2D-MELPP: A two dimensional matrix exponential based extension of locality preserving projections for dimensional reduction

Zixun Xiong[1,2] , Minghua Wan[1,2]*, Rui Xue[3] , and Guowei Yang[1,2]
[1] School of Information Engineering, Nanjing Audit University
Nanjing 211815,P.R. China
[2]Jiangsu Key Laboratory of Image and Video Understanding for Social Safety, Nanjing University of Science and Technology, Nanjing, 210094, P.R. China
[3] School of Electronics and Information Engineering, Beihang University
Beijing, P. R. China
[e-mail :181090918@stu.nau.edu.cn, wmh36@sina.com, xuerui@buaa.edu.cn, ygw_ustb@163.com]
*Corresponding author: Minghua Wan

Received December 15, 2020; revised July 14, 2021; accepted August 30, 2021; published September 30, 2022

Abstract

Two dimensional locality preserving projections (2D-LPP) is an improved algorithm of 2D image to solve the small sample size (SSS) problems which locality preserving projections (LPP) meets. It's able to find the low dimension manifold mapping that not only preserves local information but also detects manifold embedded in original data spaces. However, 2D-LPP is simple and elegant. So, inspired by the comparison experiments between two dimensional linear discriminant analysis (2D-LDA) and linear discriminant analysis (LDA) which indicated that matrix based methods don't always perform better even when training samples are limited, we surmise 2D-LPP may meet the same limitation as 2D-LDA and propose a novel matrix exponential method to enhance the performance of 2D-LPP. 2D-MELPP is equivalent to employing distance diffusion mapping to transform original images into a new space, and margins between labels are broadened, which is beneficial for solving classification problems. Nonetheless, the computational time complexity of 2D-MELPP is extremely high. In this paper, we replace some of matrix multiplications with multiple multiplications to save the memory cost and provide an efficient way for solving 2D-MELPP. We test it on public databases: random 3D data set, ORL, AR face database and Polyu Palmprint database and compare it with other 2D methods like 2D-LDA, 2D-LPP and 1D methods like LPP and exponential locality preserving projections (ELPP), finding it outperforms than others in recognition accuracy. We also compare different dimensions of projection vector and record the cost time on the ORL, AR face database and Polyu Palmprint database. The experiment results above proves that our advanced algorithm has a better performance on 3 independent public databases.

Keywords: Feature extraction, Discriminant analysis, Matrix exponential, Small sample size (SSS) problems, Two dimensional locality preserving projections (2D-LPP)

http://doi.org/10.3837/tiis.2022.09.009 **ISSN : 1976-7277**

1. Introduction

In pattern recognition, massive high dimensional image inputs are computationally challenging to analysis[1-6]. We hope to retain as much information as possible while reducing data dimensions to improve the accuracy and efficiency of later processing such as inference [7-10]. Generally, dimensionality reduction methods can be classified into two classes: feature extraction and feature selection. The former is focused on creating new features and the later is focused on selecting most relevant subsets. And there are two feature extractions classifications: (1)Nonlinear methods and (2)Linear methods. Nonlinear ones like locality linear embedding(LLE)[11], Laplacian egienmaps(LE)[12] pay attention to preserving local structure, while other methods like Isomap[13] is focused on global structure. It also tries to maintain geodesic distances between different samples. Linear methods, on the other hand, has become extremely important for its simple and computationally efficient classification strategies. Principal component analysis(PCA)[14-15] as unsupervised method and linear discriminant analysis(LDA)[16-17] as supervised method are two representative linear methos dimensionality reduction. locality preserving projections(LPP)[18], which can be seen as linearization of LE, is an alternative to PCA. LPP is obtained during the process of searching for the optimized approximate value of eigenfunctions of Laplace Beltrami operator on the manifold when there is a low dimensional manifold embedded in relatively high dimensional spaces. It needs to point out that LPP possesses two non linear advantages while former linear algorithms don't:

1.New minimal criterion preserves neighborhood structure, considering that Euclidean distances are meaningful only when they are local. It makes sense particularly when data lies on nonlinear manifold embedded in the ambient space.

2.For the 'neighborhood structure preserved' advantage above, LPP is also suitable for information retrieval applications

It is also very important that LPP is defined on more than just training data points and thus can be simply applied to new data points compared with traditional non-linear methods.

Two dimensional principal component analysis(2D-PCA)[19] saves the trouble of transforming images into one dimensional vectors and it inspires subsequent matrix based algorithm. And two dimensional linear discriminant(2D-LDA)[20] analysis extracts proper features directly from images based on Fisher's linear discriminant analysis. In the process especially like image recognition, the data is usually insufficient and having high dimensions causing singularity problems which make algorithms (e.x. LPP) unable to be applied directly. Two dimensional locality preserving projections (2D-LPP) is therefore proposed. It works on 2D images directly. Two dimensional locality preserving projections(2D-LPP)[21-23]perfectly solves the singularity that traditional LPP faced and is proved to perform better on face recognition and palmprint recognition , it turns out to be faster than LPP and achieves higher recognition rates.

Matrix based exponential method is also used to solve such small sample sized problem (SSS problems). ELDE[24] is based on local discriminant embedding (LDE)[25], which was put forward to overcome the limits that global linear discriminant analysis method meets. Matrix based exponential method also helps to improve the flaw of LPP's sensitivity to neighborhood sized k[26] and fix singularity problems.

