# Local reconstruction and local fisher discriminant based semi-supervised dimensionality reduction algorithm

**Xiangyang Chen***

*School of Electronic Engineering, University of Electronic Science and Technology, Chengdu, Sichuan Province, China, 611731*

*School of Mathematics and Computing Sciences, Guilin University of Electronic Technology, Guilin city, Guangxi Province, China, 541004*

*Received 28 October 2014, www.cmnt.lv*

**Abstract**

Local reconstruction and global preserving based semi-supervised dimensionality reduction (LRGPSSDR) algorithm gives no consideration to data locality when processing intra-class relationship and class relationship. Enhanced semi-supervised local fisher discriminant analysis algorithm (ESELF) also neglects locality of data manifold structure when maintaining data manifold structure. To address these problems, the local reconstruction and local fisher discriminant based semi-supervised dimensionality reduction (LRLFSDR) algorithm was proposed in this paper. It depicts significance of sample distance with an improved thermonuclear weight. In this way, intra-class relationship and class relationship of the same cluster attracts more attentions, thus enabling to shorten or widen intra-class distance or class distance firstly. Moreover, it uses idea of LLE algorithm to make neighbourhood linear reconstruction relationship of each point in low-dimensional space to be similar with that in high-dimensional space, which takes locality of data manifold structure into account. Test result confirmed that the proposed LRLFSDR algorithm is superior to other semi-supervised dimensionality reduction algorithms in classifying standard libraries like COIL20, Extended YaleB and CMU PIE.

*Keywords:* local fisher discriminant, local reconstruction, semi-supervised learning, dimensionality reduction

## 1 Introduction

"Curse of dimensionality" caused by "small data sample but high dimensionality" is a common problem in machine learning, data mining and pattern recognition. Dimensionality reduction is an effective solution to the "curse of dimensionality" [1-3]. It includes supervised dimensionality reduction and unsupervised dimensionality reduction according to whether training data provides dimensionality reduction algorithm on label. Supervised dimensionality reduction algorithm can make full use of data category structure to apply dimensionality reduced data better in data classification. Linear discriminant analysis (LDA) [4] is a typical supervised dimensionality reduction algorithm. Unsupervised dimensionality reduction algorithm reduces data dimensionality mainly based on data manifold structure or sample relationship. It is mainly used as pre-processing of clustering or supervised dimensionality reduction algorithm. Principal component analysis (PCA) [4] and local linear embedding (LLE) [5] are typical unsupervised dimensionality reduction algorithms.

In most practical application cases (e.g. face recognition and speaker recognition), supervised dimensionality reduction algorithm requires many labelled samples, which is very difficult. On the contrary, it is very easy to acquire unlabelled samples. Under this background, many scholars began to study how to use unlabelled samples to increase effect of supervised dimensionality reduction algorithm, such as semi-supervised discriminant analysis (SDA) [6] and LRGPSSDA [8].

Advantages and disadvantages of these algorithms are:

1) SDA [6] can embody discriminant information in labelled samples and maintain manifold structure of all samples. However, it is sensitive to neighbourhood parameter setting and limits dimensions of the projection subspace within sample categories.

2) LRGPSSDA [8] has advantages of SDA and overcomes two disadvantages of SDA. It moves faraway sample points farther so that dimensions of the projection subspace are not limited within sample categories basically. However, this is not overcome completely because faraway samples are less significant to classification result in many classification algorithms.

3) LRGPSSDR [7] has advantages of SSDR and overcomes some disadvantages of SSDR. However, it uses global covariance structure to describe negative and positive constraints, but neglects local structure. To address these problems, LRLFSDR algorithm was proposed in this paper. Improved from ESELF and LRGPSSDR, it has two advantages. Firstly, LRLFSDR algorithm requires neighbour linear reconstruction relationship of each point in low-dimensional space to be similar with that in high-dimensional space to maintain data manifold structure. Different from ESELF, it takes locality of data manifold structure into account. Secondly, LRLFSDR algorithm uses weighted covariance matrix instead of intra-class scattering matrix, thus promising

* ***Corresponding author*** e-mail: xiangyangchen0428@163.com

high effect on objective function optimization even under small intra-class sample size. This is different from LRGPSSDR algorithm.

