation:

$$I_I(x, y) = W \times I_i(x + a, y + b) + W1 \times I_f(x - a, y - b), \quad (18)$$

$$V_n(x, y) = \{I_{I1}(x, y), I_{I2}(x, y), \dots, I_{In}(x, y)\}, \quad (19)$$

where $V_n(x, y)$ is the intermediate frame of the source to the destination images. After getting the interperson frames, there is some noise. To remove this noise, use the spatial filter with the following 3×3 mask MS_k [25]:

$$MS_k = \begin{pmatrix} 0.022 & 0.044 & 0.066 \\ 0.089 & 0.111 & 0.133 \\ 0.156 & 0.178 & 0.200 \end{pmatrix}. \quad (20)$$

2.3.2. Feature extraction

Facial features play a key role in face recognition systems. To date, different methods have been used for this purpose. Some authors have used local features (eyes, nose, mouth, etc.) while others make use of global features, i.e. taking the whole face into consideration. Although the individually extracted method works fine, to improve the recognition performance, hybrid approaches are being used.

2.3.2.1. Extraction of the global facial features

It has already been mentioned in the previous discussion that the high-dimensional feature space becomes complex when used for pattern recognition, and in particular, face recognition. For this purpose, convert the image into its low-dimensional feature space, so that the normalized features represent the face. Even if the dimension reduction techniques are applied, the image still contains some high-value noisy pixels, which affect the overall impact of the recognition.

To overcome this weakness, the following steps are applied on the face images for converting a 2D image into a hexagonal image. The first step in this process involves following the mathematical formulation [26]:

$$f_h(x, y) = \sum_m^{M-1} \sum_{n=0}^{N-1} f_s(m, n)h(x - m, y - n), \quad (21)$$

where an original image $f_s(m, n)$ is a square lattice with $M \times N$ points, $f_h(x, y)$ is the desired hexagonal image, and h is the interpolation kernel. For the reconstruction method, (m, n) and (x, y) are the sample elements of the original square and converted hexagonal images, respectively. This can be seen in Figure 2.

These boundary points, extracted from the reshaped image, are mapped to 2D images for the final elliptical face boundary. Consider the set X that contains the elliptical faces, also known as the mask-based face (MBF).

$$X = \{e_1, e_2, e_3, \dots, e_n\} \quad (22)$$

These faces are fed into the LDA routine discussed earlier for extracting the global features. In this procedure, only those pixels are considered that lie within the boundary of the elliptical face, while the rest of the pixels are marked as zeros. Each face e_i of the i th class contains the hybrid local and some global features that preserve the maximum discriminating information. To check the robustness of the method, the Euclidean distance of the intermediate frames and original image is taken into account to judge the projection of face patterns in low-dimensional space, with and without the MBF approach.

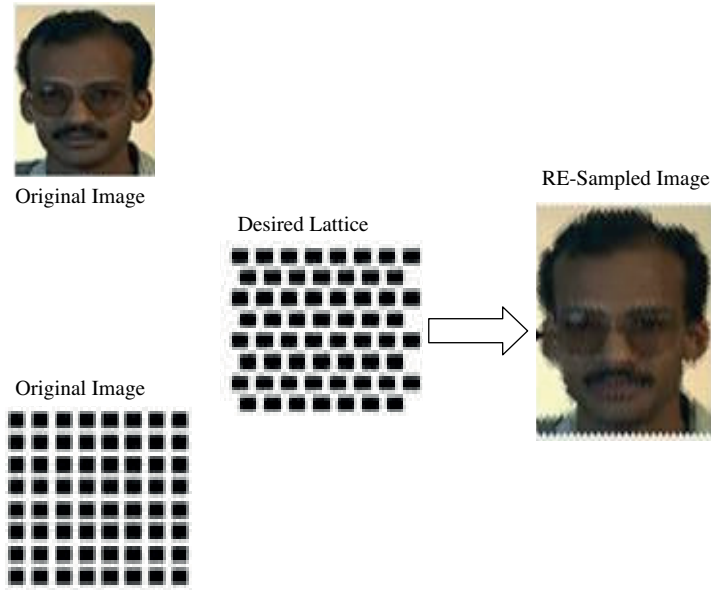


Figure 2. Process of converting a square image into a hexagonal image.

