Multimedia Tools and Applications
https://doi.org/10.1007/s11042-020-08892-9

Akin-based Orthogonal Space (AOS): a subspace learning method for face recognition

Anu Singha[1] · Mrinal Kanti Bhowmik[1] · Debotosh Bhattacherjee[2]

Received: 1 September 2018 / Revised: 14 March 2020 / Accepted: 27 March 2020 /
Published online: 11 May 2020

Abstract
A projection learning space is an approach to mapping a high-dimensional vector space to a lower dimensional vector space. In this paper, we proposed an algorithm, namely, AOS: Akin based Orthogonal Space. The algorithm is driven with two major targets - (i) to choose most representative image(s) from a group of face images of an individual, (ii) finally to produce a learning space which follows a Gaussian distribution to reduce the influence of grosses like non-Gaussianly distributed data noises, variations in facial expression and illumination. To improve the recognition performance, we proposed another approach i.e. fusion between AOS features and a custom VGG features. We justify the effectiveness of the proposed approaches over five benchmark face datasets using two classifiers. Experimental results show that the proposed learning algorithm has obtained maximum of 92.22% recognition rate, as well deep learning based fusion approch greatly improves the recognition accuracy. The comparative performances demonstrate that the proposed method could significantly outperform other relevant subspace learning methods.

Keywords Subspace · Orthogonalization · Akinity · Liability · Representative score · Custom VGG

1 Introduction

In the face analysis, we often assume that the data is noise-free; in practical scenarios the image data are polluted by gross non-Gaussian noise and corruptions like illumination,

✉ Anu Singha
anusingh5012@gmail.com

Mrinal Kanti Bhowmik
mrinalkantibhowmik@tripurauniv.in

Debotosh Bhattacherjee
debotosh@ieee.org

[1] Department of Computer Science and Engineering, Tripura University, Suryamaninagar, Agartala, India

[2] Department of Computer Science and Engineering, Jadhavpur University, Kolkata, India

expression and pose variations. To assuage these drawbacks and being motivated by J. Brendan et al. [12], we have proposed a new subspace learning method in order to choose representative face image(s) to reduce dimensionality, and if such space follows Gaussian distribution, it will reduce the aforementioned grosses. At this state, a question may arise: why another subspace method is there? All the state-of-the-art methods are basically based on optimizations, eigenvector evaluations which are highly intricate. Instead, our proposed method is ease of implementation, easily understandable, there is no complicated mathematical formulation like eigenvector decomposition or independent vector optimization.

The common subspace learning methods are mostly based on similarities [7, 8, 10, 11, 20], correlations [13, 21, 22], eigenvector decompositions [23, 42], laplacianfaces [6], orthogonal basis vectors [18, 19], independent basis vectors and optimization [1–3] of signals separation. The classic subspace learning method, Trunk et al. [37, 38] proposed an unsupervised method i.e. Principal Component Analysis (PCA) for face recognition, which uses total scatter matrix to obtain maximum variance among all the data. PCA eigenface model can be resolved using Singular Value Decomposition (SVD) [23] of the covariance matrix, where PCA is based on eigen values and SVD is based on singular values. Modified extended version of PCA which known as Incremental PCA (IPCA) [9, 15, 25, 33] and Candid Covariance-free IPCA (CCIPCA) [43], all are allowed to add upcoming images and update the PCA representation accordingly. These methods estimate the principle components of a sequence of samples incrementally without computing the covariance matrix. Independent Component Analysis (ICA) [1], a generalized version of PCA, is first applied to face recognition by Bartlet et al. [3]. It is concerned with second-order dependencies as well as highorder dependencies between variables. The authors also suggested two ICA architectures, namely, ICA Architecure I and ICA Architecure II. In last decades, J. Yang et al. [48] re-evaluate the performance of two ICA architectures using standard PCA, and find that ICA Architecture I involves a vertically centered PCA process (PCA I), while ICA Architeture II involves a whitened horizontally centered PCA process (PCA II).

Similarity or correlation analysis has also been an attractive research domain in the field of image analysis and pattern recognition. An unsupervised subspace learning method-Canonical Correlation Analysis (CCA) [26] is based on similarity metric which can simultaneously deal with two sets of data, in contrast to PCA and other subspace methods. A general framework using the linear approximation form for the non-linear methods in the mode of graph based subspace learning, such as Linear Graph Embedding (LGE), Locality Preserving Projection (LPP) [16, 27] has many applications due to its effectiveness and simpleness. The LPP is a low-dimensional embedded space where the locality relationship of samples can be preserved. Further, LGE and LPP are non-orthogonal and this makes them difficult to reconstruct the data. Their counterparts naming Orthogonal Linear Graph Embedding (OLGE) and Orthogonal Locality Preserving Projection (OLPP) [6] extract features from the original data set by solving common eigenvalue problem. These orthogonal basis have more discriminating power than their base methods.

The contributions of this paper are as follows:

1) An algorithm called ‘Akin based Orthogonal Space (AOS)’ has proposed with the following intention: (i) to generate a list of representative face subspaces (RFS). Each subspace is encoded by choosing the representative face(s) from a group of face images of an individual; (ii) to make non-orthogonal ‘RFS’ subspaces to an orthogonal space.

2) To improve the performances with the help of proposed learning space, we adapt another approach i.e. fusion approach of proposed method with a custom VGG convolutional neural network model.
3) To establish the performances of proposed approaches over other existing state-of-the-art dimension reduction subspaces, numerous experiments and comparative studies have been conducted over IRIS, USTC-NVIE, ORL, YALE, and FRGC for face recognition over grosses like expressions and illuminations.

The remainder of the paper is structured as follows: Section 2 contains related literature works, Section 3 describes the theoretical background and algorithmic steps of proposed algorithm, and Section 4 presents a fusion approach with convolutional neural network. Section 5 evaluates and analyses the algorithm over five face-image databases and compares their performances. Section 6 offers conclusions and outlines future work.

2 Related works

Most common approach towards face recognition problem is done through dimension reduction by Eigen spaces [1–3, 6, 9, 18, 19, 23, 25, 37, 38]. Face recognition without Eigen spaces, namely, pairwise face recognition, has been extensively studied during past decay [7, 8, 10, 11, 13, 20–22]. Nevertheless, it is still an on-going in terms of state-of-the-arts complex nature as well as several practical factors such as illumination, pose, expression, and so on. Cai et al. [6] proposed a face recognition linear algorithm 'orthogonal laplacian faces' which can also be performed in reproducing kernel Hilbert non-linear space. Here the space has assumed to be connected but the real world data could be disconnected and thus different components have different dimensionality. Kim et al. [19] also worked on orthogonal subspace which is an on-line learning update procedure. One of very well-known classic Eigen space methods introduced by Trunk et. al. [37] performs with high accuracy although substantially its rejection rate is 'unknown', and not robust to changes in lighting, head size, and orientation. An extension robust incremental learning version of PCA [33] which adopts to input images as they arrive, is able to treat each pixel and each image differently. Nonetheless the re-learning for more recent images has not been investigated. Another method proposed by Weng et al. [43] also estimates principle components from incoming data without computing the covariance matrix. Recently, Tang et al. [34] proposed a regression algorithm in order to solve the problem of insufficient utilization of the sample information. Tang et al. [35] proposed another subspace classification for face recognition that can make the sample points to be linear separable by utilizing the kernel function. But it has own weakness that it cannot classify the sample points that distribute around the intersection very well. In the concern of high-order dependencies between variables, Bartlet et al. [3] introduced ICA. To reduce overlapping probability densities of the composite variables, Bhowmik et al. [5] proposed a next generation variant of ICA. The weakness of this variant is that it highlight the enhanced result in log-ICA II only.

Facial similarity makes classification difficult due to its non-rigid nature of human faces. There are methods to the cases that only consider pairwise similarity or correlations between samples available in literature. Gomez et al. [11] aimed at mimicking the human ability to differentiate people that proposed algorithm composed of different one-dimensional projection lines. Whereas Cao et al. [7] proposed a new approach to learning the embedding of faces in a gallery with human perception of facial similarity. Due to annotation cost, they experimented on a small set of the face dataset. For high-dimensional data, Klami et al. [22]

introduced a novel model which not only extracts the statistical correlations between data sets but also decomposes the data into shared and data set-specific components. While the model seems more complex due to having more unknown latent variables, it has advantage of diagonal noise that reduces the risk of over-fitting.