## 2 Related algorithms

The general form of semi-supervised dimensionality reduction algorithm is given in this section because LRLFSDR algorithm is a semi-supervised dimensionality reduction algorithm. Later, two related algorithms related to the LRLFSDR algorithm are introduced briefly: LRGPSSDR and ESELF.

### 2.1 GENERAL FORM OF SUPERVISED DIMENSIONLITY REDUCTION ALGORITHM

Most of existing semi-supervised dimensionality reduction algorithms can be depicted by the following general mathematical description. Given n labelled samples $\{x_1, x_2, ...x_n\} \in R^m$ and their labels $\{l_1, l_2, ...l_n\}$ as well as $p$ unlabelled samples $\{x_1, x_2, ...x_p\} \in R^m$, semi-supervised dimensionality reduction algorithm implements mapping of $(X \in R^m) \to (X \in R^d)$, $d \ll m$ using labelled samples and unlabelled samples. To achieve mapping, most of existing semi-supervised dimensionality reduction algorithms have the following or similar objective function [1-3, 6-8]:

$$W^* = \arg \max \frac{W^T SbW}{W^T SwW + \lambda W^T SmW}, \tag{1}$$

where $W$ is the mapping matrix of data from high-dimensional space to low-dimensional space. $W^T SbW$ is to make data points of different classes in the low-dimensional space as far as possible. $W^T SwW$ is to make data points of same data class in the low-dimensional space as close as possible. $W^T SmW$ is to make data points close to each other in high-dimensional space still close to each other in low-dimensional space. Different semi-supervised dimensionality reduction algorithms have different definitions to above three items or add some new items in Equation (1).

### 2.2 LRGPSSDR ALGORITHM

The objective function of LRGPSSDR algorithm [7] is:

$$W^* = \arg \max_{W^T W = 1} \frac{W^T SbW + \lambda_1 W^T SfW}{W^T SwW + \lambda_2 W^T SmW}, \tag{2}$$

where $W^T SfW$ is to make data points away from each other in high-dimensional space still far away from each other in low-dimensional space. Rest items are for similar purpose with those in Equation (1). To be more specific:

$$W^T SbW = \sum_{l_i \neq l_j} \left(W^T x_i - W^T x_j\right)^2, \tag{3}$$

$$W^T SwW = \sum_{l_i = l_j} \left(W^T x_i - W^T x_j\right)^2, \tag{4}$$

$$W^T SfW = \sum_{x_i \notin N(x_j) \text{ and } x_i \notin N(x_j)} \left(W^T x_i - W^T x_j\right)^2, \tag{5}$$

where $N(x_i)$ is the domain of $x_i$. It can be known from Equations (3)-(5) that $W^T SbW$, $W^T SwW$ and $W^T SfW$ are the distance sum of different point sets in low-dimensional space. Then, $W^T SmW$ can be defined as:

$$W^T SmW = \sum_{x_i \notin N(x_j) \text{ and } x_i \notin N(x_j)} \left(W^T x_i - W^T x_j\right)^2, \tag{6}$$

However, LRGPSSDR algorithm [7] doesn't use Equation (6). This is because Wei Jia stated in Reference [7] that in LLE algorithm [5], Roweis et al. assume that the neighbourhood of input space is locally linear. In other words, each point in high-dimensional space can be reconstructed through linear combination of the points in its neighbourhood and requires similar linear reconstruction relationship in low-dimensional space. Under this hypothesis, LLE algorithm achieved satisfying test result. As a result, Wei Jia [7] redefined $W^T SmW$ based on this hypothesis:

$$W^T SmW = W^T XMX^T W, \tag{7}$$

where $M = (I - A)^T (I - A)$. $A$ can be calculated from:

$$\varepsilon_1(A) = \sum_i \left\| x_i - \sum_{j:x_j \in N(x_i)} A_{ij} x_j \right\|. \tag{8}$$

If $\sum_{j:x_j \in N(x_i)} A_{ij} = 1$, the linear relationship between local neighbourhoods ($A$) can be calculated by least square method.

