Diagonal Locality Preserving Projection as Dimensionality Reduction Technique with Application to Face Recognition

Veerabhadrappa[1,2]
[1]Department of Computer Science,
University College, Mangalore- 575 001,
Karnataka, India

Lalitha Rangarajan[2]
[2]Department of Studies in Computer Science,
University of Mysore, Manasagangothri,
Mysore- 570 006, India

ABSTRACT

In this paper, a new dimensionality reduction technique called Diagonal Locality Preserving Projections (DiaLPP) is proposed. In contrast to Locality Preserving Projection (LPP) and Two Dimensional Locality Preserving Projection (2DLPP), DiaLPP directly seeks the optimal projection vectors from diagonal images without vector transformation. The 2DLPP method seeks optimal projection vectors by using the row information of the image and the Alternate 2DLPP method seeks optimal projection vectors by using the column information of the image, whereas the DiaLPP seeks optimal projection vectors by interlacing both the rows and column information of the images. Experimental results on subset of UMIST and ORL face database shows that the proposed method achieves higher recognition rate than 2DLPP, Alternate 2DLPP and DiaPCA (Diagonal Principal Component Analysis).

Keywords: Locality Preserving Projection (LPP), Two-dimensional LPP, Principal Component Analysis (PCA), Dimensionality Reduction, Diagonal image, face recognition.

1. INTRODUCTION

In most of the pattern recognition applications measurements made are inherently multidimensional in nature and a representation of data in fewer dimensions can be advantageous for processing the data. Therefore, a procedure called Dimensionality Reduction (DR) is required to find the intrinsic low dimensional structures hidden in the high dimensional observations. Many statistical learning problems involve some form of dimensionality reduction either explicitly or implicitly. The goal may be one of feature selection in which, we aim to find linear or nonlinear combination of the original set of variables or one of the variable selection in which we wish to select a subset of variable from the original set. The dimensionality reduction is concerned with the problem of mapping data points that lie on or near low-dimensional manifold in a high dimensional data space to a low dimensional embedding space. Hence much importance has been attributed to the process of dimensionality reduction which is the most fundamental and one of the important stages in the field of Pattern Recognition. Principal Component Analysis (PCA) [7] is a well known dimensionality reduction technique widely used in the area of pattern recognition, computer vision, signal processing etc. In PCA, the main idea is to project the original data on the reduced number of orthogonal projection axes that restores the largest possible variance in the original data and it is a well known method for dimensionality reduction. But PCA is not optimal for general classification problems because it is unsupervised and ignores the valuable class label information. The Linear Discriminant Analysis (LDA) [1] is used to maximize the ratio of between-class variance to the within-class variance thereby guaranteeing maximal separability between the classes. The performance of LDA is degraded when encountering the small sample size and singularity problems. The Independent Component Analysis (ICA) [6] is a higher order method that seeks linear projection not necessarily orthogonal to each other that is as nearly statistically independent as possible. Recently several novel methods have been proposed to tackle the non linear data namely Kernel PCA (KPCA)[9], Locally Linear Embedding (LLE) [8], Isomap [13] and Supervised Isomap (S-Isomap) [13] etc. The KPCA is a non linear approach to extend PCA such that it can find non linear subspace with high variance. The Kernel PCA finds principal components which are nonlinearly related to the input space by performing PCA in the space produced by the nonlinear mapping (through Kernels), where the low dimensional latent structure is easier to discover. Both Isomap and LLE have attempted to preserve the local neighborhood features while trying to obtain highly non linear embeddings (Local Embeddings). But LLE and Isomap fail when data are complex and noisy. Hence S-Isomap is proposed to recover the true manifold of the noisy data and to preserve the class label information. Laplacianfaces [5] is based on a technique called Locality Preserving Projection (LPP) [4] which finds an embedding that preserves local information and obtain face subspace which best detects the essential face manifold structure. If training samples are insufficient and data dimension is high, especially for image data LPP can't be used directly due to singularity of matrices. In the above mentioned subspace based models, images are transformed into 1D vector by adjoining either column by column or row by row. This leads to high dimensional vector space and evaluation of covariance matrix and eigenvectors is time consuming. To overcome this difficulty, two dimensional PCA (2DPCA) [14] and 2DLPP [3, 10] are proposed in which an image covariance matrix is constructed from the image matrices for feature extraction. It evaluates the image covariance matrix more accurately and computes the corresponding eigenvectors more efficiently than PCA and these methods operates on image matrix and works in row direction of the image. To consider the column direction of the images, alternate 2DPCA and alternate 2DLPP [11] were proposed. Two Directional two Dimensional LPP $(2D)^2$LPP[12] was also proposed which simultaneously consider both the row and column direction of the image. Motivated by the work Diagonal PCA(DiaPCA) by Daiquiang and Zhou[2], in this paper we propose Diagonal LPP(DiaLPP) which seeks optimal projection vectors from diagonal images so that the correlation between the variations of both rows and columns of images can be preserved. The experimental results indicate that DiaLPP give higher recognition accuracy with less recognition time when compared to 2DLPP, Alternate 2DLPP and DiaPCA.