All the aforementioned methods are based on complex nature such as estimation of Eigen vectors, optimization and so on. Besides, if we consider the large number of images per person, it will also a space consuming. Therefore, our motive is to present a simple and effective subspace learning method.

3 The proposed learning space

Following the motivational exertion of Brendan [12], we tried a different way of subspace or dimension reduction solution based on a method *Akin*. *Akin* is called as someone to whom you are comparable by similarity or blood, for e.g. brothers are a case of people who are akin. Earlier, Xin et al. [46] presented an exertion on kin relationships, where authors depicted the family faces like son-father, son-mother, daughter-father, daughter mother, and so on. In our case, *Akin* indicates a face which has the maximum analogous characteristics in a group of face images of an individual. In Fig. 1, we show how *Akin* based graph can be used to extract representative faces (the highlighted face images in the figure) from a gallery of images. The representative face images are those images which scored the maximum similar values. We recovered the representative images from each group of face images of an individual in an iterative way. In other words, if we have 3 individuals as shown in Fig. 1, and if each individual has 7 images i.e. a total of 21 images, then there will be 3 iterations. There may be more than one representative images with similar score values from each group. To generate a subspace for each group, we concatenates the representative face image(s). These subspaces should be orthogonal or de-correlated because they are representative subspaces from separate group of individuals. At the end, we merge up these subspaces to produce an orthogonal space, which follows a Gaussian distribution. Figure 2 shows a brief overview of our proposed subspace learning system.

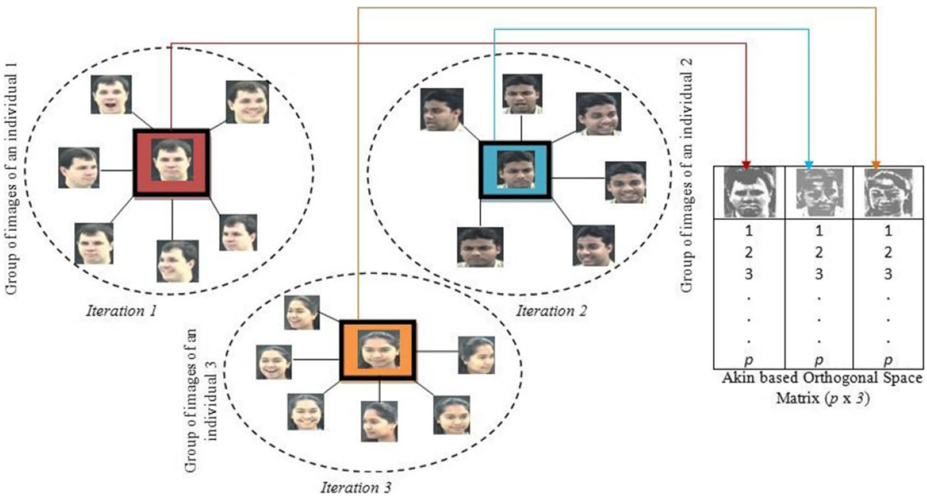

Fig. 1 Motivational graph of our proposed system

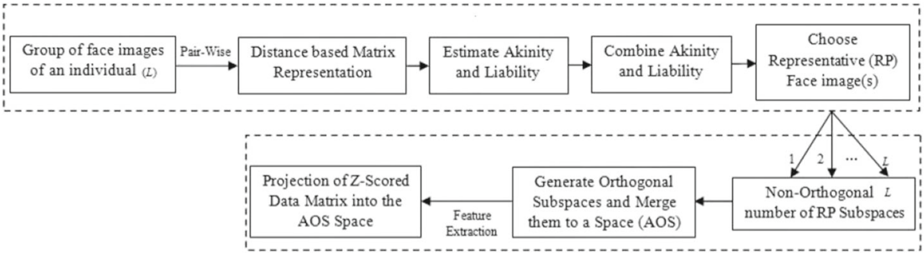

Fig. 2 The proposed system overview

The proposed idea is articulated through an algorithm called akin based orthogonal space (AOS) with two motives: (i) to create a list of representative face subspaces (RFS) in order to choose the representative face(s) from each group of face images of individuals, and (ii) to make non-orthogonal 'RFS' subspaces to an orthogonal space.

For a clear illustration of how to pull out representative face image(s) from each group, and how to orthogonalize the representative subspaces, we present a detailed procedure in Algorithm 1 and discuss step-wise afterwards.

Algorithm 1 Akin based Orthogonal Space (AOS).

Input: A set of gallery images divided into L groups and stored in a data matrix, $RP^L = [x_1^1, x_2^1, ..., x_k^1] \in \Re^{p\times k}$. Each group is a set of k images of an individual consists of p number of real valued pixels.
R = Number of representative images ($R \leq k$)

Output: We can find subspaces $[RP_o^1, RP_o^2, ..., RP_o^L] \in \Re^{p\times R}$ which are mutually orthogonal and have the same span. And finally a space $S \in \Re^{p\times N}$ which is concatenation of the entire representative orthogonal subspaces that generates a total span of N representative faces of p pixels

for $g \leftarrow 1\ to\ L$ **do**
- Creation of a distance based matrix, and convert the distance matrix to a similarity measure normalized matrix
- Estimation of Akinity, initialize $a(i, j) \leftarrow 0$
- Estimation of Liability, initialize $l(i, j) \leftarrow 0$
- Calculate the Representative (RP) Score, initialize $c(i, j) \leftarrow 0$

$$c(i, j) = a(i, j) + l(i, j)$$

$$RPScore = sort(diag(c))$$

- Choose the Representative (RP) Images, r
- **for** $r \leftarrow 1\ to\ R$ **do**

$$RP^g(:, r) = X^g(:, RPScore(r))$$

Orthogonalize the Representative Face Subspaces by initialization $RP_o^1 = RP^1$

Concatenation of generated orthogonal subspaces as $S = [RP_o^1 \bigcup RP_o^2 \bigcup ... \bigcup RP_o^L]$

Input and Output: A set of input gallery images divided into L groups. Each group is consisting of k images of an individual. Initially, it is necessary to provide the number of representative images R as user choice. Finally, the algorithm will give an orthogonal space S which is a union of L number of mutually orthogonal subspaces RP_o^1, RP_o^2, ..., RP_o^L.

Distance based Matrix Creation (Step 2): Each group of k images is stored in a data matrix X^L where each column representing a face image vector. In an iterative way, the algorithm will run for L groups as $g \leftarrow 1$ *to* L.

The input distance based matrix of the k images is structured as follows: each face image is taken as data point over row axis and as akin point over column axis, as shown in Fig. 3. The distance matrix gets as input a gathering of real valued Euclidean distances between data points x_i and akin points x_j. Hence, each data point x_i or akin point x_j is represented by a face image vector. The Euclidean norm, denoted as $||.||$, estimates a relationship between face vector pairs x_i and x_j as

$$Dis(i, j) = EuclideanDistance||x_i - x_j|| \tag{1}$$

where *Dis(i, j)* indicates how much appropriate the akin point j is suited for data point i. Here, i and j are the row and column dimensions of k images of an individual at a time. A linear normalization of an image based distance matrix is executed to eliminate the effect of certain gross influences, in a way of comparing corresponding distance values between image pairs or data and akin point pairs, as in (2).

$$DSim(i, j) = 1 - \frac{Dis(i, j) - \min Dis(:, :)}{\max Dis(:, :) - \min Dis(:, :)} \tag{2}$$

notation ':,:' indicates all data values of corresponding distance matrix rows and columns. The '1 minus' indicates of converting the distance to similarity measure i.e. maximum similarity towards 1 and minimum similarity towards 0.