Based on difference among Equations (3)-(5) and (7), it fails to achieve good dimensionality reduction effect by defining items in Equation (2) with distance sum of different point sets in low-dimensional space. Equations (3)-(6) reflect that distances between samples are equally significant. This will cause evident shortcomings in many cases.

1) For class scattering matrix in low-dimensional space defined by maximum Equation (3), further increase of the distance between two faraway samples is less significant to classification.

2) For intra-class scattering matrix in low-dimensional space defined by minimum Equation (4), further decrease of the distance between two faraway samples is also less significant to classification. When data of same class has many distribution intervals in high-dimensional space, decreasing distance between these data in low-dimensional space will affect optimization of Equation (2) significantly, thus making the mapping matrix not always beneficial to

classification. Similar disadvantages exist in Equations (5) and (6).

The LRGPSSDR algorithm still has most of abovementioned disadvantages although it has overcome some by using Equation (7) instead of Equation (6), most are remained.

### 2.3ESELF ALGORTHM

ESELF noticed that distances between different samples shall be given with different weights, which can overcome some disadvantages of LRGPSSDR algorithm. A brief introduction of ESELF algorithm is made in this section. Similarly, the objective function of ESELF algorithm is given:

$$W^* = \arg\max_{W^TW=1} \frac{W^T SbW - W^T SwW}{W^T StW}. \tag{9}$$

Reference demonstrates that $W^T SbW$ in (9) is to maximize distance between data points of different classes within certain high-dimensional neighbourhood in low-dimensional distance. $W^T SwW$ is to minimize distance between data points of same class within certain high-dimensional neighbourhood in low-dimensional distance. $W^T StW$ is to make relationships of data points in low-dimensional space similar with those in high-dimensional space. To be more specific:

$$W^T SbW = \sum_{l_i \neq l_j} (W^T x_i - W^T x_j)^2 S_{ij}^b, \tag{10}$$

$$W^T SwW = \sum_{l_i = l_j} (W^T x_i - W^T x_j)^2 S_{ij}^w, \tag{11}$$

$$W^T StW = \sum (W^T x_i - W^T x_j)^2, \tag{12}$$

where $S_{ij}^b$ and $S_{ij}^w$ in Equations (10) and (11) are weight of relationship between $i$ and $j$. Generally speaking, the further the $i$ and $j$ are, the smaller the $S_{ij}^b$ and $S_{ij}^w$ will be. Equations (10) and (11) give certain consideration to data locality when defining $W^T SbW$ and $W^T SmW$. However, ESELF algorithm still neglects locality of data manifold structure and its data manifold structure in high-dimensional space goes against local linear data.

## 3 LRLFSDR algorithm

This section introduces the proposed LRLFSDR algorithm which can overcome disadvantages of both ESELF and LRGPSSDR algorithms.

### 3.1 OBJECTIVE FUNCTION

Objective function of the proposed LRLFSDR algorithm is:

$$W^* = \arg\max_{W^TW=1} \frac{W^T SbW}{W^T StW + \lambda_2 W^T SmW}, \tag{13}$$

where $W^T SbW$ is defined in (10) and $W^T SmW$ is defined in (7). $W^T StW$ is defined as:

$$W^T StW = \sum \left(W^T x_i - W^T x_j\right)^2 s_{ij}^t, \tag{14}$$

where $s_{ij}^t = B_{ij}$ and,

$$B_{ij} = \exp\left(\frac{-d^2(x_i, x_j)}{\sigma_i^2 \sigma_j^2}\right), \tag{15}$$

where $\sigma_i$ and $\sigma_j$ are distances from $x_i$ and $x_j$ to their $k_c$ neighbouring samples. Different from SELF algorithm, the proposed LRLFSDR algorithm defines $S_{ij}^b$ in (10) as:

$$S_{ij}^b = \begin{cases} B_{ij} & if \ , l_i \neq l_j \\ 0, & others \end{cases}. \tag{16}$$

