IEEE Access
The journal for rapid open access publishing

Received June 4, 2014, accepted July 18, 2014, date of publication August 14, 2014, date of current version September 8, 2014.

Digital Object Identifier 10.1109/ACCESS.2014.2348018

Gradient-Orientation-Based PCA Subspace for Novel Face Recognition

GHEORGHITA GHINEA[1], (Member, IEEE), RAJKUMAR KANNAN[2], AND SURESH KANNAIYAN[3]
[1]Department of Computer Science, Brunel University, Uxbridge UB8 3PH, U.K.
[2]College of Computer Sciences and Information Technology, King Faisal University, Al-Ahsa 31982, Saudi Arabia
[3]Department of Computer Science, Bishop Heber College (Autonomous), Tiruchirappalli 620017, India

Corresponding author: G. Ghinea (george.ghinea@brunel.ac.uk)

ABSTRACT Face recognition is an interesting and a challenging problem that has been widely studied in the field of pattern recognition and computer vision. It has many applications such as biometric authentication, video surveillance, and others. In the past decade, several methods for face recognition were proposed. However, these methods suffer from pose and illumination variations. In order to address these problems, this paper proposes a novel methodology to recognize the face images. Since image gradients are invariant to illumination and pose variations, the proposed approach uses gradient orientation to handle these effects. The Schur decomposition is used for matrix decomposition and then Schurvalues and Schurvectors are extracted for subspace projection. We call this subspace projection of face features as Schurfaces, which is numerically stable and have the ability of handling defective matrices. The Hausdorff distance is used with the nearest neighbor classifier to measure the similarity between different faces. Experiments are conducted with Yale face database and ORL face database. The results show that the proposed approach is highly discriminant and achieves a promising accuracy for face recognition than the state-of-the-art approaches.

INDEX TERMS Face recognition, object recognition, pattern recognition.

I. INTRODUCTION

Face recognition, which is nothing but recognizing human faces from digital images and videos, has been an important area of research in the past three decades. *Automatic Face Recognition* (AFR) has been widely used for biometric authentication, video surveillance and tagging personal image and video collections. Most of the previous face recognition methodologies failed to handle various poses and lighting conditions in face images. Studies in this field have focused much on reducing the effect of changes in illuminations and poses [1]–[5]. Comparing two face images in the original image dimension is computationally expensive and inaccurate. The well-known Principle Component Analysis (PCA) method reduces the image dimensions and is widely used for face recognition. It represents the correlated face images into *Eigenfaces* which are linearly uncorrelated. The PCA often uses matrix decomposition methods such as Singular Value Decomposition (SVD), QR and Schur decomposition. The Schur decomposition method was used for face recognition in [6] and [7]. It is preferred most of the time because of its numerical stability. The stability issues of Schur decomposition are discussed in [8]. It handles the defective matrices while computing the linearly independent eigenvalues and eigenvectors. The deity face recognition by Balakrishnan *et al*. [9] is a recently proposed face recognition approach that uses the schur decomposition with PCA and their experiments showed the discriminant power of the schur decomposition for face recognition.

The pose and illumination variations of face images affect the performance of the face recognition methods. One of the ways to handle these variations in face images is to represent them as gradient oriented image, because gradient oriented features are invariant to *illuminations* and small *pose*s. At the same time, they also preserve the important structural properties of an image.

Apart from feature extraction and decomposition, distance measures also have significant impact on the accuracy of a face recognition system. Wang *et al*. [10] have shown that the performance of the k-Nearest Neighbor classifier can be improved when Hausdorff distance (HD) measure is used to find the similarity between training and testing images. While there have been several researches on PCA, to the best of our knowledge no research has leveraged the power of HD and Schur decomposition and ours is the first attempt to integrate Schur decomposition and Hausdorff distance for dimensionality reduction based face recognition.

This paper proposes a novel approach for face recognition. The image gradient orientation is used to extract the structural property initially. The subspace projection is obtained using Schur decomposition with PCA dimensionality reduction. Through this, the original high dimensional images are represented in low dimensional features. We call these features as *Schurfaces*. To achieve a better discriminative power, HD distance measure with nearest neighbors is applied to classify the face images. Here we use the same with *Schurfaces*. It is novel and unique in the sense that it is tolerant to *pose* and *illumination* changes while recognizing face images. It is an adaptive version of the traditional PCA. To show the efficiency of our proposed approach, we compare our methodology with the state-of-the-art face recognition methods: Eigenfaces [11], Fisherfaces [12], Laplacianfaces [13] and Fisher discriminant with Schur decomposition (FDS) [6].