Estimation of Akinity (Step 3): The proposed algorithm mainly conceptualized on two theoretical aspects: *Akinity* [31, 32] and *Liability* which are attempted to establish, here. As mentioned earlier, *Akin* indicates a most appropriate similar face which has maximum analogous characteristics in a group of face images of an individual. While penetrating a candidate akin point for a data point, we taking into account other candidate akin points as competitor - which has termed as *'Akinity'*. Figure 4 shows the Akinity *a(i, j)* sent from data point i to candidate akin point j (black points) i.e. akin point j is to serve as the most similar candidate for data point i, while others candidate akin points j' (green points) will compete for data point i. The Akinity *a(i, j)* is set to the input similarity *DSim(i, j)* between data point i and akin point j. While measuring similarity between a candidate akin point for a data point, the rest of candidate akin points will consider as competitors for holding that data point. So the question is, how much superior the score of akin point than other

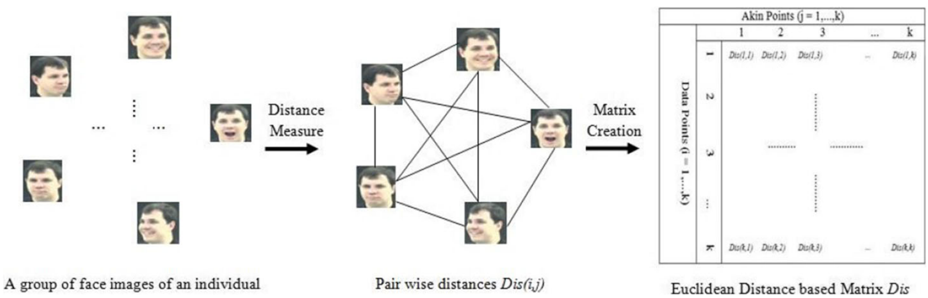

Fig. 3 Distance based matrix representation

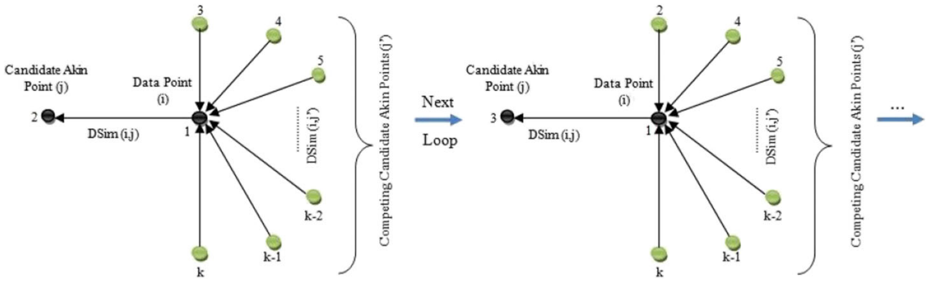

Fig. 4 Akinity *a(i,j)* Graph for data point *i* to akin point *j*

competing akin points? To answer this, we have subtracted the largest of the similarities among competing candidate akin points j' as

$$a(i,j) = DSim(i,j) - \max_{j' \neq (i,j)} DSim(i,j') \qquad \text{if i} \neq \text{j} \tag{3}$$

$$a(j,j) = \max_{j' \neq j} a(j,j') \qquad \text{if i=j} \tag{4}$$

For $i = j$, the Akinity $a(j, j)$ is set to maximum of akinity values from all the estimated values for $i \neq j$ defined as (4). That means, how appropriate it would be for data point i is chosen as an akin itself.

Since the akinity is estimated among a group of faces of same person, in some circumstances the values will be replicate, called oscillation of numerical values. It is important that they should be damped to avoid numerical oscillation. Each updated damped akinity value is set to λ times its previous value plus (1 − λ) times its current akinity value

$$a(i,j) = \lambda \times a(i,j-1) + (1-\lambda) \times a(i,j) \qquad \text{if j} \neq 1 \tag{5}$$

In our case, the damping factor (λ) value is 0.5. Then, a linear normalization has been performed. The idea behind akinity normalization is to bring the values into a range of 0 and 1, defined as (6).

$$a(i,j) = \frac{a(i,j) - \min a(:,:)}{\max a(:,:) - \min a(:,:)} \tag{6}$$

where notation ':,:' indicates all data values of corresponding akinity matrix rows and columns.

Estimation of Liability (Step 4): On the other hand, Liability indicates how much a chosen point is answerable. Figure 5 shows the liability *l(i, j)* sent from candidate akin point *j* to data point *i*. The earlier concept Akinity is concentrated on a data point at a time where the akin points are rival for occupancy to rein that data point. Through akinity when a akin point is decided for a data point, liability update the fact that data point *i* pick candidate *j* as the most appropriate akin point which has been decided by other supporting data points. It is set as the average sum of the akinity values received by candidate akin point *j* from other supporting data points i'

$$l(i,j) = \frac{1}{k-1} \sum_{i' \neq (i,j)} a(i',j) \qquad \text{if i} \neq \text{j} \tag{7}$$

$$l(j,j) = \max_{i' \neq i} l(i',j) \qquad \text{if i= j} \tag{8}$$

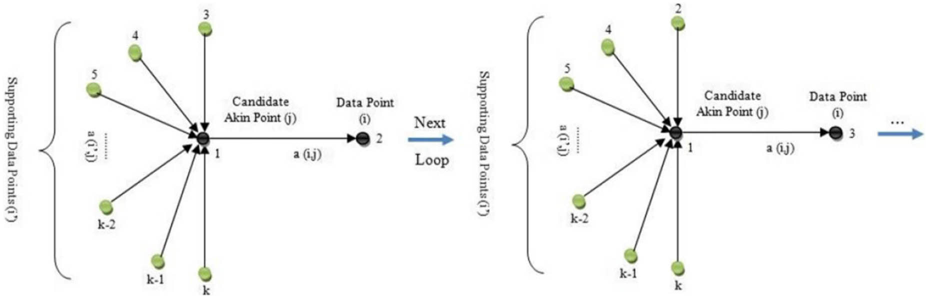

Fig. 5 Liability *l(i,j)* Graph for akin point *j* to data point *i*

Like akinity values, the liability values are damped as defined (9) and finally normalized.

$$l(i,j) = \lambda \times l(i-1,j) + (1-\lambda) \times l(i,j) \qquad \text{if i} \neq 1 \tag{9}$$

Generate Representative Face Subspaces (Steps 5-8): To make a choice the representative images, the akinity and liability values has been combined for diagonal positions and chosen the higher values in a variable *RPScore*.

$$c(i,j) = a(i,j) + l(i,j) \qquad \text{if i = j} \tag{10}$$

We can pick more than one i.e. *R* number of representative faces based on the next higher values of *RPScore*. Based on index of higher *RPScore* values and corresponding image vectors from data matrix X^g, the representative subspace or matrix will be generated (Step 8). After complete execution of Step 1 i.e. *L* number of iterations for *L* number of separate group of individuals, we can recover a list of representative face subspaces $RP^1 = [x_1^1, x_2^1, ..., x_R^1] \in \Re^{p \times R}$, $RP^2 = [x_1^2, x_2^2, ..., x_R^2] \in \Re^{p \times R}$, ..., $RP^L = [x_1^L, x_2^L, ..., x_R^L] \in \Re^{p \times R}$.

Generate an Orthogonal Space (Steps 9-10): Since the representative face subspaces are matrices from *L* number of separate group of individuals, it should be de-correlated or orthogonal from each other. For clear illustration of how to orthogonalize the subspaces, we present the detailed procedure in these steps with the help of Gram-Schmidt process [14]. The visual stages have been shown in Fig. 6

The basis subspace matrix (RP_o^1) will be deducted from the summation value of relationship with previous subspace matrices i.e. inner products between current subspace and previously formed orthogonal subspaces as defined in (11). The inner product denotes by

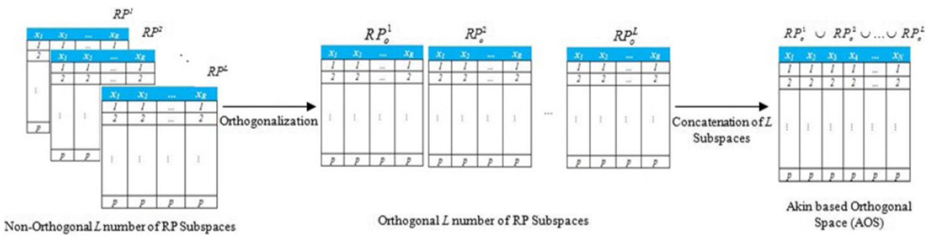

Fig. 6 Representative subspaces based orthogonal space creation

trace between two subspaces as $tr((RP^j)^T \times RP_o^i)$. The *tr* operator projects the subspace RP^j onto the line spanned by subspace RP_o^i to extract the association value between these two subspaces.Then the resultant traced scalar value will be normalized by $||RP^j||$, and finally project onto RP^j.

$$RP_o^i = RP_o^i - \sum_{j=1}^{i-1} \frac{tr((RP^j)^T \times RP_o^i)}{||RP^j||} RP^j; \textit{ if } i = 2 \textit{ to } L \tag{11}$$

$$RP_o^i = \frac{RP_o^i}{||RP_o^i||} \tag{12}$$

The sequences RP_o^i are required system of orthogonal subspaces, and the subspaces RP_o^i form an orthonormal set through normalizing i.e. dividing the subspaces itself by norm as defined in (12). At last, we generated a space S by concatenate these L number of representative subspaces, which has been termed as Akin based Orthogonal Space (AOS).