### 3.2 THEORETICAL BASIS

a) Equation (13) uses $W^T StW$ instead of $W^T SwW$ in Equation (2). This is because $W^T SwW$ influences denominator less when there are few intra-class training samples. To overcome disadvantages of $W^T SbW$ is added to both numerator and denominator of Equation (1) [6]:

$$W^* = \arg\max_{W^TW=1} \frac{W^T SbW + W^T SbW}{W^T SwW + W^T SbW + \lambda W^T SmW}. \tag{17}$$

Equation (17) can be simplified into Equation (13). Therefore, when the numerator is $W^T SbW$, $W^T SwW$ can be minimized by minimizing $W^T StW$.

b) Numerator of Equation (13) involves no $W^T SfW$. This is because $W^T SbW$ and $W^T SfW$ which represent sample distance repeat for many times when there's a small intra-class sample size. When there are abundant intra-class samples, $W^T SfW$ may have to be far away from the sample which belongs to the sample class but is beyond one point domain. This disagrees with the objective of $W^T StW$.

c) Equation (13) keeps data manifold structure by using Equation (7), which can overcome two disadvantages of ESELF algorithm, which uses covariance of all samples to keep data manifold structure.

d) Advantages of $W^T StW$ and $W^T SbW$ definitions in the proposed LRLFSDR algorithm are analysed. Weight of samples $B_{ij}$ has various definitions [6-8]

$$B_{ij} = 1, \tag{18}$$

$$B_{ij} = \exp\left(\frac{-d^2(x_i, x_j)}{\sigma}\right). \quad (19)$$

Equation (18) is 0-1 weight, while Equation (19) is thermonuclear weight and $\sigma$ is the variance of thermonuclear.

If $B_{ij}$ adopts 0-1 weight, Equation (10) is equal to (4). In other words, $B_{ij}$ in LRGPSSDR algorithm uses 0-1 weight, which is the cause of its disadvantages according to analysis in Section 2.2.

SDA algorithm can represent sample relationship with thermonuclear weight defined in Equation (19). However, thermonuclear weight has to set $\sigma$. Each dataset has only one $\sigma$, which is inadequate to describe all sample relationships in a dataset because many datasets have different data distribution patterns in different neighbourhoods.

To address this problem, Lihi Zelnik-Mano represent sample relationship by using Equation (15). It doesn't need to preset parameters and has two advantages. Firstly, weight of sample relationship changes with data distribution. Secondly, data clustering structure is more obvious. When calculating $W^T StW$ and $W^T SbW$ through such weight of sample relationship, priori attention will be paid to increase distance of samples belonging to different classes of same cluster during optimizing Equation (13). Moreover, attention also will be paid to distance between samples belonging to sample class in same cluster. As a result, samples of same class can form clusters in low-dimensional space rather than accumulate all samples together when optimizing Equation (13), thus making Equation (13) more conducive to data classification. The proposed LRLFSDR algorithm can overcome disadvantages of LRGPSSDR algorithm in Section 2.2.

Based on previous analysis, we can conclude that $W^T SbW$ in Equation (13) is mainly to maximize distance of data points belonging to different classes within certain high-dimensional cluster in low-dimensional space. $W^T StW$ is mainly to minimize distance of data points belonging to different classes within certain high-dimensional cluster in low-dimensional space. $W^T SmW$ is mainly to maintain the local linearity in high-dimensional space to low-dimensional space. Additionally, weights of sample relationship within certain cluster in $W^T StW$ and $W^T SbW$ are set zero, which maintains a stable global data structure. Apparently, objective of the proposed LRLFSDR algorithm is different from that of LRGPSSDR and ESELF algorithms. It overcomes disadvantages of both LRGPSSDR and ESELF algorithms.