2.3.2.2. Extraction of local facial features

For extracting the local features, use $f_h(x, y)$ images as discussed previously, so as to mark the prominent local features. These marked features are mapped to their corresponding 2D images. Suppose that Lf_N is the set of features containing N samples, i.e.

$$Lf_N = \{Lf_1, Lf_2, Lf_3, \dots, Lf_N\}. \tag{23}$$

Apply the proposed LDA approach for obtaining the discriminating facial features. Not all of the features provide good impact on the recognition process. Hence, some kind of threshold t value should be defined within which most of the reliable information is preserved and provides a prominent effect on the recognition process. This threshold is extracted after number of random estimations that are helpful to select the best features. The mathematical equation for determining the most discriminating features is computed as follows:

$$MDF_N = \sum_{i=1}^N Lf_i > t, \tag{24}$$

where Lf_i is the sample feature vectors that contribute to the local features obtained earlier.

2.3.2.3. Combined feature space

In order to train the interperson and original face, both the local and global facial features must be combined. The local and global features are extracted via the aforementioned techniques. To combine these features, the weighted sum rule is used for the projection of the features into the combined subspace that converts N dimensional subspace into n dimensions, where $n \ll N$ projection features are treated as weights. Here, $m + 1$ projection vectors are considered as the weighted sum for both the local and global features in a combined subspace. Finally, these projection vectors are projected using the linear projection method.

The main objective when combining the interperson with the original face is that the isolated interperson has lost some discriminated information. When the original face is combined with the interperson frame, the

loss of the discriminate information is minimized. The combined face is projected into the linear subspace by preserving its most discriminate information.

3. Experiments and results

In this section, the experimental results are illustrated while evaluating the proposed system for face recognition. In this paper, 2 techniques are used to evaluate the experimental results on 4 different databases. The first technique is applied to the CMU-PIE and FERET databases, in which the pose subsets of both databases are selected for the experimental evaluation. From the CMU-PIE, 68 subjects under varying poses and neutral expressions are considered. The estimated pose difference varies from 0° to $\pm 90^\circ$ between the images of an individual. On the other hand, the FERET set contains 200 subjects with 9 poses that vary from 0° frontal to $\pm 60^\circ$ profile view, where every image has a difference of 15° . The second technique involves the analysis of the proposed system by reducing the size of the image dimensions on both the Olivetti Research Lab (ORL) and YALE B databases. The detailed experimental procedure is given below.

3.1. Experiments using the combined local and global facial features

3.1.1. CMU-PIE dataset

The proposed system is evaluated in 3 different ways. For the experimental results, we first took the CMU-PIE dataset and performed experiments in many ways. From the CMU-PIE, 68 subjects under varying poses and neutral expressions are considered. The estimated pose difference varies from 0° to $\pm 90^\circ$ between the images of an individual. At the end, the final experimental results are drawn on the basis of a comparison with the existing approaches.

Experiment 1 In the first experiment, randomly, 2 poses (between 0 and ± 90) and 1 normal image of an individual are selected as the probe set from the CMU-PIE database. Other images of the face database are selected as the probe set. In the next step, generate intermediate frames from 2-pose faces (left and right), with regard to the normal face of an individual. A total of 10 frames are generated: 5 frames from the 1st pose and 5 from the 2nd. These frames (5 from the left and 5 from the right) are those that are more varied than the previous. Hence, 13 faces are generated (3 from the input images and 10 from the virtual faces) using the above-mentioned morphing technique. There are $13 \times 68 = 884$ virtual and actual faces used for the training. The robustness of our method is based on the minimization of the 3S problem. The discriminate information from the pose face is reduced using the global and local features and the virtual faces.

Experiment 2 In the second experiment on the CMU-PIE database, select 4 random faces of an individual with a different pose variation and normal image as the base image for the interpersonal frames. In this experiment, we generate 15 virtual frames from the base face. There are 18 virtual and original faces constructed. Hence, a total of $19 \times 68 = 1292$ faces are generated from 4 individuals' faces. The training process is performed in an incremental way. Each iteration contains 19 sets of images for each individual. The incremental process of the LDA is discussed earlier. Each group (19 subjects) is treated as a single class and is assigned a single label.