Proposition 1 *Linear orthonormal transformation of non-orthonormal representative subspaces $RP^L \in \Re^{p\times R}$ generate a symmetric matrix $M \in E^{L\times L}$ such that each element of M have densities of Gaussian.*

Proof According to AOS process, the orthonormal transformation of a representative subspaces RP^L linearly form a basis space $S \in \Re^{p\times N}$ where S consists of a set of L uncorrelated subspaces $RP_o^1, RP_o^2, ..., RP_o^L$ such that $tr((RP_o^i)^T \times RP_o^j) = 0$ if $i \neq j$ and $tr((RP_o^i)^T \times RP_o^j) = 0$ if $i = j$.

In Gaussian Orthogonal Ensemble (GOE), $distributed[M, GaussianOrthogonal MatrixDistribution(\sigma, L)]$ can be used to assert that the distribution of a random squared symmetric matrices $M = m_{ij}$ have probability densities proportional to a Gaussian distribution $(0, \sigma)$ [44]. It is evidence from the discussion that each element (m_{ij}) symmetric matrix M follows a distribution of Gaussian where ensuing of m_{ij} is coming from the product of two orthonormal basis subspaces RP_o^i and RP_o^j of S. □

3.1 Space projection to feature extraction

In feature extraction, the aim of the generated space, namely, AOS is based on the theory that the high-dimensionally observed faces will be lived in a low-dimensional space. But, our method of generating projection space is a making over the susceptible power of non-Gaussian, as argued in the *Proposition 1*. In order to extract the face features, faces data are needed to be projected into the generated space S. Earlier, it was also necessary that faces data should follow a standard form of signal close to Gaussian distribution. To standardize the faces data, we have utilized the concept of Z-Score. Z-scores are linearly transformed data values having a mean of zero and standard deviation of 1 that follow a standard Gaussian distribution. The Z-score is defined on a data matrix X of dimension $p \times n$ as in (13), where X contains n number of face column vectors and each column is a face vector of p number of pixels. The notation '1:p,:' indicates all data values of corresponding face data matrix rows 1 to p and columns.

$$\bar{X} = \frac{X(1:p,:) - \mu(1:p)}{\sigma(1:p)} \tag{13}$$

i.e. subtracting the mean

$$\mu(1:p) = \sum_{j=1}^{n} \frac{X_{pj}}{n} \tag{14}$$

of face data matrix X from each pixel intensities, and then dividing by the standard deviation

$$\sigma(1:p) = \sqrt{\sum_{j=1}^{n} \frac{(X_{pj} - \mu(1:p))^2}{n-1}} \tag{15}$$

Therefore, the extraction of face features will be followed as

$$\tilde{X} = S^T \bar{X} \tag{16}$$

3.2 Analysis of representative face(s)

Representative face image(s) of one individual should be similar to all the face images of that person, but it should be dissimilar to the representative face image(s) of other individuals. The Fig. 7 shows ten non-orthogonal representative face images as well as their corresponding orthogonal representative face images from ORL [36] dataset. The orthogonal version of representative faces dramatically alters the appearance of the images while preserving their texture information. The altered version decorrelates faces from each other because they represent separate group of individuals.

To illustrate the benefits of representative (RP) score, we adopt the concept of best fitted regression line between two feature points. In this regard, fourteen face images of two persons are randomly selected. After linear projection of these faces on AOS, we carried out two representative points as features: one RP point in x-axis and second RP point in y-axis. Then, we drew the least square fitting line between these RP points which defines the residual for the i^{th} RP image as the error associated with that image. Figure 8 presents RP images arranged in decreasing order of RP scores, which also depicts the error associated with each RP image. We have noticed that the higher the RP score has lower is the error rate except in RP image-2 in Fig. 8a.

Fig. 7 (top) Ten representative face images of ten individual groups from the ORL Database and (bottom) corresponding orthogonal representative face images

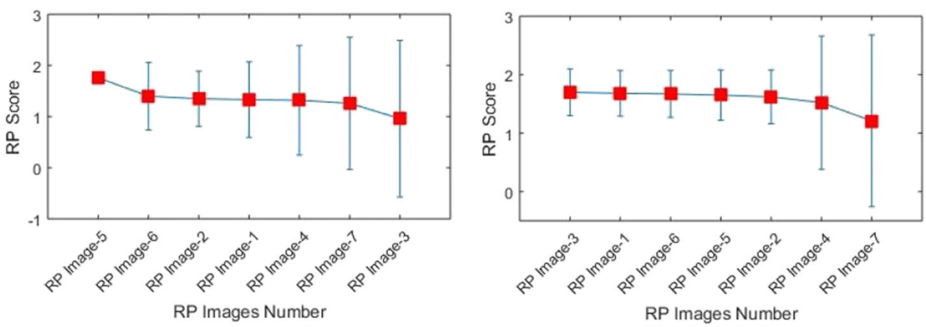

Fig. 8 **a** Residual Error estimation of first individual **b** Residual Error estimation of second individual from ORL Database

3.3 Distribution of AOS space

In literature, the orthonormal vector space is used to Gaussianize data, but there is no article that claims that the vector space itself is a Gaussian distribution. So, how has the orthonormal space S Gaussian distribution? In order to better display the influence of learned space i.e. AOS on distribution conception and a testimony of previously mentioned $Proposition$, we present a statistical test called Quantile-Quantile (Q-Q) plot [45]. The Q-Q plot is just a graphical tool to help us assess if a set of data plausibly came from some theoretical distribution such as Gaussian. If we run a statistical analysis assuming our space matrix is normally distributed, we can use standard normal Q-Q plot to check that assumption. A Q-Q plot is a scatter plot created by plotting two sets of quantiles against one another. If both sets of quantiles came from the same distribution, we would see the points forming a line that is roughly straight.

We have conducted an analysis from ORL and YALE [4] databases. From the datasets, the AOS space matrices are generated by combining number of representative face images ($R = 1$) of 30 individuals from ORL and 15 individuals from YALE. In above segments of Fig. 9a or b, the curve indicates density of a normal distribution and that histogram area of original space i.e. space without orthogonal between representative images is not fitted properly to normal curve. It becomes clearer when we plot out the Q-Q of original sample data vs. standard normal. From the points of both sample data and standard normal, we see that the points are not forming a straight line. In below segments of Fig. 9a or b, we have seen the probability densities of AOS data values that histogram area is almost laid down with classic symmetrical bell curve standard normal distribution and corresponding Q-Q plot also showing that the points of AOS data and standard normal forming a line that's roughly straight. Therefore, the AOS can transfer the original space to a signal close as Gaussian distribution that will facilitate to reduce the influence of grosses like variations in non-Gaussian or arbitrary distribution due to noises, lighting, and facial expressions.