### 3.3 OPTIMIZATION OF OBJECTIVE FUNCTION

Optimization of Equation (13) is a generalized rayleigh quotient problem. If $S_t + \lambda_2 S_m$ is non-singular, solution of Equation (13) is the eigenvector corresponding to the maximum generalized eigenvalue of Equation (20):

$$S_b w = \lambda(S_t + \lambda_2 S_m)w. \quad (20)$$

According to Equation (7),

$$Sm = XMX^T. \quad (21)$$

Then $S_b$ and $S_t$ can be calculated from Equations (10) and (14). For Equation (10),

$$\begin{aligned} &W^T SbW = \sum_{l_i \neq l_j} \left(W^T x_i - W^T x_j\right)^2 S_{ij}^b = \\ &2\sum_i W^T x_i D_{ii}^b x_i W - 2\sum_{ij} W^T x_j D_{ij}^b x_j W = \\ &2W^T X\left(D^b - S^b\right)X^T W = 2W^T XL^b X^T W, \\ &D_{ii}^b = \sum_j S_{ij}^b, L^b = D^b - S^b. \end{aligned} \quad (22)$$

$S_b$ can be calculated:

$$S^b = XL^b X^T. \quad (23)$$

For Equation (14),

$$\begin{aligned} &W^T StW = \sum\left(W^T x_i - W^T x_j\right)^2 S_{ij}^b = 2W^T XL^t X^T W, \\ &D_{ii}^t = \sum_j S_{ij}^t, L^t = D^t - S^t. \end{aligned} \quad . \quad (24)$$

$S_t$ can be calculated:

$$S^t = XL^t X^T. \quad (25)$$

Therefore, steps of the proposed LRLFSDR algorithm can be concluded:

Algorithm 1 LRLFSDR:

Input: $n$ labelled samples $\{x_1, x_2, ...x_n\} \in R^m$ and their labels $\{l_1, l_2, ...l_n\}$, $p$ unlabelled samples $\{x_1, x_2, ...x_p\} \in R^m$, neighbourhood parameter $k_m$ of Equation (8), and neighbourhood parameter $k_c$ of equation (15).

Output: Mapping matrix $W$.

Calculate weight of sample relationship (B) from equation (15).

Calculate $S_m$, $S_b$ and $S_t$ from Equations (21), (23) and (25).

Calculate eigenvalues and eigenvectors in Equation (20) and rank eigenvectors from high eigenvalue to small. Then, the mapping matrix $W$ can be gained.

## 4 Simulation experiment

Data classification is the main goal of supervised or semi-supervised dimensionality reduction algorithm. In this part, several classification experiments will be conducted to verify the effectiveness of the proposed LRLFSDR algorithm. All experiments use 1-NN classifier. The

proposed LRLFSDR algorithm is compared with SDA [6], PCA [4], LDA [4], LRGPSSDR [7] and ESELF algorithms. Only SDA, LRGPSSDR and the proposed LRLFSDR algorithms have to set parameters. The SDA algorithm can jump over parameter setting by using Equation (15) to represent weight of sample relationship. LRGPSSDR algorithm applies parameters recommended by Reference [7]: $\lambda_1 = \lambda_2 = 0.1$, $k_m = 1$ (the neighbourhood parameter of Equation (8)) and $k_f = 5$ (the neighbourhood parameter of Equation (5)). Parameters of the proposed LRLFSDR algorithm include $\lambda_2 = 0.1$, $k_m = 1$ and $k_c = 5$.

## 4.1 EXPERIMENTAL DATASET

Three actual datasets are used in this paper.