Experiment 3 Third experiment on the CMU-PIE database uses 5 randomly selected images per person. The 5 sets of images are divided as 4 images from different poses faces and 1 normal, i.e. as a base face. Now, generate 5 different virtual faces according to the abovementioned method. In this experiment, 25 images of an individual group (virtual faces and original faces) are used. The training process is also used as that defined in the second experiment. The recognition rate in the third experiment shows very good results.

The comparison in Table 1 defines the above-mentioned experiments on the CMU-PIE database. The comparison is performed through extracting the recognition rates that are defined by different authors in their research work on pose variation.

Table 1. CMU-PIE database comparison results in different pose variations.

Pose difference	Exp. 1	Exp. 2	Exp. 3
23°	85%	91.5%	97.9%
45°	80.1%	84%	91.5%
Random selection with different orientation	78.5%	81.9%	87.9%

3.1.2. FERET datasets

For evaluating the proposed scheme, the same strategy is applied as discussed in the CMU-PIE dataset. In the FERET dataset, 200 subjects with 9 poses that vary from 0° frontal to $\pm 60^\circ$ profile view were selected, where every image has a difference of 15°.

Experiment 1 In the first experiment, 4 images out of 10 faces are selected for training, which contains 3 pose faces and a natural face. Generate the 5 interperson frames, with regard to the natural face per pose. For the training, 18 frames are used per individual. As we know, the FERET database is naturally complex, and so it takes some extra time for the training of the total number of $200 \times 18 = 3600$ virtual and original faces. Another 6 images are used as the probe images. After testing, the system's outcome produces outperformed results in the combined global and local features of the individuals. The same training process is used as mentioned in the CMU-PIE database experiments.

Experiment 2 In the second experiment, half of the faces are taken from the problem set. The probe set includes 4 random left or right pose variation faces per person and 1 normal image. The same training process is applied as discussed in the previous experiments. The more the training samples are used, the more the recognition rate will be.

Experiment 3 In the third experiment, 60% of the images are selected as the probe set for training the different pose faces. As far as the construction of the virtual faces and incremental training process are concerned, the same strategy is used as discussed in the aforementioned experiments. The comparison and recognition rate is shown in Table 2.

Table 2. FERET database comparison results in different pose variations.

Pose difference	Exp. 1	Exp. 2	Exp. 3
15°	85.5%	89%	99.5%
45°	83%	87.1%	95.1%
Random selection with different orientation	80.5%	86.5%	94.9%

3.1.3. Comparison

A comparison is made using the proposed method with the other pose invariant face recognition methods proposed by different authors. The comparison is performed on 2 databases, namely the CMU-PIE and FERET databases. The comparison cannot be made directly because every author has a different architecture of the pose-based face recognition. For training a database, authors use a different number of faces per individual to train a database for recognition purposes. Table 3 shows the comparison of the proposed method with other

methods that use minimum to maximum faces per individual. In this paper, the comparison is performed on the average recognition rates defined by different authors.

Table 3. Comparison of the different pose invariant face recognition rates on the CMU-PIE and FERET databases.

Pose diff. Methods	CMU-PIE Average pose diff.	FERET Average pose diff.
Method in ... [27]	78.5	75
Method in ... [28]	98.5/89.7	-
Method in ... [29]	100/91	100/99
Method in ... [30]	91.5/87.9/81	100/90.3/82.5
Proposed method	97.5/91.5/87.9	99.5/95.1/94.9

3.2. Experiment using the reducing size of the image

Experiments are also performed with different image sizes to check the robustness of the system. This check is performed to determine the effectiveness of the presented work when varying image sizes are available. The databases under consideration are the ORL and YALE B databases, as discussed in the previous section. Table 4 below shows both databases' resolutions used to obtain the results.

Table 4. ORL and YALE B databases' resolution used for recognition.

S#	Image resolution	
	ORL	YALE B
1	112 × 92	168 × 192
2	56 × 56	125 × 125
3	37 × 23	100 × 100

The recognition rates of the proposed framework (FAPSP) are compared with standard LDA and PCA algorithms, which were rewritten by us. In addition, for training purposes, the same number of sample sizes was chosen for the LDA, PCA, and the proposed method, i.e. 3 images per person, as discussed below.