4 The fusion approach between AOS learned features and custom VGG trained features

The accomplishment of deep neural networks is attributed to their ability to learn rich image representations, but they rely on estimating millions of parameters as well as require

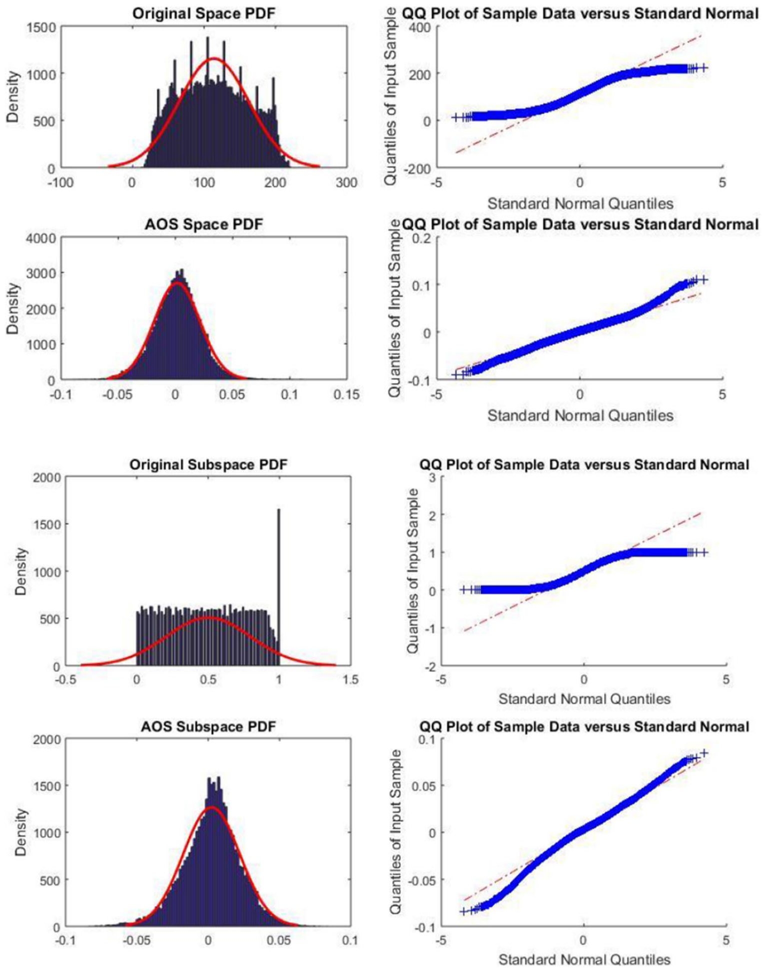

Fig. 9 A sample of distribution of the original space matrix and its AOS space matrix

a very bulk of images. In our case, we have very less amount of face data available when compared to the requirements of convolutional neural network (CNN). An extensively used alternative is to fine-tune the CNNs, which have been pre-trained using large amount of image datasets (e.g. ImageNet). The well-known CNN architectures in literature are VGG, ResNet, AlexNet and so on. One of the first CNN networks to push classification accuracy is AlexNet. It is composed of only 5 convolutional layers followed by 3 fully connected layers which is a smaller network when comapred to VGG16 network. These small networks sometimes cannot represent and explore face images in a concrete way. The VGG16, on the otherhand, makes improvement over AlexNet by replacing large kernel sized filters with multiple 3x3 kernel sized filters one after another. These multiple stacked smaller size kernel along multiple non-linear layers increases the depth of the network, which enables it to learn more complex and concrete features. The training with deeper networks, for example in ResNet consists of 152 layers performing optimization on huge parameter space and therefore naively adding the layers leading to higher training error. Hence, we are motivated to choose VGG16 with moderate number of layers and get satisfactory accuracy. However, our trail experiments show that only the use of pre-trained architectures with fine-tuning cannot provide satisfactory performance in small face datasets. To address this problem, we propose a fusion approach based on AOS features and a custom VGG features, as shown in Fig. 10. The AOS and custom VGG fused features are to be used together with fine tuning to further boost the face recognition performance. Furthermore, the custom VGG features also improve the performance of our proposed AOS features to a very great degree, as shown in the comparative study in Tables 3 or in 4.

In custom VGG, we removed the last *Dense* classification *Softmax* layer outputting class scores, and modified the fully connected (FC) layers according to our problem domain, and used the rest of pre-trained model as a fixed feature extractor. In addition, the number of output feature points of the last FC layer has been changed according to the number of individual persons (classes) as is present in our face datasets. To fuse the extracted features from both i.e. AOS (*fAOS*) and custom VGG (*fVGG*) features, we adopted the following weighted sum rule:

$$f = \alpha \times fAOS + \beta \times fVGG \tag{17}$$

where α and β indicate weights for AOS feature values and VGG feature values, respectively. In our case, α=0.4 and β=0.6. The resulting fused feature values can be used for subsequent classification.

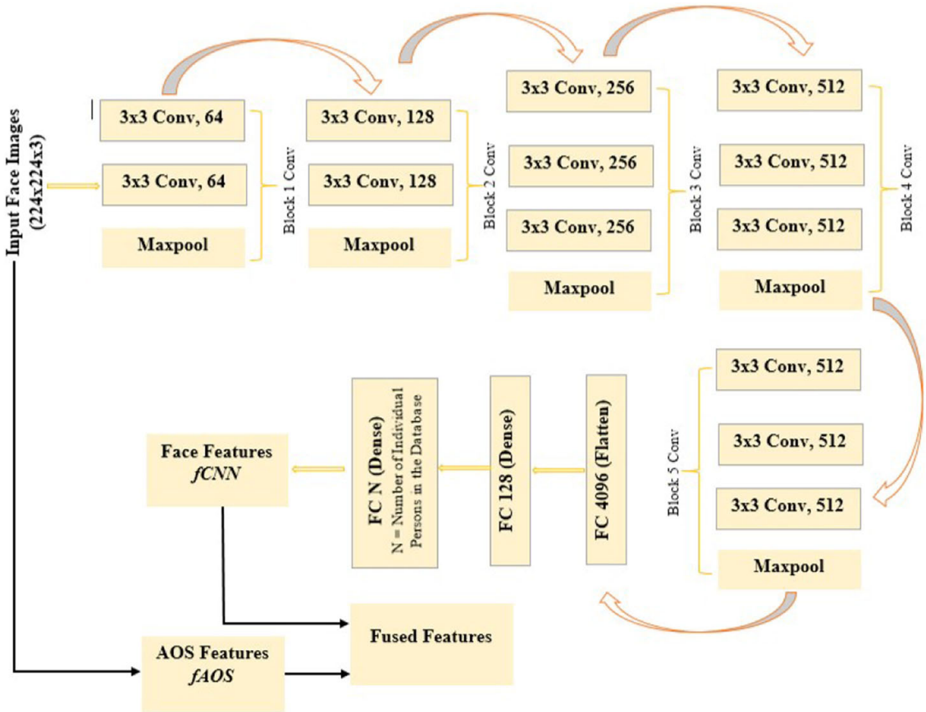

Fig. 10 Framework of the proposed fusion approach

5 Experimental results and performance evaluation

In this section, the proposed learning method AOS and deep features fusion approach are tested, and compared with other subspace learning methods using five visual face databases. The IRIS[1] and USTC-NVIE [39, 40] databases are used to evaluate the performance of proposed method over facial expressions by varying the number of representative (RP) face images. The illumination dataset of the IRIS database is also used to show the influence of our method over lighting variations. The ORL, YALE and FRGC databases are evaluated to compare the proposed methods with those relevant subspace learning methods. For comparative study, we only consider ORL, YALE and FRGC databases because most of the recent articles on this domain are basically made on these databases. Since we initially estimate the pair wise distance between face images in the AOS algorithm, transforming both images into one coordinate system is an essential issue. Therefore, we have used control point registration process of all group images of an individual based on one base frontal face image.

5.1 Experiments using the IRIS and USTC-NVIE visual face datasets

In these experiments, we mainly present the performance of AOS on the faces under expression variability, and lighting variability to demonstrate the effectiveness of proposed method. In this perspective, the face portion of each original image is manually cropped and resized to an image of 50× 50 pixels. The random forest [17] classifier with 50 number of trees is employed for this classification. Since the dimensionality is less than the sample size, the random forest offers a more complex model, which might be better in certain cases. Our parameter selection strategy is to allow the number of representative face images (RPs) vary from 1 to 5. Tables 1 and 2 shows the recognition rates of AOS versus variations in RP number.