The rest of this paper is organized as follows: Section 2 presents the related work on the existing face recognition methods. Section 3 presents an overview of the face recognition methods namely Eigenfaces, Fisherfaces, Laplacianfaces and FDS. Section 4 presents the proposed Schurfaces method and shows how it can be applied for face recognition. Section 5 discusses the experimental results and the efficiency of the proposed method, while section 6 concludes the paper and describes the future work.

II. RELATED WORK

Face recognition methodologies deal with large image dimensions and thus make the recognition task very complex. The concept of dimensionality reduction was introduced to reduce the image dimensions. It also reduces the computational complexity while manipulating high dimensional data. The PCA is a widely used method [14]–[16] for dimensionality reduction and subspace projection which is also known as *Karhunen–Loeve Transformation* (KLT). It usually follows SVD method to decompose the image (matrix) and is utilized for the human face representation [17], [18]. Kirby and Sirovich [17] have proved that any face image can be approximately reconstructed by a mean image and *eigenfaces* of the original face images. Following this approach, the well-known *eigenfaces* method for the face recognition is presented by Turk and Pentland [11]. Since then, PCA became one of the most successful approaches for face recognition [12].

Almost all PCA based face recognition methods follow SVD technique to decompose the covariance matrix of centered training images. In contrast, QR decomposition based PCA methods were proposed for face recognition [19]. Adaptive modified PCA using Sanger's algorithm and QR decomposition that was proposed by Ghassabeh *et al.* [14] is a more time consuming technique because of its iterative nature of calculating eigenvalues and eigenvectors. Schur decomposition method is a fast and numerically stable method [8], [20] for computing linearly independent eigenvalues of a matrix. Also, it is suitable for the *defective matrix* decomposition.

Belhumeur *et al.* [12] had introduced the *Fisherfaces* approach for subspace learning by considering both within-class scatter and between-class scatter of original data. The *Laplacianfaces* is an appearance based face recognition approach introduced by He and Yan [13]. It is based on Locality preserving projections (LPP). Fisher discriminant with Schur decomposition was introduced by Song *et al.* [6]. However FDS did not show better recognition accuracy and it is a time consuming approach.

To improve the face recognition rate, the authors in [15], [16], and [21] applied preprocessing technique before learning subspace from the original images. Down-sampling is considered as preprocessing step in [21]. Whitening filter is used as a preprocessing step in [15]. Hsieh *et al.* [16] proposed an image partition technique which combines vertically centered PCA and whitened horizontally centered PCA to get better recognition performance compared to the traditional PCA. Gradient orientation preserves the important structure of the image. It is much useful in face recognition. To improve the face recognition accuracy, Tzimiropoulos *et al.* [22] used gradient orientation with PCA.

The Hausdorff distance (HD) is a similarity measure that is applied in many face recognition approaches since it is less sensitive to noise. Huttenlocher *et al.* [23] have proposed a method based on HD to compare face images. The HD measure with k-Nearest Neighbors classifier was proposed by Wang *et al.* [10] which improves the face recognition accuracy when gradient features are used with Nearest Neighbors classifier.

III. FACE RECOGNITION METHODS–OVERVIEW

This section gives an overview of the standard face recognition methodologies namely Eigenfaces, Fisherfaces, Laplacianfaces and Fisher discriminant with Schur decomposition.

A. EIGENFACES FOR FACE REPRESENTATION

The *Eigenfaces* is the commonly used face recognition method in computer vision through dimensionality reduction. It reduces the original image dimension and creates a linear subspace projection. It also maximizes the scatter of all projected images [12].