3.2.1. Experiments on the ORL database

The ORL database holds 400 images of 40 discrete subjects having 10 samples each. The images contain different lighting conditions, occlusions, and facial expressions, with a dark setting environment. In this section, 3 images are selected out of 400 different individuals for each resolution image. Out of these 3 images, 1 image is a left pose, the 2nd is a right pose, and the 3rd one is a frontal image. The left and right images generate 10 intermediate frames with respect to a normal face. As a consequence, 13 images are generated by each person. For the training phase, 2 factors are important: the 1st is the number of subjects for the training and 2nd is the size of the image. For an analysis that exposes a higher number of samples per person, the more the recognition rates will be. On the other hand, the resolution plays an important role. A decrease in the image resolution causes the recognition rate to minimize. The main reason for the decrease in the recognition rate is due to the decreasing the size of the image. The decrease of the image size causes the pixel values to be destroyed, due to which, the similarity between 2 subjects is scattered. The experiment results are taken on 112 × 92 dimensions and shown in Table 5.

The analysis shows that 3% to 4% recognition rates are decreased between the LDA and PCA, but the proposed FAPSP recognition rates are greater than 94% in all of the subjects at a 112 × 92 resolution.

Table 5. Recognition rates on the ORL database at a resolution of 112×92 .

No. of subjects	Image resolution 112×92		
	LDA	PCA	FAPSP
10	97%	95%	99%
20	92%	87%	99%
30	89%	88%	95%
40	84%	82%	94%

The original ORL database resolution compresses into 56×56 using the MATLAB function. To check the recognition rates, a comparison is given in Table 6.

Table 6. Recognition rates on the ORL database at a resolution of 56×56 .

No. of subjects	Image resolution 56×56		
	LDA	PCA	FAPSP
10	93%	90%	96%
20	86%	81%	93%
30	80%	77%	89%
40	74%	69%	86%

As discussed above, the LDA, PCA, and FAPSP depend on global facial features, and so by decreasing the resolution, the recognition results are going to be decreased.

The next analysis is performed at a very low resolution, 37×23 , converted by the MATLAB function. The comparison is shown in Table 7.

Table 7. Recognition rates on the ORL database at a resolution of 37×23 .

No. of subjects	Image resolution 37×23		
	LDA	PCA	FAPSP
10	88%	85%	92%
20	83%	81%	87%
30	81%	79%	84%
40	77%	73%	81%

3.2.2. Experiment on the YALE B database

There are 73 images per subject, having 9 different poses and 64 illuminated images, in which 3 images are randomly selected from 9 pose images. Out of 3 images, 1 image is the neutral face, while the remaining 2 images are the left and right poses, respectively. From these pose images, 13 frames are generated, in which 5 frames are generated from the left pose and 5 are from the right pose. The dimensions of the images are reduced to 168×192 , as provided in Table 8.

Table 8. Recognition rates on the YALE B database at a resolution of 168×192 .

No. of subjects	Image resolution 168×192		
	LDA	PCA	FAPSP
13	85%	79%	92%
26	79%	73%	90%
39	74%	67%	84%

This experiment is performed on its original size. The next experiment is performed in the same way, as discussed in Table 8, but the results shown in Table 9 were taken at a dimension of 125×125 .

Table 9. Recognition rates on the YALE B database at a resolution of 125×125 .

No. of subjects	Image resolution 125×125		
	LDA	PCA	FAPSP
13	82%	77%	90%
26	74%	71%	89%
39	70%	65%	82%

The third experiment, the results are taken at a dimension of 100×100 and shown in Table 10. The subjects are used with the difference of 13, i.e. first, 13 images are trained, and then 26, and 39, and so on.

Table 10. Recognition rates on the YALE B database at a resolution of 100×100 .

No. of subjects	Image resolution 100×100		
	LDA	PCA	FAPSP
13	79%	75%	88%
26	73%	68%	84%
39	68%	62%	79%

The results are taken in 2 ways that reveal that the overall system performs well compared to other systems. Initially, the results are drawn on a small training dataset. Next, the system randomly chooses pose images and creates intermediate frames, due to which the number of images per person increases. By increasing the number of images, the recognition rate increases. To analyze the system, the second approach used in this paper is based on reducing the size of the images. The dimensionality reductions of the images damage the pixels, due to which the recognition rate decreases.