According to[27], 2D methods don't necessarily perform better than 1D methods. It implies the limitations that 2D method meets and the direction for feature work. 2D-LPP is simple and efficient, but it don't perform well when training samples are limited and it seems to meet limitations in some public databases. We apply matrix exponential methods on 2D-LPP to improve its behavior and we get its enhanced version--2D-MELPP. However, the computational complexity of 2D-MELPP is extremely high, so we replace some of matrix multiplications with multiple multiplications to get a more efficient way for solving 2D-MELPP. We also use a more efficient way to calculate matrix exponential to save time. The new 2D-MELPP has high memory costs, so we replace some matrix multiplications with multiple multiplications and we investigate the enhanced version on a random 3D data set, ORL, AR Face Database and Polyu Palprint Database, 2D-PCA,2D-LDA and 2D-LPP are also tested.

The remainder of paper is arranged as follows: The mainly content is about reference to related works in Section 2. Section 3 is the retrospect of innovative work in this paper. Section 4 is composed of 4 experiments designed to show the efficiency and accuracy of our new algorithm. Section 5 is separated into two parts, respectively describing conclusions of experiments in section 4 and the direction of our feature work.

2. Related Work

Our work is based on 2D-LPP and matrix exponential methods, which both help to solve SSS problems and improve recognition accuracy.

2.1 2D-LPP method

2D-LPP applies the projection for each data as follows:

$$x_i = A_i w (i = 1,2,3...n) \tag{1}$$

$$x \in R^{r \times d} \quad A_i \in R^{r \times c} \quad w \in R^{c \times l}$$

where r ,c are row and column dimension of A_i **r**espectively. w is the target projection matrix a, while A_i is the ith source image for testing, and n is the number of images in the datasets. It chooses[21] the criterion as：

$$\min \sum_{i,j} S_{ij} \| X_i - X_j \|^2 \; (i, j = 1,2,3...) \tag{2}$$

is generated as： S_{ij}

(1). First, we have to construct adjacency matrix $G \in R^{m \times n}$. For every G_{ij} both k-nearest neighbors and ε - nearest neighbors are appropriate. When doing the experiment in this paper, we choose the former and add label Information;

(2). For each G_{ij} not equal to zero, we suppose ith and jth image are “close” ;

(3). For all “close” ones, we apply the Heat Kernel method to build the S matrix.

Looking back to the origin criterion, it can be written as:

$$\min \sum_{i,j} \| A_i w - A_j w \|^2 \; (i, j = 1,2,3...) \tag{3}$$

where the target function can be simplified[23] as

$$\min w^T A^T (L \otimes I_n) A w \tag{4}$$

where A is in the form like $[A_1^T, A_2^T ... A_n^T]^T$, D_{ii} is column sum of S, I_n is a $n \times n$ identity matrix and $\otimes$ is kronecker product.
To eliminate an arbitrary scaling element of the process, 2D-LPP dispose it as followed :

$$\sum_i D_{ij} X_i^T X_i = 1 \tag{5}$$

which can also be presented as:

$$w^T A^T (D \otimes I_n) A w = 1 \tag{6}$$

And the whole minimization problem is transformed into a classical generalized eigenvalue problem as follows:

$$A^T (L \otimes I_n) A w = \lambda A^T (D \otimes I_n) A w \tag{7}$$

So we can get final target projection matrix w.

2.2. Matrix exponential Process

In the sub-section, we will briefly look into the definition and some basic properties of matrix exponential process for the prerequisite of 2D-MELPP:

The Matrix exponential for M($M \in R^{m \times n}$) is defined as follows:

$$\exp(M) = 1 + M + \frac{M^2}{2!} + ... + \frac{M^m}{m!} + ... \tag{8}$$

where 1 can be treated as an $n \times n$ identity matrix, the matrix exponential has properties as followed:

(1). $\exp(M)$ is the sum of a sequence of matrixes with finite numbers;

(2). $\exp(M)$ is a full rank matrix;

(3). For arbitrary square matrix M , there exists the inverse of its matrix exponential;

(4). Supposing that R is a nonsingular matrix, we can have:

$$\exp(R^{-1} M R) = R^{-1} \exp(M) R$$

(5). For every eigenvectors of $M(v_1, v_2 ... v_n)$ corresponding to $\alpha_1, \alpha_2 ..., \alpha_n$ there exists $e^{\alpha_1}, e^{\alpha_2} ..., e^{\alpha_n}$ as eigenvalues of exp(M) having the same eigenvectors just like M.

3. 2D-Matrix exponential based discriminate locality preserving projection（2D-MELPP）

3.1. Notation used in 2D-MELPP

Notation	Meaning
W	0-1 affinity (weight) matrices
D	Column sum of W
L	Laplacian matrix, where $L = D - W$
A	Image set matrix
w	Projection matrix
λ	Eigen value corresponding to w
$\exp()$	Matrix exponential function
$\otimes$	kronecker product
P	P comes from SVD of A, where $A = P\Sigma V^T$
I_n	I_n is the n by n identity matrix
H	Where $H = D \otimes I_n$
M_H	$M_H = A^T Q_H \Sigma_H^{0.5}$,where $\Sigma_H^{0.5}$ comes form eigen-decomposition of H
Q_M^{τ}	Q_M^{τ} is column orthogonal matrix
L	Where $L = L \otimes I_n$
X_L	$X_L = V^T Q_L \Sigma_L^{0.5}$ where $\Sigma_L^{0.5}$ comes form eigen-decomposition of L
P_M^{τ}	P_M^{τ} is column orthogonal matrix

3.2. 2D-MELPP without further improvements

In 2D-MELPP, the objective function is:

$$\min w^T A^T \exp(L \otimes I_n) A w \tag{9}$$

To eliminate an arbitrary scaling element of the process, 2D-LPP dispose it as followed :

$$w^T A^T \exp(D \otimes I_n) A w = 1 \tag{10}$$

With Matrix based exponential $D \otimes L$ operator included accordingly.
Then the whole optimal question is equal to the eigenvalue problem as follows：

$$\exp(A^T (L \otimes I_n) A) w = \lambda \exp(A^T (D \otimes I_n) A) w \tag{11}$$

Referring to the properties (5) in section B, it is mathematically similar and can solve SSS problem in classical eigenvalue questions for (4) in section B.
According to[26], it can also emphasize the geometry features by the same time.