Recognition of Faces under Expression Changes: For this experiment, we have collected 3 different expressions from both IRIS and USTC database, namely, surprise, laughing, and anger. In case of IRIS database, the expression dataset reposes of total 660 images from 20 individual persons and there are 33 different faces of each individual from 3 different expressions. To arrange the database according to our experimental setup, we have first created 3 probe sets. For instance, in probe set 1, 'surprise' and 'laughing' of 20 individual expressive face images (20× 22=440 images) are separated for training purpose and the rest of the images of 'anger' (20× 11=220 images) are kept for testing purpose. Similarly, the details of other two probe sets are given in Table 1. In case of USTC-NVIE database, a total of 270 expressive face images of 30 individual persons consists of 9 different face images from 3 different expressions. The details of arrangement of probe sets are also given in Table 1.

Table 1 lists the recognition rates across 15 tests for 5 representative points and 3 probe sets. There are few points to be taken from the table. In case of IRIS database, the recognition accuracies get superior in probe sets 1 and 3 i.e. when testing 'anger' and 'laughing' expressions against 'surprise-laughing' and 'surprise-anger' training expression sets respectively. Almost contrary scenario has been seen in case of USTC-NVIE dataset i.e. probe

[1] http://www.cse.ohio-state.edu/otcbvsbench/Data/02/download.html.

Table 1 Recognition Results (Accuracy (%)) on IRIS and USTC-NVIE Databases under Expression Variations

Prob sets	IRIS database							USTC-NVIE database						
	Number of train images	Number of test images	Number of Representative Face Images (RPs)					Number of train images	Number of test images	Number of Representative Face Images (RPs)				
			1	2	3	4	5			1	2	3	4	5
Probe Set 1	Surprise, Laughing (20× 22=440)	Anger (20× 11=220)	87.77	90	88.88	91.11	92.22	Surprise, Laughing (30× 6=180)	Anger (30× 3=90)	80.90	82.72	82.72	83.63	80
Probe Set 2	Laughing, Anger (20× 22=440)	Surprise (20× 11=220)	87.77	87.77	83.33	88.88	87.77	Laughing, Anger (30× 6=180)	Surprise (30× 3=90)	84.54	90.90	89.54	90	90.90
Probe Set 3	Anger, Surprise (20×22=440)	Laughing (20× 11=220)	91.11	90	88.88	90	92.22	Anger, Surprise (30× 6=180)	Laughing (30× 3=90)	84.09	82.27	83.63	83.63	85.45

Table 2 Recognition Results (Accuracy (%)) on IRIS database under variations with illumination

Prob sets	IRIS database						
	Number of training images	Number of testing images	Number of Representative Face Images (RPs)				
			1	2	3	4	5
Probe Set 1	Lon, Ron, 2on (17× 33=561)	Off (17×11=187)	60.96	64.17	58.82	56.68	59.35
Probe Set 2	Lon, Ron, Off (17×33=561)	2on (17×11=187)	81.81	86.63	87.70	84.49	87.16
Probe Set 3	Lon, 2on, Off (17×33=561)	Ron (17×11=187)	86.09	85.56	87.70	91.44	91.97
Probe Set 4	Ron, 2on, Off (17×33=561)	Lon (17×11=187)	84.49	81.28	85.02	82.35	82.35
Probe Set 5	Lon, Ron (17×22=374)	2on, Off (17×22=374)	71.92	75.40	71.12	69.51	74.06
Probe Set 6	Lon, 2on (17×22=374)	Ron, Off (17×22=374)	66.04	73.26	69.51	74.06	69.78
Probe Set 7	Lon, Off (17×22=374)	Ron, 2on (17×22=374)	66.84	68.18	69.78	68.18	72.72
Probe Set 8	Ron, 2on (17×22=374)	Lon, Off (17×22=374)	65.77	67.11	63.63	70.05	66.04
Probe Set 9	Ron, Off (17$times$22=374)	Lon, 2on (17×22=374)	76.73	76.73	74.59	80.74	77.54
Probe Set 10	2on, Off (17$times$22=374)	Lon, Ron (17×22=374)	79.25	82.44	81.91	85.37	81.91

set 2 giving better results than other two probe sets. With respect to the number representative face images, we have seen that RP=4, 5 are promising than other representative points. Overall, the proposed learning space is showing highest recognition rate of 92.22% for IRIS and 90.90% for USTC-NVIE.

Recognition of Faces under Illumination Changes: In this experiment, we consider the IRIS illumination dataset since USTC-NVIE is available only for expressive face images. The IRIS illumination dataset reposes of 748 images of 17 individual consisting of 44 different images from 4 illumination sets. These 4 illumination sets are left light on (Lon), right light on (Ron), both light on (2on), and left & right light off (Off). From these 4 illumination sets, we have created 10 probe sets for training and testing purposes, as shown in Table 2.

Table 2 lists the percentage of accuracies across 10 probe sets. Here, the probe set 3 i.e. testing face features of right light on (Ron) alongside training face features of left light on (Lon), both light on (2on), left & right light off (Off) gets more improvement. The lowest accuracies are highlighted in probe set 1 when testing face features of left & right light off (Off). A possible reason could be because the method is hard to recognize due to more dim faces. On the whole, a performance of highest 91.97% recognition was shown under variations in lighting environment.

5.2 Experiments using the ORL, YALE and FRGC face databases

For ORL or AT&T database, we have considered a total of 270 face images from 27 subjects. For some subjects, the images were taken at different times, varying the lighting, and facial expressions (open / closed eyes, smiling / not smiling).

For Yale database, we have considered a total of 150 face images from 15 subjects. There are 10 images per subject, one for each of the following facial expressions: normal, center-light, left-light, right-light, happy, sad, sleepy, surprised, and wink.

The proposed method is also tested on the Face Recognition Grand Challenge (FRGC) ver2.0 database [29], which consists of real facial images from uncontrolled conditions. We made use of a subset of FRGC for experiments such as: for each selected subject, 10 images are randomly selected for training and the rest are for testing. From the whole dataset, we have considered the uncontrolled conditions of illumination, expression, and pose.

In this section, we compare the performance of face recognition system over two categories. First, the performance over several distances' metrics will be evaluated. As we know, our proposed system starts with a distance matrix where each element value is a result of pair wise distance between two face images. Therefore, we consider a comparative assessment among several distance metrics, namely, euclidean, city block, minkowski, cosine, correlation, spearman, hamming, and jaccard etc. In order to alleviate the effect of the recognition performance caused by the choice of the distance metric's, we run the system over ORL and YALE databases. The resultant performances have been shown in Fig. 11. In both databases, we have seen that all the distance metrics are promising almost nearby performances, although correlation metric screening somehow slightly more outerperforms over all the representative (RP) points except RP-4 in case of ORL and RP-1 in case of YALE. Consequently, the second comparative study is carried out based on correlation distance metric.

To show the performance of proposed learning space and corresponding fusion approach with custom VGG network, we also present a comparative revise with 10 popular face recognition learning subspaces which are somehow relevant to our method. These are - orthogonal subspace based learning methods PCA, CCIPCA, SVD, independent subspace based learning methods ICA Architecture I, ICA Architecture II, PCA I, PCA II, graph

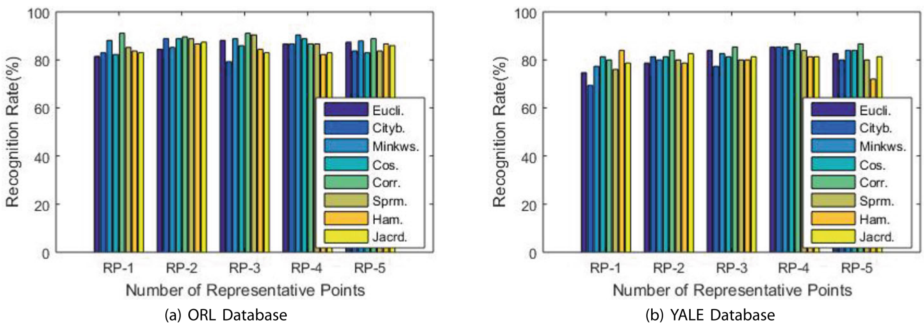

Fig. 11 Distance metric based comparative evaluation. Euclidean (Eucli.), CityBlock (Cityb.), Minkowski (Minkws.), Cosine (Cos.), Correlation (Corr.), Spearman (Sprm.), Hamming (Ham.), Jaccard (Jacrd.)