1) COIL20 image library. It includes 20 objects and each object has 72 images, totally 1,440 images. These images are zoomed into 32×32 in this paper. Therefore, the data dimension is 1024.

2) Extended YaleB. It is a face dataset, including 38*64 256-gray images with different illuminations. Each image is cut into 32×32.

3) CMU PIE. It is also a face dataset, including 41,368 face images of 68 persons. These images are collected under different postures, illuminations and expressions. Each image is cut into 32×32 and identified as 256-gray image.

## 4.2 EXPERIMENT SETTINGS

The experiment uses three algorithms, namely, unsupervised dimensionality reduction algorithm (e.g. PCA), supervised dimensionality reduction algorithm (e.g. LDA) and semi-supervised dimensionality reduction algorithm (e.g. SDA, LRGPSSDR, ESELF and the proposed LRLFSDR).

During the training process, unsupervised and semi-supervised dimensionality reduction algorithms can use both labelled and unlabelled training samples, while supervised dimensionality reduction algorithms can only use labelled ones. n labelled training samples and unlabelled training samples are selected randomly from each class for each experiment (Table 1). To analyse effect of the amount of unlabelled samples on the proposed LRLFSDR algorithm and other semi-supervised dimensionality reduction algorithms, experiments 1-3 have same amount of labelled samples, but their unlabelled samples increase successively. Experiments 4-6 have more labelled samples in order to discuss effect of the amount of labelled samples on the proposed LRLFSDR algorithm and other semi-supervised dimensionality reduction algorithms. Moreover, the amount of unlabelled samples in experiments 4-6 also changes to maintain same proportion of unlabelled samples with that in experiments 1-3. Each experiment is repeated for 50 times, taking the mean as the final result.

TABLE 1 Amount of labelled and unlabelled samples in different experiments

| n | Experiment 1 | Experiment 2 | Experiment 3 | Experiment 4 | Experiment 5 | Experiment 6 |
|---|---|---|---|---|---|---|
| $n_l$ | 5 | 5 | 5 | 10 | 10 | 10 |
| $n_{ul}$ | 3 | 6 | 9 | 5 | 15 | 20 |

## 4.3 CLASSIFICATION RESULT

Classification result is represented by accuracy in this paper [7]. The calculation formula of accuracy is:

$$precision = \frac{\#(ecorrectly\_recognized\_recordings)}{\#(all\ recognized\_recordings)}. \quad (26)$$

Classification accuracies of Experiments 1-3 under different dimensions are shown in Figure 1. Maximum classification accuracies under different dimensions are listed in Table 2. Figure 1a is Experiment 1's classification accuracy of COIL20 dataset. Figure 1b is Experiment 2 classification accuracy of COIL20 dataset. Figure 1c is Experiment 3 classification accuracy of COIL20 dataset. Figure 1d is Experiment 1 classification accuracy of CMU PIE dataset. Figure 1e is Experiment 2 classification accuracy of CMU PIE dataset. Figure 1f is Experiment 3 classification accuracy of CMU PIE dataset. Figure 1g is Experiment 1's classification accuracy of Extended YaleB dataset. Figure 1h is Experiment 2 classification accuracy of Extended Yale B dataset. Figure 1i is Experiment 3 classification accuracy of Extended Yale B dataset.

Experimental results reveal that the proposed LRLFSDR algorithm achieves better classification accuracy of all datasets compared to other algorithms. This proves effectiveness of the proposed LRLFSDR algorithm. Based on analysis of Experiments 1-3, the proposed LRLFSDR algorithm and LRGPSSDR algorithm improve the semi-supervised dimensionality reduction effect significantly by using certain unlabelled samples. However, when there are more unlabelled samples than labelled ones, further increase of unlabelled samples will greatly lower the classification accuracy of semi-supervised dimensionality reduction algorithms. Hence, the proposed LRLFSDR algorithm and LRGPSSDR algorithm shall choose reasonable amount of unlabelled samples. Moreover, the proposed LRLFSDR algorithm can achieve stable recognition rate at lower dimensions than other algorithms. This indicates that the proposed LRLFSDR algorithm can provide classifier lower-dimensional features, which can accelerate operation of the classifier significantly.