3.3. Efficient procedure to solve 2D-MELPP

The computational complexities of 2D-MELPP is computed $\exp(A^T(L \otimes I_n)A)$ and $O(r^3)$ respectively. A paper[28] meets the same problem as working with one dimension framework as follows:

$$w = \arg\min(w^T \exp(XDX^T)w)^{-1}(w^T \exp(XLX^T)w) \tag{12}$$

To solve this, we can apply the similar strategy.
As what it's shown in the part followed, an efficient procedure formula is build to solve 2D-MELPP.
First, we assume the SVD decomposition of A is

$$A = P\Sigma V^T \tag{13}$$

Let $S_a = P^T \exp(A^T(L \otimes I_m)A)P$, $S_b = P^T \exp(A^T(D \otimes I_m)A)P$, suppose that the eigenvector is u and λ is the eigenvalue.

Then we obtain:

$$\exp(A^T(D \otimes I_n)A)Pu = \lambda \exp(A^T(L \otimes I_n)A)Pu \tag{14}$$

According to (4) the computational complexity is reduced. However, the computational complexity of computing $\exp(A(D \otimes I_n)A)$ and $\exp(A^T(L \otimes I_n)A)$ is still $O(d^3)$. Then, out target has changed into compute $\exp(A^T(D \otimes I_n)A)$ and $\exp(A^T(L \otimes I_n)A)$ efficiently.

The eigen-decomposition of $H = D \otimes I_n$ is supposed :

$$H = Q_H \Sigma_H Q_H^T \tag{15}$$

Let $M_H = A^T Q_H \Sigma_H^{0.5}$,and its economic decomposition is

$$M_H = Q_M \Sigma_M V_M^T \tag{16}$$

Then we obtain:

$$A^T HA = Q_M \Sigma_M^2 Q_M^T \tag{17}$$

Let column orthogonal matrix Q_M^τ be that $[Q_M, Q_M^\tau]$ is orthogonal.

According to the definition, we can have that:

$$\exp(A^T HA) = [Q_M, Q_M^\tau][Q_M, Q_M^\tau]^T + \dots \tag{18}$$

$$+\frac{[Q_M, Q_M^{\tau}]\begin{bmatrix}\Sigma_M^2 & O_{r_M\times(r-r_M)}\\ O_{(r-r_M)\times r_M} & O_{(r-r_M)\times(r-r_M)}\end{bmatrix}}{m!}+\dots$$

$$=Q_M\exp(\Sigma_M^{\tau})Q_M^T+1-Q_MQ_M^T$$

where $Q_{c\times d}$ are zeros matrices, c,d are its row and column numbers respectively. And r_M is the dimension of Σ_M .

Then, we have:

$$P^T\exp(A^THA)P=P^T\exp(\Sigma_H^2)Q_M^TP+I-P^TQ_MQ_M^TP \tag{19}$$

Similarly, the eigen-decomposition of $L=L\otimes I_n$ is assumed:

$$L=Q_L\Sigma_LQ_L^T \tag{20}$$

Let $X_L=V^TQ_L\Sigma_L^{0.5}$, and its economic decomposition is:

$$X_L=Q_X\Sigma_XV_X^T \tag{21}$$

Then we obtain:

$$A^TLA=PQ_X\Sigma_X^2Q_X^TP^T \tag{22}$$

Let column orthogonal matrix P_M^{τ} be that $[PQ_X, P_X^{\tau}]$ is orthogonal. Then we have:

$$\exp(A^TLA)=[PQ_X, PQ_X^{\tau}][PQ_X, PQ_X^{\tau}]^T+$$

$$\dots+\frac{[PQ_X, PQ_X^{\tau}]\begin{bmatrix}\Sigma_M^{2m} & O_{r_x\times(d-r_x)}\\ O_{(r-r_x)\times r_x} & O_{(r-r_x)\times(r-r_x)}\end{bmatrix}}{m!}$$

$$=PQ_X\exp(\Sigma_X^2)Q_X^TP^T+P_X^{\tau}(P_X^{\tau})^T \tag{23}$$

Then we have:

$$P^T\exp(A^TLA)P=Q_X\exp(\Sigma_X^2)Q_X^T+I-Q_XQ_X^T \tag{24}$$

Now we have the following new process to solve 2D-MELPP:

Algorithm :2D-MELPP
Input: sample A
Output: The projected matrix of G

(1). Constructing the matrix A, $H=\exp(A^T DA)$, $L=\exp(A^T LA)$ just like normal 2D-LPP;

(2). Calculating the SVD of A as $A=P\Sigma V^T$ and the matrix H as $H=Q_H\Sigma_H Q_H^T$;

(3). Compute the matrix $M_H=A^T Q_H \Sigma_H^{0.5}$ and its economic SVD as $M_H=Q_M\Sigma_M V_M^T$;

(4). Compute the eigen-decomposition of L as $L=Q_L\Sigma_L Q_L^T$;

(5). Calculate the matrices $X_L=V^T Q_L \Sigma_L^{0.5}$,and its economic decomposition is $X_L=Q_X\Sigma_X V_X^T$;

(6). Compute the matrix

$J=P^T\exp(A^T HA)P=P^T\exp(\Sigma_H^2)Q_M^T P^T+P^T Q_M Q_M^T P$

$U=P^T\exp(A^T LA)P=Q_X\exp(\Sigma_X^2)Q_X^T+I-Q_X Q_X^T$ respectively;

(7). Solve the eigenvalue problem $Ju_i=\lambda_i Uu_i$;

(8). Let G = PU_k where U_k is the matrix combined of k eigenvectors corresponding to k largest eigenvalues, and then orthogonalize the projection matrix and get the G we wanted(k is a constant value) .