The rest of this paper is organized as follows: Section 2 briefly reviews the 2DLPP method; The proposed DiaLPP method is introduced in Section 3; In Section 4, some experiments on COIL object dataset are given to reveal the performance and superiority of the proposed method over 2DLPP, Alternate 2DLPP and DiaPCA. Conclusion is presented in the Section 5.

2. TWO-DIMENSIONAL LOCALITY PRESERVING PROJECTION (2DLPP): A REVIEW

Let **A** = $[A_1, A_2, \ldots, A_N]$ be the N sample images taken from an (m x n) dimensional image space. The 2DLPP is applied on these N sample images to obtain an n-dimensional unitary column vectors: $W=(w_1, w_2, \ldots, w_n)$. The feature matrix X_i for the image A_i is obtained by projecting onto W using the following transformation:

$X_i = A_i W$, i=1,2,…,N.

The computation of transformation function W is outlined below. It shall be observed here that the computation of transformation function W does not require the transformation of an image matrix into vector form and hence is two dimensional rather than one dimensional.

1). **Constructing the nearest-neighbor graph**: Let G denote a graph with N nodes, i^{th} node corresponding to image A_i. Insert an edge between nodes i and j if A_i and A_j are nearer. Either the method a) or b) given below can be used as a measure of closeness.

a) *k-nearest neighbors*: Nodes i and j are connected by an edge if i is among k nearest neighbors of j or j is among k nearest neighbors of i.

b) *ε-neighborhoods*: Nodes i and j are connected if $\|A_i - A_j\| < \varepsilon$ where the distance between two matrices $\|*\|$ is just the Euclidean distance between their vectorization representation in R^{mn}

2). **Choosing the weights:** If there is an edge between nodes i and j, put a similarity weight S_{ij} on it, otherwise let $S_{ij}=0$. Then a sparse symmetric similarity matrix $[S]_{NxN}$ is obtained and the similarity weight S_{ij} can be any one of the following

a) *Simple-minded*: $S_{ij}=1$ if and only if nodes i and j are linked by an edge

b) *Heat kernel*: If nodes i and j are linked, then

$$S_{ij} = e^{-\frac{\|A_i - A_j\|^2}{t}}$$ where t is some constant.

3). **Eigenmap**: Compute the eigenvectors and eigenvalues for the generalized eigenvalue problem:

$$\mathbf{A^T(L \otimes I_m)AW = \lambda A^T(D \otimes I_m)AW} \quad \textbf{(1)}$$

where **D** is diagonal matrix with $D_{ii} = \sum_j S_{ij}$; **L = D - S** is the Laplacian matrix; **A** is an (*mNxn*) matrix generated by arranging all the image matrices in column $A = [A_1^T, A_2^T, \ldots, A_N^T]^T$, Operator $\otimes$ is the Kronecker product of the matrices and $\mathbf{I_m}$ is the identity matrix of order m.

Let $w_1, w_2, \ldots, w_d$ be the first *d* unitary orthogonal solution vector corresponding to the *d* smallest generalized eigenvalues, ordered according to their magnitude $0 \le \lambda_1 \le \lambda_2 \le \ldots \le \lambda_d$. These eigenvalues are nonnegative because $\mathbf{A^T(L \otimes I_m)A}$ and $\mathbf{\lambda A^T(D \otimes I_m)A}$ are both symmetric and positive semi definite. Thus the embedding is as follows:

$X_i = A_i W$, $W=(w_1, w_2, \ldots, w_d)$, $i=1,2,\ldots,N$,

Here, X_i is (*m x d*) feature matrix of A_i and W is the (*n x d*) transformation matrix. These matrices are used for classification purpose using nearest neighbor classifier.