Note: Different papers show different results using the PCA and LDA [31–35], which are higher than the results shown in this paper. The reason for the higher results is that these papers used a minimum of 5 training samples per person and they presented the general recognition results, and not specifically for the pose variation, while in this paper, we have used 3 samples per person and handled the pose variation and 3S problem. The reason for selecting a lower number of samples is to save storage and processing costs and handling the pose variation leads to more efficient face recognition results compared to getting the results generally from the face. Therefore, our approach is far better than the existing approaches in terms of storage, processing, and recognition results. In addition, we have rewritten our own programs for the LDA and PCA using the standard algorithms and gathered the results using them. We then compared the results of these programs with the results of our proposed work.

4. Conclusion

In this paper, 2 contributions were made. First, the 3S problem is handled by making a small contribution in the LDA. Second, the proposed framework attempts to solve the pose variation problem. To deal with the singularity of the LDA, an alteration is made in the diagonal entries of the within-class scatter matrix. This is later used for modifying the Fisher's criterion equation. This framework is then applied to a face recognition scenario for dealing with the pose variation problem. The uniqueness of this paper is based on the use of mesh morphing for computing the intermediate frames and the combined subspace of both local and global facial

features. The experimental results reveal that the anticipated framework shows fine performance as far as the pose variation and 3S problem are concerned.

References

- [1] W.C. Zhao, R. Phillips, A.P.J. Rosenfeld, ‘Face recognition: a literature survey’, *ACM Computing Surveys (CSUR)*, Vol. 35, pp. 399–458, 2003.
- [2] X. Zou, J. Kittler, K. Messer, ‘Illumination invariant face recognition: a survey’, *1st IEEE International Conference on Biometrics: Theory, Applications, and Systems*, pp. 1–8, 2007.
- [3] X. Tan, S. Chen, Z.H. Zhou, J. Liu, ‘Face recognition under occlusions and variant expressions with partial similarity’. *IEEE Transactions on Information Forensics and Security*, Vol. 4, pp. 217–230, 2009.
- [4] X. Zhang, Y. Gao, ‘Face recognition across pose: a review’, *Pattern Recognition*, Vol. 42, pp. 2876–2896, 2009.
- [5] B. Fasel, J. Luetttin, ‘Automatic facial expression analysis: a survey’, *Pattern Recognition*, Vol. 36, pp. 259–275, 2003.
- [6] M. Kirby, L. Sirovich, ‘Application of the Karhunen-Loeve procedure for the characterization of human faces’, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 12, pp. 103–108, 1990.
- [7] H. Zhao, P.C. Yuen, J.T. Kwok, ‘A novel incremental principal component analysis and its application for face recognition’, *IEEE Transactions on Systems Man, and Cybernetics, Part B: Cybernetics*, Vol. 36, pp. 873–886, 2006.
- [8] P.N. Belhumeur, J.P. Hespanha, D.J. Kriegman, ‘Eigenfaces vs. fisherfaces: Recognition using class specific linear projection’, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 19, pp. 711–720, 1997.
- [9] H. Zhao, P.C. Yuen, ‘Incremental linear discriminant analysis for face recognition’, *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, Vol. 38, pp. 210–221, 2008.
- [10] M.E. Wibowo, D. Tjondronegoro, ‘Face recognition across pose on video using eigen light-fields’, *Proceedings of the International Conference on Digital Image Computing: Techniques and Applications*, pp. 536–541, 2011.
- [11] L. Wiskott, J.M. Fellous, N. Kuiger, C. von der Malsburg, ‘Face recognition by elastic bunch graph matching’, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 19, pp. 775–779, 1997.
- [12] A. Lanitis, C.J. Taylor, T.F. Cootes, ‘Automatic face identification system using flexible appearance models’, *Image and Vision Computing*, Vol. 13, pp. 393–401, 1995.
- [13] T.F. Cootes, G.J. Edwards, C.J. Taylor, ‘Active appearance models’, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 23, pp. 681–685, 2001.
- [14] D. Mazumdar, K. Chanda, M. Bhattacharya, S. Mitra, ‘A pose recovery algorithm to minimize the effects of pose variation on the matching performance of elastic bunch graph based multiview face recognition system’, *International Conference on Computing: Theory and Applications*, pp. 377–381, 2007.
- [15] T. Gernoth, A. Gooßen, R.R. Grigat, ‘Face recognition under pose variations using shape-adapted texture features’, *17th IEEE International Conference on Image Processing*, pp. 4525–4528, 2010.
- [16] F. Yang, A. Krzyzak, ‘Face recognition under significant pose variation’, *Canadian Conference on Electrical and Computer Engineering*, pp. 1313–1316, 2007.
- [17] J. Hu, W. Deng, J. Guo, ‘2D projective transformation based active shape model for facial feature location’, *18th International Conference on Fuzzy Systems and Knowledge Discovery*, Vol. 4, pp. 2442–2446, 2011.
- [18] T.F. Cootes, G.V. Wheeler, K.N. Walker, C.J. Taylor, ‘View-based active appearance models’, *Image and Vision Computing*, Vol. 20, pp. 657–664, 2002.
- [19] N.S. Vu, A. Caplier, ‘Efficient statistical face recognition across pose using local binary patterns and Gabor wavelets’, *IEEE 3rd International Conference on Biometrics: Theory, Applications, and Systems*, pp. 1–5, 2009.