Then the time complexity is reduced to O(rn^2).

4. Experimental results

In the experiment, we will evaluate the performance of 2D-MELPP with PCA+LPP, 2D-PCA, ELPP, 2D-LDA, and 2D-LPP on the random 3D data set and three different public image databases. We pick up KNN as the classifier for three image databases. Four experimental results are completed in the same experimental environment (CPU: 2.04 GHz, RAM: 3.9GB).

4.1 The Random 3D data set

In the first experiment, we will propose 2D-MELPP and other methods on a synthetic 3D matrix M. The data set is made up of 20 matrices, each of which is generated by a 20 by 20 matrix having 400 points subject to normal distribution. And the data set is divided into 2 class as follows(The data set is normalized) :

$$class(A_{kij})=\begin{cases} class-a & A_{kij}\geq 0.999 \\ class-b & A_{kij}\leq 0.001 \end{cases} \tag{25}$$

where A_{kij} is value of ith row and jth column of the kth matrix, $(0 \le i, j, k \le 20)$, $class(A_{kij})$ is the class of A_{kij} and the class is either class-a or class-b. The distribution of two classes can be seen in **Fig. 1**.

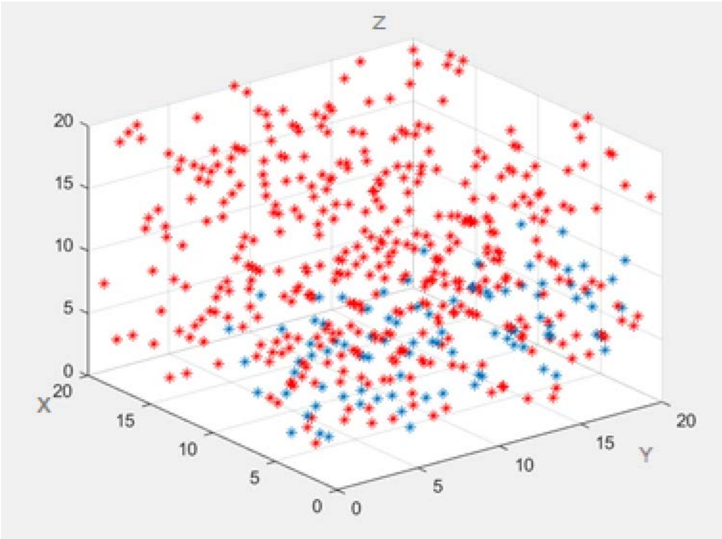
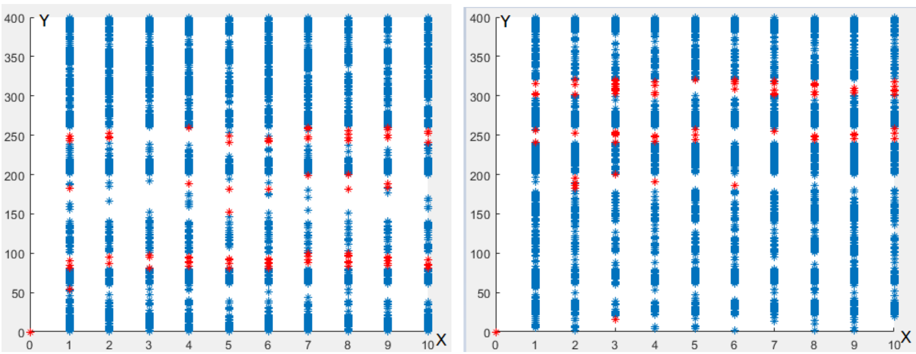
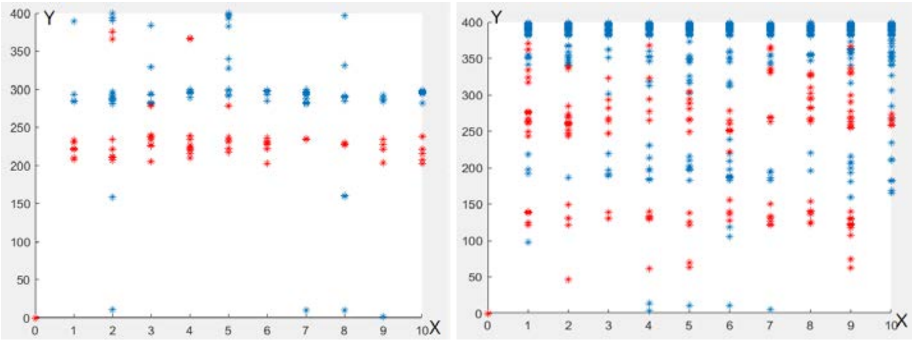


Fig. 1. The class distribution of 3D data set. Each point is represented by (x_i, y_i, z_j) in Cartesian coordinates. Variable i ranges from 0 to 20. Two classes are marked by blue and red respectively. Each point is assigned with random values ranging from 0 to 1 and the rules to decide its class are shown in (25).