3. PROPOSED MODEL

3.1 Diagonal Locality Preserving Projection (DiaLPP)

From the literature we observe that, the 2DPCA [14] and 2DLPP [3, 10] reflects only the information between rows, which implies column information is missing. Similarly, Alternate 2DPCA and Alternate 2DLPP [11] reflects only column information by omitting the row information. This implies that with standard 2DLPP or Alternate 2DLPP method, we can only capture either row or column information at the same instant. We attempt to solve this problem by transforming the original faces into corresponding *diagonal face images*. Because the rows (or columns) in the transformed diagonal images simultaneously integrate the information of rows and columns in the original images, it can reflect both information between rows and columns.

Suppose that there are M images denoted by m by n matrices A_k (k = 1,2,…,M). For each image, define the corresponding *diagonal image* as follows:

(1) If the height m is equal to or smaller than the width n, use the method illustrated in Figure 1 to generate the diagonal image D for the original image A.

(2) If the height m is bigger than the width n, use the method illustrated in Figure 2 to generate the diagonal image D for the original image A.

(3) After obtaining the diagonal images as shown in Figure 3(b) for all the M images, apply the 2DLPP algorithm explained in Section 2.

Algorithm :DiaLPP[Training Phase]

Input: Set of images: $\mathbf{A} = \{A_i^j \mid i=1...N, j=1...M\}$ where N is the number of images; M is the number of views and each image is of size m x n.

Output: Knowledge base : $\mathcal{F} = \{F_i^j \mid i=1...N, j=1...M\}$

Method:

[A] Computation of optimal projection axes in row direction : W

- Compute the diagonal image **D** of size (mN x n) by arranging all images column wise.
- Find eigenvectors by solving the equation (1).
- Choose d eigenvectors W = $(w_1, w_2, \ldots, w_d)$ associated with first d smallest eigenvectors .

[B] Create knowledge base by projecting W on the set of images as follows:

$\mathcal{F} = \{F_i^j = A_i^j W \mid i = 1...N,\ j = 1...M\}$

Algorithm DiaLPP [Training Phase] ends

Algorithm : DiaLPP[Recognition Phase]

Input : Test Image I(m x n),
Knowledge base $\mathcal{F}$,
Optimal projection axes : W

Output : Class label of I

Method:

1. Obtain the feature matrix I^f of the input image I using W and Z, $I^f = Z^T I W$
2. Find F_r^S such that
$\| I^f - F_r^S \|_2 = \arg\min(\| I^f - F_i^j \|_2, \forall i = 1...N,\ j = 1...M)$,
where $\| . \|_2$ denotes the Euclidean distance
3. Classify the test image I as a member of the r^{th} class

Algorithm DiaLPP[Recognition Phase] ends

4. EXPERIMENTAL RESULTS

In this section, we present the experimental results of proposed model. To corroborate the success of the proposed model, the well known existing dimensionality reduction techniques such as 2DLPP, Alternate 2DLPP and Diagonal PCA have been considered for comparative study. The superiority of the proposed model is established through the recognition accuracy and recognition time. We performed all experiments on the standard UMIST face dataset. All our experiments are carried out on a Core2 Duo PC machine (2.20 GHz, 2.19 GHz) and 1GB RAM under Matlab 7.5 platform.

4.1. Results on partial UMIST face database

The actual UMIST Face Database consists of 564 images of 20 people. Each covers a range of poses from profile to frontal views. The files are all in PGM format, approximately 220 x 220 pixels in 256 shades of grey. For the purpose of reducing the computation burden of Kronecker product in 2DLPP, Alternate 2DLPP and DiaLPP, we have taken the faces of all 20 people with 8 different views contains both left views and right views , each with reduced size of 50 x 50. We have conducted experiments on this dataset in order to corroborate the success of the proposed method for face recognition. We have used the disjoint set for training and testing. It can be observed from Table 1 and Figure 4 (a)-(d) that the proposed DiaLPP method has better recognition rate with very less recognition time when compared to other methods.