Let us consider a two dimensional face image $I_{N\times N}$. It can be represented as a vector with a dimension N^2. This high dimensional vector can be represented in a low dimensional subspace without losing much data of the original face image. The idea of PCA is to find the vectors that best suit to represent the training set of face images. These vectors are then used to compute the subspace of the face images. Here each vector has the length of N^2 that describes an $N \times N$ image. The subspace is the linear combination of the original face images. The face like appearance of eigenvectors is called Eigenfaces. Let the training set of images be $I_1, \ldots, I_N$. Then the average of the original images will be

$$\boldsymbol{\mu} = \frac{1}{\mathrm{N}} \sum_{i=1}^{\mathrm{N}} \mathrm{I_N} \quad (1)$$

The eigenvectors λ_v and eigenvalues λ_i are computed from the scatter (covariance) as

$$\phi_c = \frac{1}{\mathrm{N}} \sum_{i=1}^{\mathrm{N}} (\mathrm{I_i} - \boldsymbol{\mu})(\mathrm{I_i} - \boldsymbol{\mu})^{\mathrm{T}} \quad (2)$$

Here, $(\mathrm{I_i} - \boldsymbol{\mu}) = \tilde{v}$ is the variance. The eigenvectors are then used to represent the face like appearance called eigenfaces. More details on how eigenfaces are used to compare the testing face image can be found in [11]. The drawback of this approach is that it only considers between-class-scatter in subspace projection not within-class-scatter and, more about this can be found in [12].

B. FISHERFACES FOR FACE REPRESENTATION

The *Fisherfaces* are originally derived from the Fisher Linear Discriminant (*FLD*) method. Unlike the *Eigenfaces* technique discussed in the previous section, *Fisherfaces* considers both with-in-class and between-class scatter matrix for subspace projection [12] of the original images. The between-class scatter matrix of the original matrices can be defined as

$$\mathrm{S_B} = \sum_{i=1}^{c} \mathrm{N_i}\,(\boldsymbol{\mu}_\mathrm{i} - \boldsymbol{\mu})\,(\boldsymbol{\mu}_\mathrm{i} - \boldsymbol{\mu})^{\mathrm{T}} \quad (3)$$

and the within-class scatter matrix is defined as

$$\mathrm{S_w} = \sum_{i=1}^{c} \sum_{\mathrm{x_k} \varepsilon \mathrm{X_i}} (\boldsymbol{\mu}_\mathrm{i} - \boldsymbol{\mu})\,(\boldsymbol{\mu}_\mathrm{i} - \boldsymbol{\mu})^{\mathrm{T}} \quad (4)$$

where $\boldsymbol{\mu}_\mathrm{i}$ is the mean image of the original class $\mathrm{X_i}$, and $\mathrm{N_i}$ is the number of samples in class $\mathrm{X_i}$. The maximized ratio of the determinant of the between-class scatter to the determinant of the within-class scatter is

$$\mathrm{W_{opt}} = \arg\max_{\mathrm{w}} \frac{|\mathrm{W^T s_B W}|}{|\mathrm{W^T s_w W}|} = [\mathrm{w_1 w_2 \cdots\ w_m}] \quad (5)$$

where $\{w_i | i = 1, 2, \ldots, m\}$ is the set of eigenvectors of $\mathrm{S_B}$ and $\mathrm{S_w}$ corresponding to the m largest eigenvalues $\{\lambda_i | i = 1, 2, \ldots, m\}$. It uses the PCA to reduce the dimension of the feature space and further it uses FLD method as defined in [12].

$$\mathrm{W}_{opt}^{T} = \mathrm{W}_{fld}^{T}\ \mathrm{W}_{pca}^{T} \quad (6)$$

where

$$\mathrm{W}_{pca} = \arg\max_{\mathrm{W}} |\mathrm{W}^T \mathrm{S}_T \mathrm{W}| \quad (7)$$

$$\mathrm{W}_{fld} = \arg\max_{\mathrm{W}} \frac{|\mathrm{W}^T \mathrm{W}_{pca}^T \mathrm{S}_B \mathrm{W}_{pca} \mathrm{W}|}{|\mathrm{W}^T \mathrm{W}_{pca}^T \mathrm{S}_W \mathrm{W}_{pca} \mathrm{W}|} \quad (8)$$

The major problem with this approach is that the number of training samples per subject should be more than one. It is a very time consuming method and uses large memory space than *Eigenfaces*.