based orthogonal subspace learning methods OLGE, OLPP, and similarity metric based learning subspace method CCA. In order to alleviate the effect on the face recognition performance caused by the choice of principle components (PCs), independent components

Table 3 Average recognition results (accuracy (%)) on ORL, YALE and FRGC databases over random forest and decision tree

Methods	Random forest			Decision tree		
	ORL	YALE	FRGC	ORL	YALE	FRGC
Proposed Fusion (AOS + Custom VGG)	93.62	87.19	89.33	66	60	65.29
	92.73	*88.79*	*88.74*	*70.37*	*68*	*70.98*
Proposed Method (AOS)	87.40	82.66	81.47	65.62	60	62.13
	88.59	*77.33*	*79.33*	*66.07*	*67.20*	*69.25*
PCA	84	69.06	69.76	58.07	50.66	55.79
	86.66	*68.53*	*71.23*	*64.14*	*46.93*	*53.23*
CCIPCA	81.18	73.86	73.08	58.22	54.40	60.22
	79.70	*67.20*	*69.28*	*62.37*	*54.13*	*59.71*
SVD	84	70.93	71.25	57.62	50.66	51.23
	82.66	*73.06*	*71.66*	*64.14*	*46.93*	*49.21*
ICA Architecture-I	68.44	46.93	51.39	55.25	34.40	41.89
	72.44	*43.73*	*50.79*	*54.22*	*36.80*	*41.08*
ICA Architecture II	66.37	46.66	52.44	49.62	41.86	46.98
	72.44	*45.86*	*54.14*	*57.03*	*36.80*	*45.93*
PCA I	88	78.66	79.65	52	44.26	61.91
	89.33	*77.33*	*80.06*	*56.29*	*51.20*	*68.97*
PCA II	73.62	46.66	60.12	54.51	37.33	53.20
	69.77	*48.26*	*59.23*	*55.70*	*41.06*	*51.18*
OLGE	87.40	72	73.22	66.66	54.66	59.21
	78.51	*72*	*71.76*	*61.48*	*56*	*58.81*
OLPP	91.11	78.66	78.46	67.59	64	62.15
	92.59	*74.66*	*77.22*	*68.25*	*52*	*69.21*
CCA	84.59	81.60	81.09	52.44	41.06	59.23
	87.25	*77.06*	*79.68*	*54.96*	*55.46*	*68.68*

Italic text results represents second training phase, and normal text results represents first training phase

(ICs), or representative (RPs) points, we run each system 5 times with a gap of 20 PCs/ICs or 5 RPs (i.e. 1, 2, 3, 4, 5 RPs respectively, in our case) and then average them. There are two phases of training samples. In the first phase, half of the images per person are used for training and the remaining images are used for testing from the considered databases. In the second phase, the training and testing datasets are exchanged.

Table 3 shows the average recognition rate on ORL, YALE and FRGC databases over Random Forest (RF) and Decision Tree (DT) [30] classifiers. There are few main points to be taken. In case of ORL dataset, the proposed method shows at least 3% and 7% more accuracy in training phase 1 and phase 2, and at least 2% and 2% more in training phase 1 and phase 2 over RF and DT respectively. It is also noticed that the evaluation results of DT has not performed well in comparison to RF. The reason may be due to the continuous real-valued features which are not well suited for decision trees. Some exceptions are also noticed. The performance of AOS is slightly lesser than PCA I and OLPP in terms of average recognition rate. A possible reason is that PCA I reducing the pixel redundancy among different faces can enhance the robustness of PCA [43], and OLPP preserve local sparse structure of data in the transformed domain as well. In case of YALE and FRGC datasets, proposed method significantly outperforms than state-of-the-arts as well as their extended versions using RF. In both datasets, the recognition accuracies behave approximately similar. The performances of PCA I and CCA are most competitive with our proposed method than other state-of-the-arts. Classifying through DT, we analyse little improvement for FRGC than YALE. As usual, PCA I and OLPP get competitive results. At the end, the strategy to fuse AOS features and custom VGG features are analysed. The fusion approach outperforms above all and greatly improves the proposed learning space by maximum 11% of recognition rate.

The comparative studies on relevant methods with our proposed one depends on many environment setups: (i) number of training and testing images (ii) resolution size of images

Table 4 Comparison of the recognition accuracies (%) on ORL and YALE Databases

Method	ORL database		YALE database	
Proposed fusion approach	95.55		92	
Proposed Method (AOS)	89.62		86.66	
Lei Zhang et al. [49]			m=2, n=16	75.56
	–		m=1, n=32	74.44
			m=1, n=32	74.44
Jing Wang et al. [41]	t=2	81.69	t=4	76.67
	t=3	88.68	t=5	77.00
			t=6	82.53
K.Papachristou et al. [28]	p=10%	73.33	–	
	p=20%	85.00		
	p=30%	89.29		
Yong Xu et al. [47]	t=2	86.88	–	
	t=3	89.64		
J. Liang et al. [24]	–		t=5	86.28
			t=7	87.00

m, n is Block Size, t = Number of Images per Person to form the Training Set, p = % of Images per Person for Training

(iii) parameters of subspace techniques (iv) classification techniques (v) parameters of classification techniques and so on. For these reasons, we have presented a comparative study with hands-on implementation on our experimental setup in Table 3. As well we present another comparative study with direct imitative optimum recognition accuracy results of referred papers as shown in Table 4. The study shows that AOS produces comparable outcomes with other recent works, and, as usual, fusion approach outerperforms.

6 Conclusions

In this paper, we present an algorithm (AOS) with two key parts (i) part 1: representative face subspace (RFS) to choose the representative face(s) from a group of face images of an individual, which in turn works as a dimension reduction (ii) part 2: generating an orthogonal space by combining all non-orthogonal subspaces of part 1. This space follows a distribution which is close to Gaussian. Later on, we proposed another approach i.e. fused method with a custom VGG pre-train model. Classification experiments on both IRIS and USTC-NVIE databases demonstrate that AOS could work better on face under expression variants than lighting variety. The comparative studies of proposed learning space with other relevant spaces over ORL, YALE, and FRGC databases show that proposed method performs comparative to other state-of-the-art as well as topical methods. The fused results also show that convolutional neural network features can greatly improve the face recognition performances. In future work, we will extend our proposed work and explore the other possible grosses such as occlusion and pose variant problems. Full program is available for the research community, contact email or website respectively: mkb.cse@gmail.com or mrinalkantibhowmik@tripurauniv.in and www.mkbho-wmik.in.

Acknowledgements The work presented here is being conducted in the Biometrics Laboratory of Department of Computer Science and Engineering of Tripura University (A central university), Tripura, Suryamaninagar-799022.

Funding There is no funding for this research.

Compliance with Ethical Standards

Conflict of interests The author declares no potential conflict of interest with respect to the authorship and/or publication of this article.

References

1. Aapo H, Oja E (2000) Independent component analysis: algorithms and applications. Neural Netw 13:411–430
2. Bartlett MS, Lades HM, Sejnowski TJ (1998) Independent component representation for face recognition. In: Proc. SPIE symp. electron. imaging: sci. technol., SPIE, pp 528–539
3. Bartlett MS, Movellan JR, Sejnowski S (2002) Face recognition by independent component analysis. IEEE Trans Neural Netw 13:1450–1464
4. Bellhumer PN, Hespanha J, Kriegman D (1997) Eigenfaces vs. fisherfaces: recognition using class specific linear projection. IEEE Trans Pattr Anal Mach Intellig, Special Issue on Face Recognition 17:711–720
5. Bhowmik MK, Saha P, Singha A et al (2019) Enhancement of robustness of face recognition system through reduced gaussianity in log-ICA. Expert Syst Appl 116:96–107