4.2 Results on partial ORL face database

The actual ORL database contains 400 images from 40 individuals, each providing 10 different images with the size of 112 x 92. For the purpose of reducing the computation burden of Kronecker product in 2DLPP, Alternate 2DLPP and DiaLPP, we have taken 20 images with 10 different views, each with reduced size of 50 x 50. We have conducted a series of experiments to compare the performance of DIALPP, Alternate 2DLPP, 2DLPP and DiaPCA with varying number of training views and testing views. The recognition accuracy and recognition time of each method is summarized in Table 2. It shall be observed from Table 2 that the DiaLPP consumes less time for recognition with relatively higher recognition rate when compared to other methods. The recognition performance of 2DLPP, Alternate 2DLPP, DiaPCA and DiaLPP with varying number of dimension of feature vectors and varying number of training samples is shown in Figure 5 (a)-(d).

5. CONCLUSION

In this paper, a new dimensionality reduction method DiaLPP is introduced which is suitable for an efficient face representation and recognition. Here diagonal image which contain both the row and column information is considered to compute the transformation function. The main advantage of DiaLPP is that it requires fewer coefficients needed for face representation and recognition unlike PCA/2DPCA/2DLPP. The success of the proposed model is demonstrated experimentally by considering the standard face datasets UMIST and ORL. The proposed model is relatively faster and has better recognition rate when compared to the other well known dimensionality reduction approaches like 2DLPP, Alternate 2DLPP and DIAPCA.

Original Image

A A

a_{11}	a_{12}	a_{13}	a_{14}	a_{15}	a_{11}	a_{12}	a_{13}	a_{14}	a_{15}
a_{21}	a_{22}	a_{23}	a_{24}	a_{25}	a_{21}	a_{22}	a_{23}	a_{24}	a_{25}
a_{31}	a_{32}	a_{33}	a_{34}	a_{35}	a_{31}	a_{32}	a_{33}	a_{34}	a_{35}
a_{41}	a_{42}	a_{43}	a_{44}	a_{45}	a_{41}	a_{42}	a_{43}	a_{44}	a_{45}

→

Diagonal Image

D

a_{11}	a_{12}	a_{13}	a_{14}	a_{15}
a_{22}	a_{23}	a_{24}	a_{25}	a_{21}
a_{33}	a_{34}	a_{35}	a_{31}	a_{32}
a_{44}	a_{45}	a_{41}	a_{42}	a_{43}

Figure 1. Way of producing diagonal image when the width is greater than the height

Original Image

A

A

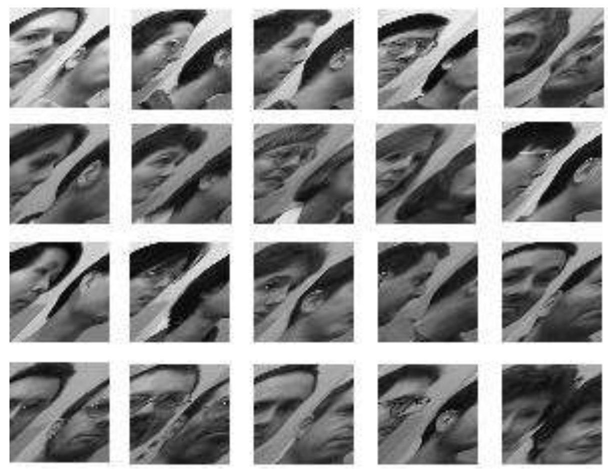
a_{11}	a_{12}	a_{13}	a_{14}
a_{21}	a_{22}	a_{23}	a_{24}
a_{31}	a_{32}	a_{33}	a_{34}
a_{41}	a_{42}	a_{43}	a_{44}
a_{51}	a_{52}	a_{53}	a_{54}
a_{11}	a_{12}	a_{13}	a_{14}
a_{21}	a_{22}	a_{23}	a_{24}
a_{31}	a_{32}	a_{33}	a_{34}
a_{41}	a_{42}	a_{43}	a_{44}
a_{51}	a_{52}	a_{53}	a_{54}

Diagonal Image
D

a_{11}	a_{22}	a_{33}	a_{44}
a_{21}	a_{32}	a_{43}	a_{54}
a_{31}	a_{42}	a_{53}	a_{14}
a_{41}	a_{52}	a_{13}	a_{24}
a_{51}	a_{12}	a_{23}	a_{34}

Figure 2. Way of producing diagonal image when the height is greater than the width