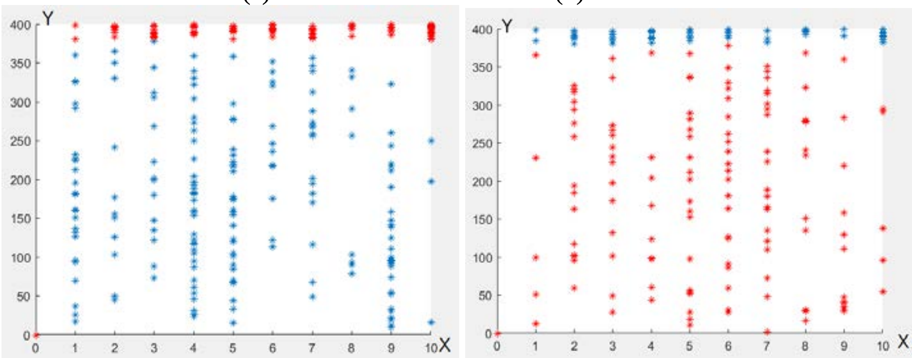
We test PCA+LPP, ELPP, 2D-PCA, 2D-LDA, 2D-LPP and 2D-MELPP on this random data set. And FIGURE 2 shows the result of dimension reduction with PCA+LPP, 2D-PCA, ELPP, 2D-LDA, 2D-LPP and 2D-MELPP (from left to right, up to down).



(a) PCA+LPP (b) ELPP



(c) 2D-PCA (d) 2D-LDA



(e) 2D-LPP (f) 2D-MELPP

Fig. 2. The projected 2D data set with (a) PCA+LPP, (b) ELPP, (c) 2D-PCA, (d) 2D-LDA, (e) 2D-LPP and (f) 2D-MELPP (from left to right, up to down). The X axis is the reduced dimension, Y is the serial number of the picture.

As can be seen from **Fig. 2**, 2D-MELPP provide projected data with excellent linear separability, while other methods fail to do so.

4.2 The ORL face database

The ORL face image database consists of 400 face images with 40 classes, each having 10 samples within. The background were all set in homogeneous darkness, and the images are featured by different light intensities, facial expressions and facial details(glassed and not) as it can be shown in **Fig. 3.**

Fig. 3. The images of one person from ORL face database

We range dimensions of project vector from 1 to 20 in **Table 1** and pick one with highest recognition accuracy each(from 2 to 6 training samples each class). **Fig. 4 s**hows recognition accuracy varies under 6 samples with the changes of dimensions. In **Table 2**, we record the time cost on each algorithm under 6 samples on the ORL face database.

Table 1. highest mean value of rECOGNITION ACCURACY (IN PERCENT) ON ORL database. There is a one-to-one correspondence between colunms and fixed training numbers. values in each PARENTHESIS ranges between 1 and 20, they represent for corresponding projected matrix dimensions.

Algorithm\Train num per class	2 Train	3 Train	4 Train	5 Train	6 Train
PCA+LPP	62.75(20)	73.75(20)	78.25(20)	83.25(20)	87.25(20)
ELPP	57.5 (20)	73.25(20)	79.75(20)	84.00(17)	89.25(20)
2D-PCA	60.63(20)	76.07(20)	76.67(20)	78.00(20)	81.88(20)
2D-LDA	73.89 (8)	84.38 (2)	90.00 (3)	90.8(10)	93.50(9)
2D-LPP	71.56 (5)	89.64(17)	85.00(12)	92 (16)	95.00(17)
2D-MELPP	**75.31(11)**	**91.79(18)**	**92.92(19)**	**95.00(16)**	**98.50 (7)**

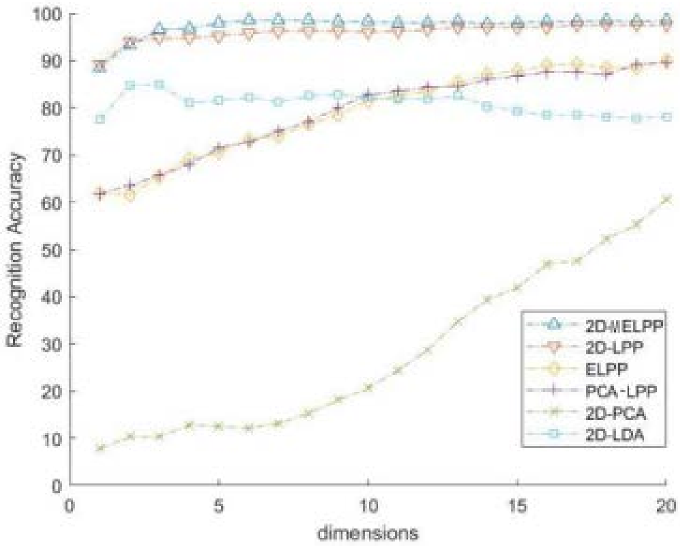


Fig. 4. The recognition accuracy with 6 samples per class on the ORL face database

Table 2. AVERAGE cost time(s) of PCA+LPP, ELPP, 2D-PCA, 2D-LDA, 2D-LPP and 2D-MELPP on The ON ORL FACE SET. EACH COLUMN CORRESPONDS TO A FIXED ALGORITHM.And THE 2D-MELPP IS THE IMPROVED VERSION.

Algorithm	PCA+LPP	ELPP	2D-PCA	2D-LDA	2D-LPP	2D-MELPP
Time(s)	0.861	1.365	2.343	2.494	0.898	**1.303**

4.3 The AR face database

AR face database is made up of face images of 120 people with 26 images of each. In this experiment we take 10 each which don't have too many occlusion problems while testing algorithm on a lager scale than ORL in **Fig. 5**.

Fig. 5. 20 image samples of one person form AR database. Pictures in second line were token 2 weeks after the first line

We range dimensions of project vector from 1 to 20 in **Table 3** and pick one with highest recognition accuracy each(from 2 to 6 training samples each class). **Fig. 6** shows recognition accuracy varies under 6 samples per class with the changes of dimensions.

Table 3. highest mean values of rECOGNITION ACCURACY (IN PERCENT) ON Ar database. There is a one-to-one correspondence between colunms and fixed training numbers. values in each PARENTHESIS ranges between 1 and 20, they represent for corresponding projected matrix dimensions.