C. LAPLACIANFACES FOR FACE REPRESENTATION

The *Laplacianfaces* has been another popular face recognition method and a good overview of the *Laplacianfaces* for subspace representation can be found in [13]. The *Laplacianfaces* is derived from the Locality Preserving Projection (*LPP*). It is a linear dimensionality reduction technique and is a linear approximation of the nonlinear Laplacian Eigenmap [24]. As presented by He *et al.* [25], LPP builds a graph incorporating the neighborhood information of the dataset. It uses the Laplacian of a graph and then computes a transformation matrix to build the subspace of the dataset. The derivations and theoretical justifications of LPP can be found in [25]. Here, for the given dataset $\mathrm{x_1}, \ldots, \mathrm{x_m}$ in $\mathrm{R^n}$, we find a transformation matrix A that maps these m points to a set of points $\mathrm{y_1}, \ldots, \mathrm{y_m}$ in $\mathrm{R^l}$ $(l \ll n)$, where $\mathrm{y_i}$ represents $\mathrm{x_i}$ and $\mathrm{y_i = A^T x_i}$. The objective function of the LPP is,

$$\min \sum_{\mathrm{ij}} (\mathrm{y_i} - \mathrm{y_j})^2\ \mathrm{S_{ij}} \quad (9)$$

where $\mathrm{y_i}$ is the one dimensional representation of the original data $\mathrm{x_i}$ and S is a similarity matrix as defined below:

$$\mathrm{s_{ij}} = \left\{ \begin{array}{ll} \exp(-\|x_i - \mathrm{x_j}\|^2/\mathrm{t}), & \|\mathrm{x_i} - \mathrm{x_j}\|^2 < \varepsilon \\ 0 & \text{otherwise} \end{array} \right\} \quad (10)$$

$$\mathrm{s_{ij}} = \left\{ \begin{array}{ll} \exp(-\|\mathrm{x_i} - \mathrm{x_j}\|^2/\mathrm{t}), & \text{if } \mathrm{x_i} \text{ is among k nearset} \\ & \text{neighbors of } \mathrm{x_j} \text{ or } \mathrm{x_j} \text{ is among} \\ & \text{k nearset neighbors of } \mathrm{x_i} \\ 0 & \text{otherwise} \end{array} \right\} \quad (11)$$

where ε is small, and greater than zero. It defines the locality i.e., radius of the local neighborhood. The LPP includes the following steps:

1. Construct the adjacency graph G with nodes m.
2. Choose the weights W. It is sparse $m \times m$ matrix having the edge joining vertices i and j, and 0 if there is no edge.
3. Compute Eigenmaps: $\mathrm{XLX^T a} = \lambda \mathrm{XDX^T a}$, where D is a diagonal matrix whose entries are column sums of W,

$$\mathrm{D_{ij}} = \sum\nolimits_{\mathrm{j}} \mathrm{W_{ij}}$$

The Laplacian matrix is $\mathrm{L = D - W}$.

D. FISHER DISCRIMINANT WITH SCHUR DECOMPOSITION (FDS)

The Schur decomposition theorem is stated in [26]. It produces a novel eigenanalysis which is very effective when used with Fisher Discriminant Analysis (FDA) for face recognition applications. The details of FDA with Schur decomposition can be found in [6]. The Fisher linear discriminant has the computation $O(mnt+t^3)$ and requires $O(mn+mt+nt)$ memory space, where m is number of samples, n is the number of features extracted from the face images and $t = min(m, n)$. It is computationally very expensive and infeasible when both m and n are very large. FDS has all the characteristics of FDA except the eigenanalysis using the Schur decomposition.

The Schur decomposition method is computationally efficient, reliable and numerically stable. For a given square matrix A, Schur decomposition can be represented as

$$A = Q\,U\,Q^T \tag{12}$$

where, Q is unitary matrix, U is an upper-triangular matrix and its diagonals $(\lambda_1, \ldots, \lambda_n)$ are schurvalues. It is then followed by the evaluation

$$\sqrt{A} = U\,\mathrm{diag}(\sqrt{\lambda_i})U^T \tag{13}$$

Here, the computation cost is $10n^3$. Like Fisherfaces, FDS also has some demerits such as complex computations for subspace creation, large memory utilization and large processing time.

IV. PROPOSED SCHURFACES FOR FACE RECOGNITION

The proposed Schurfaces is a linear subspace projection method characterizing high discriminative power and low computational complexity. It follows the traditional PCA technique with Schur decomposition for eigenanalysis. It has special characteristics such as computational efficiency, numerical stability and handling of non-diagonalizable matrix (defective matrix). The eigenvalues of a matrix $A_{n\times n}$ is computable even though it has less than n linearly independent eigenvectors. Here we call eigenvalues and eigenvectors as Schurvalues and Schurvectors.