6. Cai D, He X, Han J et al (2006) Orthogonal Laplacianfaces for face recognition. IEEE Trans Image Process 15:3608–3612
7. Cao C, Ai Pentland HZ (2015) Facial similarity learning with humans in the loop. J Comp Sci and Tech 30:499–510
8. Cazzanti L, Gupta MR (2007) Local similarity discriminant analysis. In: Proc. int. conf. machine learning (ICML). ACM, Corvallis, pp 137–144
9. Chandrasekaran S, Manjunath BS, Wang YF et al (1997) An eigenspace update algorithm for image analysis. Graph Models Image Process 59:321–332
10. Chen L, Casperson D, Liu Y et al (2013) A general probability framework for improving similarity based approaches for verification. In: Proc. IEEE int. conf. image processing (ICIP). IEEE, Melbourne, pp 3367–3371
11. Fagertun J, Gomez DD, Hansen MF et al (2011) Sparse similarity-based fisherfaces. In: Proc. Scandinavian conference on image analysis (SCIA). Lecture notes in computer science, vol 6688. Ystad, Sweden, pp 69–78
12. Frey BJ, Dueck D (2007) Clustering by passing messages between data points. Science 315:972–976
13. Fu Y, Yan S, Huang TS (2008) Correlation metric for generalized feature extraction. IEEE Trans Pattr Anal Mach Intellig 30:2229–2235
14. Gantmacher FR (1959) The theory of matrices. Chelsea Publishing Company, New York
15. Hall P, Marshall D, Martin R (2000) Merging and splitting eigenspace models. IEEE Trans Pattr Anal Mach Intellig 22:1042–1048
16. He X, Yan S, Hu Y et al (2004) Face recognition using Laplacianfaces. IEEE Trans Pattn Anal Mach Intell 27:328–340
17. Ho TK (1998) The random subspace method for constructing decision forests. IEEE Trans Pattr Anal Mach Intellig 20:832–844
18. Hou C, Nie F, Zhang C et al (2009) Learning an orthogonal and smooth subspace for image classification. IEEE Signal Processing Lttrs 16:303–306
19. Kim TK, Kittler J, Cipolla R (2010) On-line learning of mutually orthogonal subspaces for face recognition by image sets. IEEE Trans Image Process 19:1067–1074
20. Kim Y, Yoo JH, Choi K (2011) A motion and similarity-based fake detection method for biometric face recognition systems. IEEE Trans on Consum Electron 57:756–762
21. Kittler J, Li YP, Matas J (2000) On matching scores for LDA based face verification. In: Proc. British machine vision conf., Bristol, pp 42–51
22. Klami A, Virtanen S, Kaski S (2013) Bayesian canonical correlation analysis. J Mach Learn Res 14:965–1003
23. Klema VC (1980) The singular value decomposition: its computation and some applications. IEEE Trans Autom Control 25:164–176
24. Liang J, Chen C, Yi Y et al (2017) Bilateral two-dimensional neighborhood preserving discriminant embedding for face recognition. IEEE Access 5:17201–17212
25. Marshall D, Martin R (1998) Incremental eigen analysis for classification. In: Proc. British machine vision conference, Southampton, pp 286–295
26. Mian A, Bennamoun M, Owens R (2007) An efficient multimodal 2D-3D hybrid approach to automatic face recognition. IEEE Trans Pattr Anal Mach Intellig 29:1927–1943
27. Niyogi X (2004) Locality preserving projections. In: Proc. of the neural information processing systems, pp 153–160. Vancouver, arXiv
28. Papachristou K, Tefas A, Pitas I (2014) Symmetric subspace learning for image analysis. IEEE Trans Image Process 23:5683–5697
29. Phillips PJ, Flynn PJ, Scruggs T et al (2005) Overview of the face recognition grand challenge. In: Proc. IEEE computer society conference on computer vision and pattern recognition. IEEE Computer Society Press, Los Alamitos
30. Quinlan JR (1996) Learning decision tree classifiers. ACM Comput Surv 28:71–72
31. Singha A, Bhowmik MK (2018) Object recognition based on representative score features. In: Proc. of the IEEE 18th international conference on advanced learning technologies (ICALT). IIT Bombay, India, pp 419–421
32. Singha A, Bhowmik MK (2019) Salient features for moving object detection in adverse weather conditions during night time. In: IEEE Transactions on circuits and systems for video technology. In Press
33. Skocaj D, Leonardis A (2003) Weighted and robust incremental method for subspace learning. In: Proc. of the Ninth IEEE international conference on computer vision (ICCV'03). IEEE, France, pp 1494–1501
34. Tang L, Lu H, Pang Z et al (2019) A distance weighted linear regression classifier based on optimized distance calculating approach for face recognition. Multimed Tools Appl 78:32485–32501

35. Tang L, Li Z, Su J et al (2019) Kernel nearest-farthest subspace classifier for face recognition. Multimed Tools Appl 78:32007–32021
36. The AT&T Face Database. [Online]. Available: http://www.uk.research.att.com/facedatabase.html
37. Turk M, Pentland A (1991) Eigenfaces for recognition. J Cogn Neurosci 3:71–86
38. Turk M, Pentland A (1991) Face recognition using Eigenfaces. In: Proc. IEEE conf. comput. vis. pattern recog. IEEE, Maui, pp 586–591
39. Wang S, Liu Z, Lv S et al (2010) A natural visible and infrared facial expression database for expression recognition and emotion inference. IEEE Trans Multimed 12:682–691
40. Wang S, Liu Z, Wang Z et al (2013) Analyses of a multimodal spontaneous facial expression database. IEEE Trans Affect Comput 4:34–46
41. Wang J, Lu C, Wang M et al (2014) Robust face recognition via adaptive sparse representation. IEEE Trans Cybern 44:2368–2378
42. Weisstein EW Eigen decomposition. MathWorld
43. Weng J, Zhang Y, Hwang W (2003) Candid covariancefree incremental principal component analysis. IEEE Trans Pattr Anal Mach Intellig 25:1034–1040
44. Wigner E (1958) On the distribution of the roots of certain symmetric matrices. Ann Math 67:325–327
45. Wilk MB, Gnanadesikan R (1968) Probability plotting methods for the analysis of data. Biometrika Trust 55:1–17
46. Xia S, Shao M, Luo J et al (2012) Understanding kin relationships in a photo. IEEE Trans Multimed 14:1046–1056
47. Xu Y, Zhong Z, Yang J et al (2017) A new discriminative sparse representation method for robust face recognition via l2 regularization. IEEE Trans Neur Netw Learn Syst 28:2233–2242
48. Yang J, Zhang D, Jing-Yu Y, Constructing PCA (2007) Baseline algorithms to reevaluate ICA-based face recognition performance. IEEE Trans Syst Man Cybern-Part B: Cybern 37:1015–1021
49. Zhang L, Gao Q, Zhang D (2007) Block independent component analysis for face recognition. In: Proc. 14th Int. conference on image analysis and processing (ICIAP). IEEE, Modena, pp 217–222

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Anu Singha received Master's Degree in Computer Applications, and Computer Science & Engineering from South Asian University (A SAARC University), New Delhi, and Tripura University (A Central University), Agartala, India in 2013 and 2015, respectively. From 2015, he is pursuing his Ph.D. in Computer Science & Engineering from Tripura University, India. Currently, he is also working as Junior Research Fellow (JRF) under a Defense Research and Development Organization (DRDO) funded project. His research interests include computer vision, object detection, face recognition, and deep learning etc.

Mrinal Kanti Bhowmik obtained his Bachelor of Engineering in Computer Science & Engineering from Tripura Engineering College in 2004 and Masters of Technology in Computer Science & Engineering from Tripura University (A Central University), India, in 2007. In 2014, he received Ph.D. (Engineering) degree from Jadavpur University, Kolkata, India. He has successfully completed two DeitY funded project, one DBT-Twinning funded project and one SAMEER funded project as Principal Investigator. Currently, he is the Principal Investigator of the two Govt. of India projects, one DRDO funded project and one ICMR funded project. From July 2010 onward, he is serving in the Department of Computer Science & Engineering at Tripura University as an Assistant Professor. His current research interests are in the field of computer vision, medical imaging, biometrics etc. He is also a Senior Member of the IEEE (USA).

Debotosh Bhattacharjee received the MCSE and Ph.D. (engineering) degrees, in 1997 and 2004, respectively from Jadavpur University, India. Currently he is serving as Professor in Department of Computer Science & Engineering, in Jadavpur University, Kolkata, India. His research interests include pertain to fuzzy logic, genetic algorithm, cellular automata, rough set theory, artificial neural network, etc. in face recognition, OCR, and information security. He is author or co-author of over 150 journal and conference publications, and several book chapters in various areas of computer vision. He is currently a Senior Member of IEEE (USA).