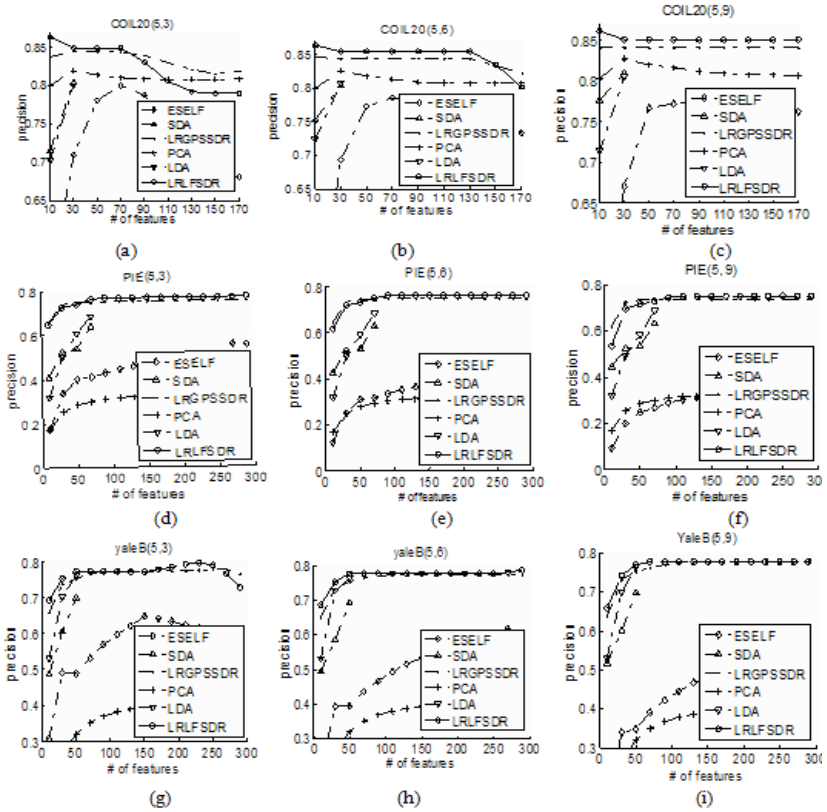


FIGURE 1 Experiments 1-3 classification accuracies of three datasets under different dimensions

TABLE 2 Experiments 1-3's maximum classification accuracies of three datasets

| | ESELF | SDA | LRGPSSDR | PCA | LDA | LRLFSDR |
|---|---|---|---|---|---|---|
| COIL20(1) | 80.60 | 80.41 | 84.44 | 81.94 | 80.06 | 86.32 |
| COIL20(2) | 78.66 | 80.41 | 84.63 | 82.61 | 80.84 | 86.43 |
| COIL20(3) | 77.28 | 81.04 | 84.12 | 82.72 | 80.52 | 86.16 |
| PIE(1) | 55.17 | 63.84 | 75.15 | 33.08 | 68.24 | 76.79 |
| PIE(2) | 40.83 | 63.08 | 75.15 | 32.71 | 68.29 | 76.45 |
| PIE(3) | 34.38 | 63.21 | 73.82 | 32.98 | 68.48 | 74.95 |
| YaleB(1) | 64.83 | 70.04 | 77.96 | 40.44 | 76.30 | 79.66 |
| YaleB(2) | 61.74 | 69.31 | 77.35 | 40.48 | 76.35 | 78.66 |
| YaleB(3) | 54.28 | 69.74 | 77.72 | 40.18 | 76.31 | 77.72 |

Maximum classification accuracies of Experiments 4-6 under different dimensions are listed in Table 3. It can be seen from Table 2 that the proposed LRLFSDR algorithm can offer optimal dimensionality reduction effect under most cases, indicating that it is superior to other algorithms under different amounts of labelled samples.