The idea of our proposed approach is to create a low dimensional subspace that better represents the original gradient oriented images with reduced features. It is achieved through finding the Schurvectors of largest schurvalues. These schurvectors define the subspace of the original images and are computed from the schur-decomposition of the covariance matrix. These vectors hold face features and we call this as *Schurfaces*. The block diagram of our proposed approach is shown in figure 1. Let us consider the gradient orientated face features $E = \{e_1, e_2, \ldots, e_N\}$, where each e_i represents face feature. These faces can be projected into a subspace as

$$\rho = \lambda_f' \tilde{v} \tag{14}$$

here, λ_f' is a transposed Schurfaces and $\tilde{v}$ is a mean centered matrix. The Schurfaces λ_f are computed as shown below:

$$\lambda_f = \tilde{v}\lambda_v \tag{15}$$

where, λ_v is an orthogonal vector of the covariance matrix. Schurvalues (λ_i) and their corresponding Schurvectors are computed by decomposing the covariance matrix. Then, the training set projection feature is computed using Schurfaces and centered matrix $(\tilde{v})$. The scatter matrix of the mean centered matrices can be computed as:

$$\phi_c = \sum_{i=1}^{N} (e_i - \boldsymbol{\mu})\,(e_i - \boldsymbol{\mu})^T \tag{16}$$

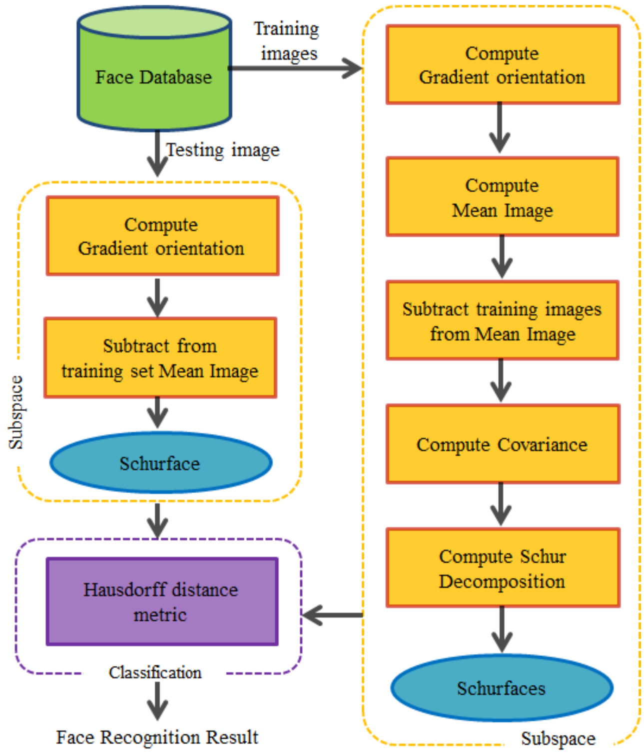


FIGURE 1. Block diagram of the proposed approach.

It takes computation time $O(nm^2)$. Here $e_i - \boldsymbol{\mu} = \tilde{v}$ is the variance, and $\boldsymbol{\mu}$ is the mean matrix that is computed as

$$\boldsymbol{\mu} = \frac{1}{N}\sum_{i=1}^{N} E \tag{17}$$

Using Schur decomposition any real square matrix $(m \times n)$ can be decomposed. Here, the covariance matrix ϕ_c is decomposed as

$$\phi_c = Q\,U\,Q^T \tag{18}$$

where U is an upper triangular matrix having schurvalues in diagonals $(\lambda_1, \ldots, \lambda_n)$. The x_{ij}'s are some upper triangular values and Q is an orthogonal matrix $(Q^TQ = QQ^T = I)$ and also $Q^T = Q^{-1}$. The diagonal elements of U are sorted to make $\lambda_1 \geq \lambda_2 \cdots \geq \lambda_n$, so that we have optimality in the reconstruction with respect to a little loss of face image information. The schurvectors are the columns of orthogonal matrix Q for the corresponding schurvalue λ_i in matrix U. We call these schurvectors as *Schurfaces*. The gradient orientation of a testing image T is extracted and projected to the subspace. The projection features of the test image T is computed by subtracting from training set mean matrix $\boldsymbol{\mu}$ and then multiplying by the transposed Schurfaces λ_f' such that

$$\rho_t = (T - \boldsymbol{\mu})\lambda_f' \tag{19}$$

The reconstructed training image features and testing image features are compared for classification using Hausdorff distance metric. This face recognition approach resembles some characteristics of the PCA and it differs in eigenvalue computations. Some of the notable features of the proposed *Schurfaces* are:

- While other methods such as *Eigenfaces, Fisherfaces, Laplacianfaces* and Fisher discriminant with Schur decomposition (*LDS*) aim to project the original face images into low dimensional space, our *Schurfaces* project the gradient orientation of the original face images into low dimensional subspace.
- An efficient algorithm for face recognition should be able to show the discriminative power with large set of face images. Our proposed *Schurfaces* has more discriminative power than the other approaches.
- In any classification problem, one requires a distance metric for classification. It is well known that Nearest Neighbors classifier is the standard and simple method for classification and often used in many pattern recognition applications. While many real world pattern recognition applications use Euclidean distance measure, in contrast we use the HD to compute the dissimilarity between two images. It is because HD is very effective when computing dissimilarity between two images.

The accuracy of the proposed approach is shown through the experiments in section 5. To show the efficiency of the proposed approach, we compared our method with the standard methods such as *Eigenfaces*, *Fisherfaces, Laplacianfaces* and *FDS*.

V. EXPERIMENTAL RESULTS

In this section, we present the experiments results using the face datasets such as Yale, ORL and Indian DEity data Set (IDES) to show the efficiency of our proposed *Schurfaces* for face representation and recognition problem.

TABLE 1. Performance comparison on Yale database.

Method	Number of Dimensions	Error rate (%)
Eigenfaces	47	25
Laplacianfaces	21	17
Fisherfaces	24	20
FDS	113	16
Schurfaces	21	14

A. PREPROCESSING

Let us consider a set of face samples $\mathrm{I} = \{\mathrm{I}_1, \ldots, \mathrm{I}_\mathrm{N}\}$ of size $m \times n$ for feature extraction. Generally, images are noisy and vary in size. Therefore, the input images need to be normalized to a standard dimension. The gradient orientation of face images are extracted and then used to create *Schurfaces*.

B. FACE RECOGNITION USING SCHURFACES

We conduct experiments with our proposed *Schurfaces*. As explained in section 4, *Schurfaces* method has been implemented. Once *Schurfaces* are created, the task moves to classification of the testing face images. We have used the Nearest Neighbor classifier because of its simplicity. The Hausdorff distance is computed between two images I_1 and I_2 using the following equation:

$$\mathrm{H}(\mathrm{I}_1,\ \mathrm{I}_2) = \max(\mathrm{h}(\mathrm{I}_1, \mathrm{I}_2), \mathrm{h}(\mathrm{I}_2,\ \mathrm{I}_1)); \tag{20}$$

where,

$$\mathrm{h}(\mathrm{I}_1,\ \mathrm{I}_2) = \max_{a\in \mathrm{I}_1} \min_{b\in \mathrm{I}_2} \|a - b\| \tag{21}$$

here, a and b are the pixel points in I_1 and I_2 respectively. Further in this section, we will show the performance of our proposed *Schurfaces* for face recognition and compare our approach with the standard methods: *Eigenfaces*, *Fisherfaces*, *Laplacianfaces* and *FDS*.

C. EXPERIMENTS ON YALE DATABASE

The Yale database contains GIF formatted 165 grayscale images of 15 individuals. There are 11 images per person with variations such as *center-light, with glasses, happy, left-light, with no glasses, normal, right-light, sad, sleepy, surprised,* and *wink*. We took six images per subject from the database for training and the remaining images for testing. As explained in section 4, *Schurfaces* were created using the training images and then the testing images were projected into a low-dimensional representation. The face recognition was performed using Nearest Neighbors classifier with HD metric. The *Schurfaces* showed high discriminative power since it uses gradient orientation, Schur decomposition and HD metric. The best face recognition results are shown in table 1.

D. EXPERIMENTS ON ORL DATABASE

The ORL database contains 10 different images of 40 distinct subjects in PGM file format with pixel resolution of 92×112. These 8-bit grey images were taken at *different times, varying lighting slightly, facial expressions (open/closed eyes, smiling/non-smiling)* and *facial details (with glasses/without glasses)*. This database was populated at the Olivetti Research Laboratory in Cambridge, UK during April 1992 to April 1994.

We randomly took 5 different images of each subject as training set images and the remaining images are considered as testing images. The subspaces were created using the training images and then the testing images were projected into a low-dimensional representation and classification was performed using Nearest Neighbors classifier with HD. The performance of all five approaches differs with the number of dimensions. The face recognition error rates are shown in table 2 and from that, it is obvious that *Schurfaces* outperforms the other approaches.