Algorithm\Train num per class	2 Train	3 Train	4 Train	5 Train	6 Train
PCA+LPP	30.42(20)	48(20)	47.17(20)	55.5(20)	65.17(20)
ELPP	64.5(20)	55.41(19)	56.75(19)	66.75(20)	74.08(20)
2D-PCA	47.75(20)	55.5(20)	62.58(20)	72.42(20)	79(20)
2D-LDA	80.83(20)	75.41(19)	75.91(20)	91.67(19)	94.75(2)
2D-LPP	89.75(20)	86.33(20)	82.16(12)	93.5(17)	96.75(19)
2D-MELPP	**75.67(1)**	**87.42(2)**	**87.83(5)**	**91.75(11)**	**98.33(7)**

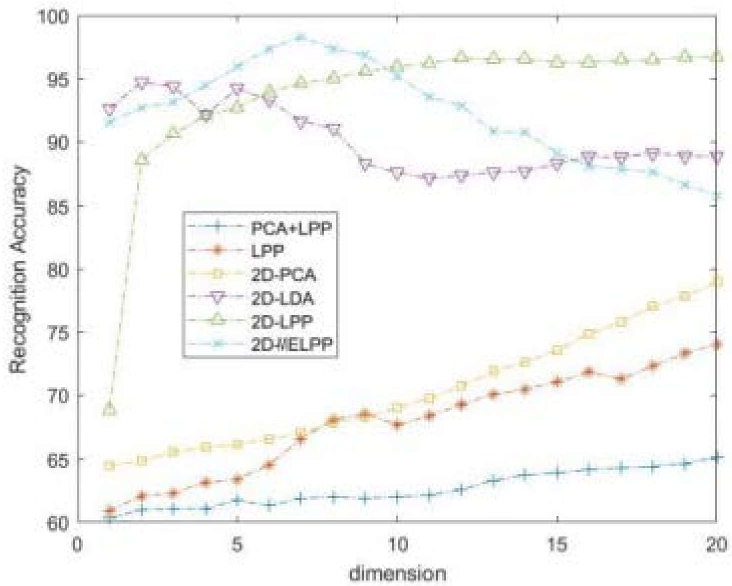


Fig. 6. The recognition accuracy with 6 samples per class on the AR face database

In **Table 4**, we record the time cost on each algorithm under 6 samples on the AR face database.

Table 4. AVERAGE cost time(s) of PCA+LPP, ELPP, 2D-PCA, 2D-LDA, 2D-LPP and 2D-MELPP on The ON ar FACE SET. EACH COLUMN CORRESPONDS TO A FIXED ALGORITHM.And THE 2D-MELPP IS THE IMPROVED VERSION.

Algorithm	PCA+LPP	ELPP	2D-PCA	2D-LDA	2D-LPP	2D-MELPP
Time(s)	4.723	6.686	4.201	4.5385	6.521	**6.516**

4.4. The Polyu palmprint database

Polyu Palmprint Database consists of 600 palmprint image with 100 classes, each having 6 samples within. **Fig. 7** shows the background were also set in homogeneous darkness and the images are featured by different light intensities. Half of the original image is taken 2 month later than the other half, and we set the training and testing division according to that.

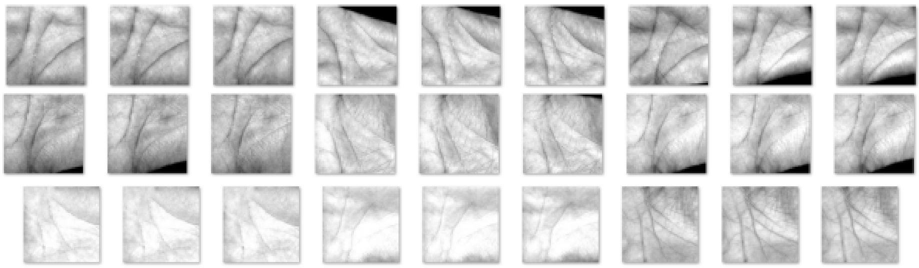


Fig. 7. The palmprint samples from Polyu palmprint database

Fig. 8 shows recognition accuracy varies under 3 samples with the changes of dimensions ranging from 1 to 50, and **Table 5** shows highest recognition accuracy of each algorithm, **Table 6** records the time cost with 50 dimensions.

Table 5. highest mean value of rECOGNITION ACCURACY (IN PERCENT) ON polyu database. There is a one-to-one correspondence between colunms and methods.

Algorithm	PCA+LPP	ELPP	2D-PCA	2D-LDA	2D-LPP	2D-MELPP
Recognition accuracy(%)	87(47)	96.33(48)	73.3(47)	90.33(5)	97(34)	**97.33(47)**

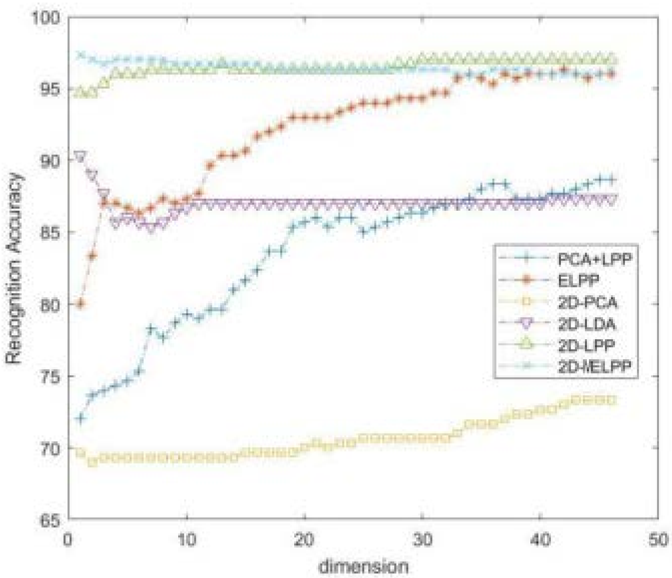


Fig. 8. recognition accuracy with 3 samples per class on Polyu Palmprint database

Table 6. AVERAGE cost time(s) of PCA+LPP, ELPP, 2D-PCA, 2D-LDA, 2D-LPP and 2D-MELPP on PolyU palmprint database. EACH COLUMN CORRESPONDS TO A FIXED ALGORITHM.And THE 2D-MELPP IS THE IMPROVED VERSION.