TABLE 3 Experiments 4-6's maximum classification accuracies of three datasets

| | ESELF | SDA | LRGPSSDR | PCA | LDA | LRLFSDR |
|---|---|---|---|---|---|---|
| COIL20(4) | 88.36 | 87.90 | 92.03 | 90.00 | 87.54 | 92.63 |
| COIL20(5) | 86.22 | 88.93 | 92.17 | 89.82 | 86.86 | 92.69 |
| COIL20(6) | 85.52 | 89.89 | 92.07 | 90.46 | 87.21 | 92.43 |
| PIE(4) | 61.73 | 79.94 | 85.21 | 47.27 | 74.01 | 87.43 |
| PIE(5) | 58.84 | 80.12 | 84.58 | 47.18 | 74.00 | 86.65 |
| PIE(6) | 57.13 | 80.65 | 85.26 | 47.30 | 74.03 | 86.32 |
| YaleB(4) | 79.57 | 86.28 | 90.43 | 56.19 | 87.23 | 90.35 |
| YaleB(5) | 70.17 | 86.57 | 90.07 | 56.37 | 87.05 | 90.50 |
| YaleB(6) | 63.99 | 86.95 | 89.92 | 55.84 | 87.44 | 90.25 |

## 5 Conclusions

Experimental results demonstrate that the proposed LRLFSDR algorithm can achieve good dimensionality reduction effect under low dimensions. However, it neglects samples beyond the neighbourhood domain when calculating local reconstruction coefficient. Meanwhile, it sets only one neighbourhood parameter to the whole dataset, which neglects locality of data distribution to a certain extent. Therefore, how to involve data locality into the building of reconstruction coefficient requires further research.

## Acknowledgement

This work was supported by the National Natural Science Foundation of China (Grant No.61261033, 41201479, 61062003 and 61162007) and Guangxi Natural science Foundation (Grant No.2013GXNSFBA019270).

## References


[1] Song De, Yao L 2013 Semi-supervised Dimensionality Reduction Based on Manifold Structure Preserving and Discriminative Locality Aligment *Microcomputer Applications* **30**(5) 17-20 *(in Chinese)*

[2] Du C, Sun J, Zhou S, Wang L, Zhao J 2013 Dimensionality reduction based on sparse representation and nonparametric discriminant analysis. *Journal of National University of Defense Technology* **35**(2) 143-7 *(in Chinese)*

[3] Liu H, Zhou C 2013 Dimensionality reduction based on sparse representation and nonparametric discriminant analysis *Computer engineering and design* **33**(2) 228-33 *(in Chinese)*

[4] Martinez A M, Kak A C 2001 PCA Versus LDA *IEEE Transactions on Pattern Analysis and Machine Intelligence* **23**(2) 228-33

[5] Roweis S T, Saul L K 2000 Nonlinear dimensionality reduction by locally linear embedding *Science* **290**(5500) 2323-6

[6] Cai D, He X, Han J 2007 Semi-Supervised Discriminant Analysis *IEEE 11th International Conference on Computer Vision 2007 (ICCV 2007)* 1-7

[7] Wei J, Peng H 2008 Local and global preserving based semi-supervised simensionality reduction method *Journal of software* **19**(11) 2833-42

[8] Wei J, Yang C, Ma Q et al. 2010 Semi-supervised siscriminant analysis method based on local reconstruction and global preserving. *Journal of South China University of Technology* **38**(7) 45-9 *(in Chinese)*


## Author


**Xiangyang Chen, born in January 1972, Guilin, Guangxi Province, P.R. China**

**Current position, grades:** lecturer at Guilin University of Electronic and Technology, Guangxi, China.
**Scientific interests:** multi-source information fusion and the optimization of sensor management.
**Publications:** more than 11 papers.
**Experience:** teaching experience of 16 years, 2 scientific research projects.