(a) Sample UMIST face dataset



(b) Diagonal image of UMIST face dataset

(c) Sample ORL face dataset

(d) Diagonal image of ORL face dataset

Figure 3. Sample dataset and their corresponding diagonal images

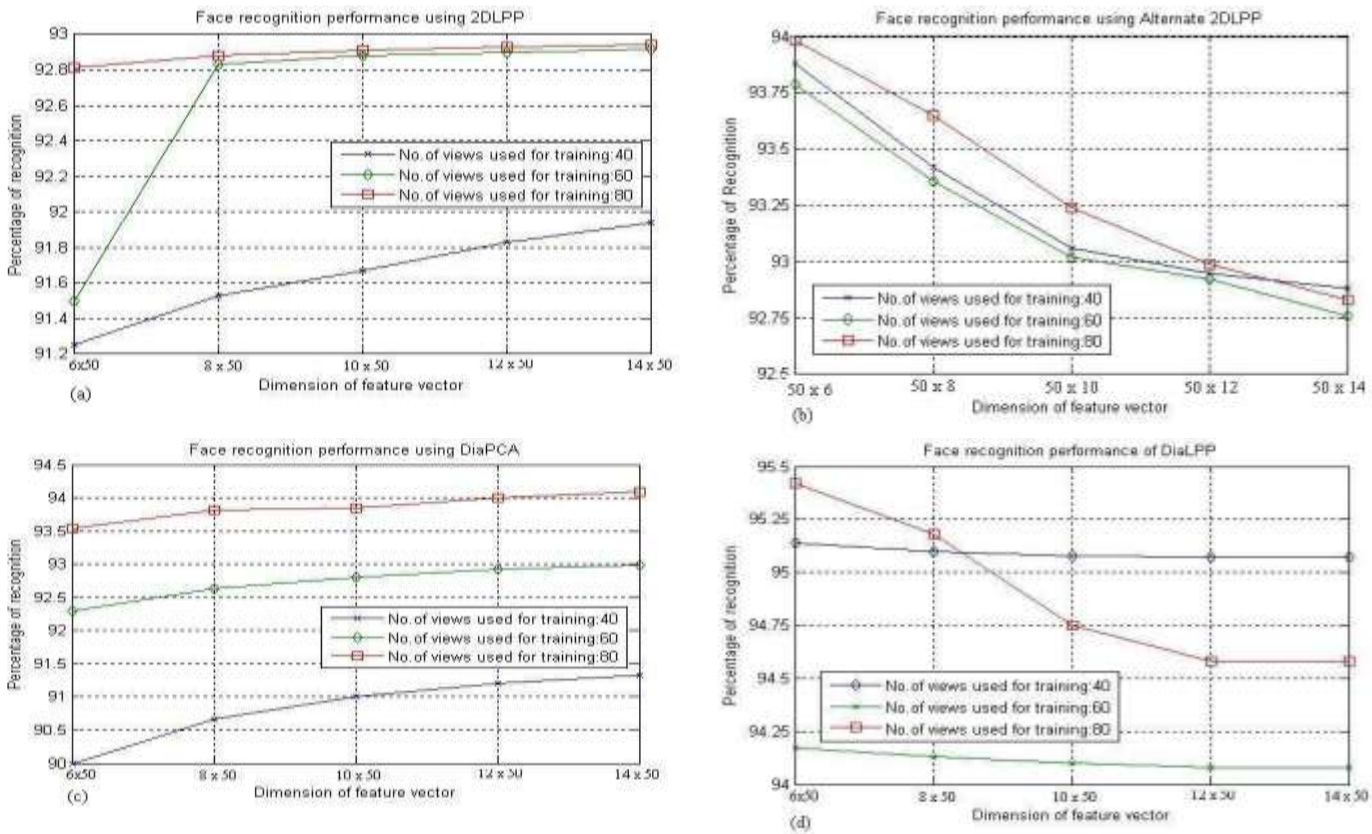


Figure 4: Performance with varying number of training samples and varying number of dimension of feature vectors (a) 2DLPP (b) Alternate 2DLPP (c) DiaPCA (d) DiaLPP for UMIST face dataset