E. EXPERIMENTS ON IDES DATABASE

The IDES dataset contains Indian deity face images with different poses and illuminations. The dataset consists of

TABLE 2. Performance comparison on ORL database.

Method	Number of Dimensions	Error rate (%)
Eigenfaces	45	23
Laplacianfaces	33	19
Fisherfaces	23	23
FDS	23	18
Schurfaces	23	15

FIGURE 2. Sample images from IDES dataset.

800 deity face images of 20 Indian deities, where each class contains 40 face images. Sample deity face images from IDES are shown in figure 2. It is a collection of illumination and pose varied deity face images. We randomly took 5 different images of each subject as training set images and the remaining images are considered as testing images. The performance of the methods is shown in table 3.

TABLE 3. Performance comparison on IDES dataset.

Method	Number of Dimensions	Error rate (%)
Eigenfaces	61	26
Laplacianfaces	17	20
Fisherfaces	21	24
FDS	21	22
Schurfaces	39	18

VI. CONCLUSION

In this paper, a novel approach for face representation which is pose and illumination invariant, namely *Schurfaces*, is introduced and experimentally evaluated for face recognition. To the best of our knowledge, this is the first approach that uses gradient orientation and Schurvectors for subspace learning. It utilizes the schur decomposition for computing schurvalues and schurvectors. The experiments showed that the *Schurfaces* has the high discriminant power and consistently outperformed the standard face recognition methods. In future, we will focus on preprocessing techniques and bidirectional subspace learning techniques to improve face recognition accuracy.

REFERENCES

[1] F. Bashar, A. Khan, F. Ahmed, and H. Kabir, "Face recognition using similarity pattern of image directional edge response," *Adv. Elect. Comput. Eng.*, vol. 14, no. 1, pp. 69–76, 2014.

[2] Z. Chai, H. Mendez-Vazquez, R. He, Z. Sun, and T. Tan, "Explore semantic pixel sets based local patterns with information entropy for face recognition," *EURASIP J. Image Video Process.*, vol. 2014, no. 1, pp. 1–11, May 2014.

[3] G. Anbarjafari, "Face recognition using color local binary pattern from mutually independent color channels," *EURASIP J. Image Video Process.*, vol. 2013, no. 1, pp. 1–11, 2013.

[4] K. Zhang, L. Zhang, and M.-H. Yang, "Real-time compressive tracking," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2012, pp. 864–877.

[5] N.-S. Vu and A. Caplier, "Face recognition with patterns of oriented edge magnitudes," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2010, pp. 313–326.

[6] F. Song, D. Zhang, and J. Yang, "A novel dimensionality-reduction approach for face recognition," *Neurocomputing*, vol. 69, nos. 13–15, pp. 1683–1687, 2006.

[7] G. Wang and W. Zhang, "Neighborhood preserving fisher discriminant analysis," *Inf. Technol. J.*, vol. 10, no. 12, pp. 2464–2469, 2011.

[8] S. Michael and P. Van Dooren, "Stability issues in the factorization of structured matrices," *SIAM J. Matrix Anal. Appl.*, vol. 18, no. 1, pp. 104–118, 1997.

[9] R. Balakrishnan, R. Kannan, S. Kannaiyan, and S. Swaminathan, "Deity face recognition using schur decomposition and hausdorff distance measure," in *Proc. IEEE 56th Int. Midwest Symp. Circuits Syst. (MWSCAS)*, Aug. 2013, pp. 1184–1187.

[10] J. Wang and Y. Tan, "Hausdorff distance with k-nearest neighbors," in *Advances in Swarm Intelligence*. Berlin, Germany: Springer-Verlag, 2012, pp. 272–281.

[11] M. A. Turk and A. P. Pentland, "Face recognition using eigenfaces," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 1991, pp. 586–591.

[12] P. N. Belhumeur, J. P. Hespanha, and D. Kriegman, "Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 7, pp. 711–720, Jul. 1997.

[13] X. He, S. Yan, Y. Hu, P. Niyogi, and H.-J. Zhang, "Face recognition using laplacianfaces," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 3, pp. 328–340, Mar. 2005.

[14] Y. A. Ghassabeh, H. A. Moghaddam, and M. Teshnehlab, "Adaptive modified PCA for face recognition," in *Proc. Conf. MVA*, 2007, pp. 126–129.