- [20] D. González-Jiménez, J.L. Alba-Castro, “Toward pose-invariant 2-D face recognition through point distribution models and facial symmetry”, *IEEE Transactions on Information Forensics and Security*, Vol. 2, pp. 413–429, 2007.
- [21] J. Heo, “Generic elastic models for 2D pose synthesis and face recognition”, PhD Dissertation, CMU, 2009.
- [22] U.H. Prabhu, J.M. Savvides, “Unconstrained pose invariant face recognition using 3D generic elastic models”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 1–1, 2011.
- [23] M. Alexa, “Local control for mesh morphing”, *Proceedings of the International Conference on Shape Modeling & Applications*, pp. 209–215, 2001.
- [24] D.T. Sandwell, “Biharmonic spline interpolation of GEOS-3 and SEASAT altimeter data”, *Geophysical Research Letters*, Vol. 14, pp. 139–142, 1987.
- [25] R.C. Gonzalez, R.E. Woods, S.L. Eddins, *Digital Image Processing Using MATLAB: and “Mathworks, MATLAB Sim SV 07”*, New Jersey, Prentice Hall, 2007.
- [26] L. Middleton, J. Sivaswamy, *Hexagonal Image Processing: A Practical Approach*, New York, Springer-Verlag , 2005.
- [27] R. Gross, I. Matthews, S. Baker, “Appearance-based face recognition and light-fields”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 26, pp. 449–465, 2004.
- [28] X. Chai, S. Shan, X. Chen, W. Gao, “Locally linear regression for pose-invariant face recognition”, *IEEE Transactions on Image Processing*, Vol. 16, pp. 1716–1725, 2007.
- [29] S.J.D. Prince, J. Warrell, J. Elder, F.M. Felisberti, “Tied factor analysis for face recognition across large pose differences”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 30, pp. 970–984, 2008.
- [30] M.S. Sarfraz, O. Hellwich, “Statistical appearance models for automatic pose invariant face recognition”, *IEEE Conference Proceedings*, pp. 1–6, 2008.
- [31] K. Tan, S. Chen, “Adaptively weighted sub-pattern PCA for face recognition”, *Neurocomputing*, Vol. 64, pp. 505–511, 2005.
- [32] H. Yin, P. Fu, S. Meng, “Face recognition based on adaptively weighted fisherface”, *Journal of Optoelectronics Laser*, Vol. 17, pp. 1405, 2006.
- [33] J. Lu, K.N. Plataniotis, A.N. Venetsanopoulos, “Face recognition using LDA-based algorithms”, *IEEE Transactions on Neural Networks*, Vol. 14, pp. 195–200, 2003.
- [34] J. Zhang, X. Fei, Y. Zhang, “A new method for face recognition based on PCA optimize strategy”, *IEEE Conference Proceedings*, Vol. 10, pp. 417–420, 2010.
- [35] W. Zuo, K. Wang, D. Zhang, H. Zhang, “Combination of two novel LDA-based methods for face recognition”, *Neurocomputing*, Vol. 70, pp. 735–742, 2007.