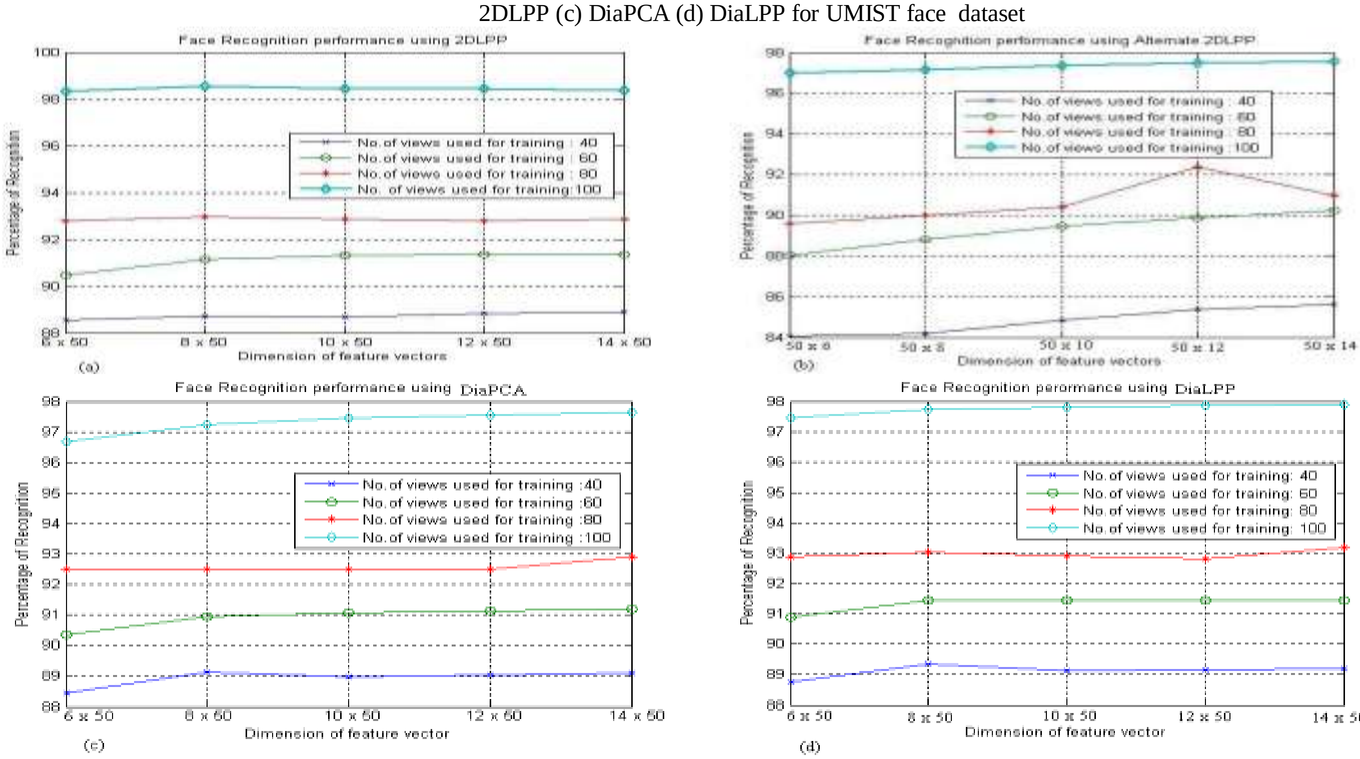


Figure 5: Performance with varying number of training samples and varying number of dimension of feature vectors (a) 2DLPP (b) Alternate 2DLPP (c) DiaPCA (d) DiaLPP for ORL face dataset

Table 1: Face Recognition Performance of 2DLPP, Alternate 2DLPP, DiaPCA and DiaLPP for UMIST face dataset

No. of Views used for training	No. of Views used for testing	Percentage of Recognition and (Recognition time in seconds)			
		2DLPP (14 x 50)	**Alt-2DLPP(50 x 6)**	**DiaPCA (14 x 50)**	**DiaLPP (6 x 50)**
80	80	92.94 (140.96)	93.98 (139.62)	94.10 (30.18)	95.42 (26.18)
60	100	92.92 (126.54)	93.88 (131.31)	92.99 (27.35)	95.14 (24.82)
40	120	91.94 (106.94)	93.79 (106.51)	91.33 (24.31)	94.17 (20.23)

Table 2: Face Recognition Performance of 2DLPP, Alternate 2DLPP, DiaPCA and DiaLPP for ORL face dataset

No. of Views used for training	No. of Views used for testing	Percentage of Recognition and (Recognition time in seconds)			
		2DLPP(10 x 50)	**Alt-2DLPP(50 x 14)**	**DiaPCA (14 x 50)**	**DiaLPP (14 x 50)**
100	100	95.73(216.56)	97.58(337.62)	97.63(98.03)	97.88(84.14)
80	120	92.96(207.31)	90.97(320.58)	92.92(93.24)	93.18(89.38)
60	140	91.36(182.51)	90.24(281.32)	91.19(88.34)	91.43(84.12)
40	160	88.91(140.53)	85.63(217.18)	89.10(72.56)	89.21(62.53)

6. REFERENCES

[1]. Balakrishnama.S.Ganapathiraju, "Linear Discriminant Analysis- A brief Tutorial", Institute for signal and Information, MS, 1998.