[15] L.-Z. Liao, S.-W. Luo, and M. Tian, "'Whitenedfaces' recognition with PCA and ICA," *IEEE Signal Process. Lett.*, vol. 14, no. 12, pp. 1008–1011, Dec. 2007.

[16] P. C. Hsieh and P. C. Tung, "A novel hybrid approach based on sub-pattern technique and whitened PCA for face recognition," *Pattern Recognit.*, vol. 42, no. 5, pp. 978–984, 2009.

[17] M. Kirby and L. Sirovich, "Application of the Karhunen–Loeve procedure for the characterization of human faces," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 12, no. 1, pp. 103–108, Jan. 1990.

[18] A. Pentland, "Looking at people: Sensing for ubiquitous and wearable computing," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 1, pp. 107–119, Jan. 2000.

[19] S. Gudivada and A. G. Bors, "Face recognition using ortho-diffusion bases," in *Proc. 20th Eur. Signal Process. Conf. (EUSIPCO)*, Aug. 2012, pp. 1578–1582.

[20] R. W. Aldhaheri, "Model order reduction via real Schur-form decomposition," *Int. J. Control*, vol. 53, no. 3, pp. 709–716, 1991.

[21] F. Chichizola, L. De Giusti, A. De Giusti, and M. Naiouf, "Face recognition: Reduced image eigenfaces method," in *Proc. 47th Int. Symp. ELMAR*, Jun. 2005, pp. 159–162.

[22] G. Tzimiropoulos, S. Zafeiriou, and M. Pantic, "Principal component analysis of image gradient orientations for face recognition," in *Proc. IEEE Int. Conf. Automat. Face Gesture Recognit. Workshops (FG)*, Mar. 2011, pp. 553–558.

[23] D. P. Huttenlocher, G. A. Klanderman, and W. J. Rucklidge, "Comparing images using the Hausdorff distance," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 15, no. 9, pp. 850–863, Sep. 1993.

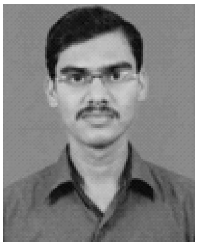
[24] M. Belkin and P. Niyogi, "Laplacian eigenmaps and spectral techniques for embedding and clustering," in *Proc. Conf. NIPS*, vol. 14. 2001, pp. 585–591.

[25] X. He and P. Niyogi, "Locality preserving projections," in *Proc. Conf. NIPS*, vol. 16. 2003, pp. 234–241.

[26] H. G. Gene and C. F. Van Loan, *Matrix Computations*, vol. 4. Baltimore, MD, USA: The Johns Hopkins Univ. Press, 2012, p. 351.

GHEORGHITA GHINEA (M'01) received the B.Sc. and B.Sc. (Hons.) degrees in computer science and mathematics, and the M.Sc. degree in computer science from the University of the Witwatersrand, Johannesburg, South Africa, in 1993, 1994, and 1996, respectively, and the Ph.D. degree in computer science from the University of Reading, Reading, U.K., in 2000. He is currently a Reader with the School of Information Systems, Computing and Mathematics, Brunel University, Uxbridge, U.K. His current research interests include multimedia computing, telemedicine, quality of service, and computer networking and security issues.

RAJKUMAR KANNAN received the B.Sc. and M.Sc. degrees in computer science from Bharathidasan University, Tiruchirappalli, India, in 1991 and 1993, respectively, and the Ph.D. degree in multimedia datamining from the National Institute of Technology, Tiruchirappalli, in 2007. He is currently with the College of Computer Sciences and Information Technology, King Faisal University, Al-Ahsa, Saudi Arabia. His research activities primarily lie at the confluence of multimedia, information retrieval, semantic web, social informatics, and collective intelligence. He is a member of the Association for Computing Machinery and the British Computer Society, and a Life Member of the Computer Society of India and the Indian Society For Technical Education.



SURESH KANNAIYAN received the B.Sc. degree in computer science from Bishop Heber College (Autonomous), Tiruchirappalli, India, in 2009, and the master's degree in computer applications from Anna University, Tiruchirappalli, in 2012. He is currently pursuing the Ph.D. degree in computer science with the Department of Computer Science, Bishop Heber College (Autonomous). His main interests include machine learning, computer vision, image processing, and information retrieval.