Algorithm	PCA+LPP	ELPP	2D-PCA	2D-LDA	2D-LPP	2D-MELPP
Time(s)	1.521	7.77	2.377	2.903	7.474	**7.680**

From results of the experiment above, we can make observations：

(1). According to **Fig.s 2, 4 and 6**, provided with identical experimental conditions, the recognition accuracy of the advanced algorithm has better performances. The ability of the algorithm to preserve more discrimination than others may account for it.
(2). Generally, 2D-MELPP achieves higher recognition accuracy than other algorithms with small projection dimensions as shown in **Table 1, 3, 5**.
(3). 2D-MELPP takes less time than ELPP in tree databases, which provides supports for our development on computational complexities does reduce the it. It can be seen from **Table 2, 4, 6**.

5. Conclusion

5.1. Results and advanced nature of experiments

We are inspired by the idea of ELPP and comparison works between 2D-LDA and LDA to employee matrix exponential on 2D-LPP and work algebra procedure to promote its behavior both in theory and in test on random 3D data set, two public face databases and one palmprint database. We also develop a way to save memory cost for solving 2D-MELPP. Our advanced 2D-MELPP performs better in recognition accuracy and it cost less time than ELPP.
2D-MELPP is an efficient solution for the SSS problem elegantly since the matrix exponential of a symmetric matrix is positive definite all the time. Since 2D-MELPP is based directly on image matrix and the matrix exponential process has distances enlarged, it can preserve more information and achieve higher accuracy.

5.2. Spectrum of future work

Although it's easy to apply linear methods for dimensionality reduction, it's nonetheless hard to attain linear data in real world feature extraction application. In view of above-mentioned reason, improvements can be made with nolinear methods or struct based methods. We will verify them in future work.

Acknowledgement

This work is partially supported by the National Science Foundation of China under Grant Nos. 61876213, 62172229, 6177227, 61976117, 71972102, 61861033, 61976118, the Key R&D Program Science Foundation in Colleges and Universities of Jiangshu Province Grant Nos.18KJA520005, 19KJA360001, 20KJA520002, the Natural Science Fund of Jiangsu Province under Grants Nos. BK20201397, BK20211295, BK20221349. Jiangsu Key Laboratory of Image and Video Understanding for Social Safety of Nanjing University of Science and Technology under Grants J2021-4. Funded by the Qing Lan Project of Jiangsu University, Funded by the Future Network Scientific Research Fund Project SRFP-2021-YB-25, and China's Jiangxi Province Natural Science Foundation (No. 20202ACBL202007). This work is funded in part by the "Qinglan Project" of Jiangsu Universities under Grant D202062032.

References

[1] Jia, Y. T., Law, S. S., & Yang, N., "Response and reliability analysis of a high-dimensional stochastic system," *Probabilistic Engineering Mechanics*, vol. 61, 2020. Article (CrossRef Link)

[2] Wei Z, Guo W, Li B, et al, "High-dimensional metric combining for non-coherent molecular signal detection," *IEEE Transactions on Communications*, Vol.68, No. 3, pp.1479-1493, 2019. Article (CrossRef Link)

[3] Pevný, Tomáš, Filler, Tomáš, & Bas, Patrick, "Using high-dimensional image models to perform highly undetectable steganography," in *Proc. of International Workshop on Information Hiding*, pp.161-177, 2010. Article (CrossRef Link)

[4] Jia, Y. T., Law, S. S., & Yang, N., "Response and reliability analysis of a high-dimensional stochastic system," *Probabilistic Engineering Mechanics*, 103077, 2020. Article (CrossRef Link)

[5] Bin, Gu, Victor, S, Sheng, & Keng, et al, "Incremental support vector learning for ordinal regression," *IEEE Transactions on Neural Networks & Learning Systems*, vol. 26, no. 7, pp. 1403-1416, 2015. Article (CrossRef Link)

[6] Gu, B., Sun, X., & Sheng, V. S., "Structural minimax probability machine," *IEEE Transactions on Neural Networks & Learning Systems*, Vol.28, No.7, pp.1646-1656, 2017. Article (CrossRef Link)

[7] Laurens Van Der Maaten, Eric Postma, & Jaap Van Den Herik, "Dimensionality reduction: a comparative review," *Review Literature & Arts of the Americas*, Vol.10, No.1, 2009. Article (CrossRef Link)

[8] Chapelle, O., Bernhard Schölkopf, & Zien, A., "Spectral Methods for Dimensionality Reduction," *Semi-Supervised Learning, MIT Press*.