[2]. Daoqiang Zhang, Zhi_Hua Zhou, Songcan Chen, Diagonal Principal Component Analysis for face recognition, Pattern recognition 2006, 39(1), pp 140-142.

[3]. Dewan Hu, Guiyu feng, Zongtan Zhou, Two Dimensional locality Preserving projections with its application to palm print recognition , Pattern Recognition 2007, Vol 40(10), pp 339-342.

[4]. He.X and Niyogi.P, "Locality Preserving Projections", Advances in Neural Information Processing Systems, 16, 2003.

[5]. He.X.F, S.YAn, Y.Hu, P.Niyogi and H.J.Zhang, Face recognition using Laplacianfaces, IEEE Trans. Pattern Analysis and Machine Intelligence, 2005, Vol 2793, pp 328-340.

[6]. Hyvarinen A, "Survey on Independent Component Analysis", Neural Computing Surveys 1999, Vol 2: 94-128

[7]. Jolliffe.I.T, "Principal Component Analysis" Springer Verlag, NY, 1986.

[8]. Roweis.S.T and L.K.Saul, "Nonlinear dimensionality reduction by locally linear embedding", Science 2000, Vol.290, pp 2323-2326

[9]. Scholkopf.B, Smola.A, and Muller.K.R, "Kernel PCA", Advances in Kernel methods, Support Vector Learning, MIT Press, 1999, pp.327-352.

[10]. Sibao Chen, Haifeng Zhao, Min Kong, and Bin Luo, 2D-LPP: a two-dimensional extension of locality preserving projections, Journal of Neuro-computing, 2007, Vol.70 (4-6), pp 912-921.

[11]. Veerabhadrappa, Lalitha Rangarajan, B.H.Shekar, Alternate 2DLPP: A new dimensionality reduction technique for clustering" Proceedings of National Conference on Recent Trends in Information and Communication Technology (RTICT2008) Tamilnadu, India,2008, pp 240-244.

[12]. Veerabhadrappa, Lalitha Rangarajan, B.H.Shekar, $(2D)^2$LPP: A new dimensionality reduction technique with application to face/object representation and recognition, International Journal of Systemics, Cybernetics and Informatics, April 2009, pp 17-22.

[13]. Xin Geng, De-Chuan Zhan, and Zhi-Hua Zhou, "Supervised Nonlinear Dimensionality Reduction for Visualization and Classification", IEEE Transactions on Systems, Man and Cybernetics, 2005, Vol 35(6), pp 1098-1107.

[14]. Yang.J, D. Zhang, A.F. Frangi, and J.Yang, Two-Dimensional PCA: A new approach to appearance based face representation and recognition, IEEE Trans. Pattern Analysis and Machine Intelligence, 2004, Vol.26(1), pp 131-137.

7. ABOUT THE AUTHORS

Veerabhadrappa obtained his B.Sc., and M.Sc., degrees in Computer Science and Technology from the University of Mysore, India, respectively, in the years 1987 and 1989. He is a faculty in the Department of Computer Science, University College, Mangalore, India and currently he is on Teacher Fellowship to carry out Ph.D in the Department of Studies in Computer Science, University of Mysore, India. He authored 8 peer-reviewed papers in journals and conferences. His area of research covers dimensionality reduction, face/object recognition and symbolic data analysis.

Dr. Lalitha Rangarajan is currently a Reader at the Department of Computer Science, University of Mysore, India. She has two master degrees one in Mathematics from Madras University, India and the other in Industrial Engineering (specialization: Operations Research) from Purdue University, USA. She has taught mathematics for five years in India soon after the completion of masters in mathematics. She is associated with the Department of Studies in Computer Science, University of Mysore, soon after completion of masters at Purdue University. She completed her doctorate in Computer Science in 2004 and since then doing research in the areas of image processing, retrieval of images, bioinformatics and pattern recognition. She has more than 30 publications in reputed conferences and journals to her credit.