[9] Yan, S., Xu, D., Zhang, B., Zhang, H. J., Yang, Q., & Lin, S., "Graph embedding and extensions: a general framework for dimensionality reduction," *IEEE Transactions on Pattern Analysis & Machine Intelligence*, Vol.29, pp. 40-51, 2007. Article (CrossRef Link)

[10] Zhang, T., Tao, D., Li, X., & Yang, J., "Patch alignment for dimensionality reduction," *IEEE Transactions on Knowledge & Data Engineering*, Vol.21, No.9, pp.1299-1313, 2009. Article (CrossRef Link)

[11] Roweis, Sam, T., Saul, Lawrence, & K, "Nonlinear dimensionality reduction by locally linear embedding," *Science*, vol. 290, n0, 5500, pp. 2323-2326, 2000. Article (CrossRef Link)

[12] Belkin, M., & Niyogi, P., "Laplacian eigenmaps for dimensionality reduction and data representation," *Neural Computation*, vol. 15, no. 6, pp. 1373-1396, 2003. Article (CrossRef Link)

[13] J. B. Tenenbaum, V. de Silva, and J. C. Langford, "A global geometric framework for nonlinear dimensionality reduction," Science, vol. 290, no. 5500, pp. 2319–2323, 2000. Article (CrossRef Link)

[14] Turk, M., "Eigenfaces for recognition," *J. Cognitive Neuroscience*, 3(1), 71-86, 1991. Article (CrossRef Link)

[15] Kirby, M., & Sirovich, L., "Application of the karhunen-loeve procedure for the characterization of human faces," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.12, No.1, pp.103-108, 1990. Article (CrossRef Link)

[16] Swets, D. L., & Weng, J. J., "Using discriminant eigenfeatures for image retrieval," *IEEE Transactions on Pattern Analysis & Machine Intelligence*, Vol.18, No.8, pp.831-836, 1996.

[17] Belhumeur, P. N., Hespanha, J. P., & Kriegman, D. J., "Eigenfaces vs. fisherfaces: recognition using class specific linear projection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.19, No.7, pp.711-720, 1997.

[18] He, X., "Locality preserving projections," *Advances in Neural Information Processing Systems*, Vol.16, No.1, pp.186-197, 2003. Article (CrossRef Link)

[19] Yang, J., Zhang, D. D., Frangi, A. F., & Yang, J. Y., "Two-dimensional pca: a new approach to appearance-based face representation and recognition," *IEEE Trans Pattern Anal Mach Intell*, Vol.26, No.1, pp.131-137, 2004. Article (CrossRef Link)

[20] Li, M., & Yuan, B., "2d-lda: a statistical linear discriminant analysis for image matrix," *Pattern Recognition Letters*, Vol.26, No.5, pp.527-532, 2005. Article (CrossRef Link)

[21] Chen, S., Zhao, H., Kong, M., & Luo, B., "2d-lpp: a two-dimensional extension of locality preserving projections," *Neurocomputing*, Vol.70, No.4-6, pp.912-921, 2007. Article (CrossRef Link)

[22] Zhi, R., & Ruan, Q., "Facial expression recognition based on two-dimensional discriminant locality preserving projections," *Neurocomputing*, No.71, Vol.7-9, pp.1730-1734, 2008. Article (CrossRef Link)

[23] Hu, D., Feng, G., & Zhou, Z., "Two-dimensional locality preserving projections (2dlpp) with its application to palmprint recognition," *Pattern Recognition*, Vol.40, No.1, pp.339-342, 2007. Article (CrossRef Link)

[24] Ran, R., Fang, B., Xuegang, W. U., & Zhang, S., "A simple and effective generalization of exponential matrix discriminant analysis and its application to face recognition," *ICE Transactions on Information and Systems*, E101.D(1), pp.265-268, 2018. Article (CrossRef Link)

[25] Chen, H. T., Chang, H. W., & Liu, T. L., "Local discriminant embedding and its variants," in *Proc. of CVPR, IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2, pp.846-853, 2005. Article (CrossRef Link)

[26] Wang, S. J., Chen, H. L., Peng, X. J., & Zhou, C. G., "Exponential locality preserving projections for small sample size problem," *Neurocomputing*, Vol.74, No.17, pp.3654-3662, 2011. Article (CrossRef Link)

[27] Zheng, W. S., Lai, J. H., & Li, S. Z., "1d-lda vs. 2d-lda: when is vector-based linear discriminant analysis better than matrix-based?," *Pattern Recognition*, 41(7), 2156-2172, 2008. Article (CrossRef Link)

[28] Lu, G. F., Wang, Y., Zou, J., & Wang, Z., "Matrix exponential based discriminant locality preserving projections for feature extraction," *Neural Networks*, vol. 97, pp.127-136, 2018. Article (CrossRef Link)

ZIXUN XIONG received the bachelor's degree in infromation institute from the Nanjing Audit University in 2022. He is a student at the University of Massachusetts Amherst. His current research interests include reinforcement learning and graphics.

MINGHUA WAN received the B.S. degree in automated institute from the Nanchang University of Aviation in 2003, the M.S. and PhD degrees in pattern recognition and intelligent systems from the Nanjing University of Science and Technology (NUST) in 2007 and 2011, respectively. He is the author of more than 50 scientific papers in pattern recognition and computer vision. He is a professor at the Nanjing Audit University. His current research interests include face recognition and detection, and image processing.

RUI XUE received the B.S. degree in electronic engineering and the Ph.D. degree in transportation information and control engineering from Beihang University, in 2001 and 2010, respectively. He is currently a Professor with the National Key Laboratory of CNS/ATM, Beihang University. His research interests include error modeling, estimation, and validation.

GUOWEI YANG received the B.S. and the M.S. degrees in mathematics from the Jiangxi Normal University, Nanchang, China, in 1985 and 1988, and the Ph.D. degree from the University of Science and Technology Beijing, China, in 2004. He was appointed as Professor in the Qingdao University, China, in 1999. Now, he is a professor in the School of Technology of Nanjing Audit University (NAU). His current research interests include artificial intelligence, artificial life, artificial neural network, pattern recognition, innovative and creative design, etc.