Journal of Theoretical and Applied Information Technology
31st July 2016. Vol.89. No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

FACE CLASSIFICATION USING WIDROW-HOFF LEARNING PARALLEL LINEAR COLLABORATIVE DISCRIMINANT REGRESSION (WH-PLCDRC)

[1] VISHWANATH P., [2]Dr. V M VISWANATHA
[1] NIMS University, Jaipur(Rajasthan), Electronics and Communication Engineering
[2] S L N College of Engineering, Raichur, Electronics and Communication Engineering
E-mail: [1]vishalpetli73@gmail.com, [2]vmviswanatha@gmail.com

ABSTRACT

Learning based face classification technique is proposed in this paper to improve Xiaochao Qu's linear collaborative discriminant regression classification (LCDRC). The LCDRC helps to find out a discriminant subspace by increasing collaborative between-class reconstruction error and decreasing the within-class reconstruction error simultaneously. With the aim of proposing the further improvement of the accuracy of probe image classification by minimizing the value of WCRC with the aid of our proposed methodology. Proposed methodology used parallel process to find an optimal projection matrix with the aid of the LCDRC method (arm A) & implement Widrow-Hoff learning method into the LCDRC (arm B). Both arms A & arm B are varying with the calculation of WCRC & arm A returns the current value of WCRC & arm B derives the future value of WCRC. Then, the optimal projection matrix is derived by the proposed methodology after the selection of suitable value of WCRC derived from the arm A & arm B method.

Keywords: *Face Classification, Linear Collaborative Discriminant Regression, Reconstruction Error, Widrow-Hoff Learning.*

1. INTRODUCTION

Biometric authentication has gotten a great deal of interest recently is due to the wide range of applications information security & access control. Face recognition system [1] is one of the methods for biometric authentication to identify from face images. The automatic face recognition system (FRS) can find the faces from still or video streams, which have been performed by the classifiers [2]. Research is focused to improve the ability of FRS by concentrating on the representation methodology, as a consequence, some algorithms are developed such as Sparse Representation Classification (SRC) [3], Collaborative Representation Classification (CRC) [4] & Linear Regression Classification (LRC) [5]. SRC & CRC convert the test images to the linear representation by utilizing every training images. The image classification is carried out by the calculation of a representation error of each class. Alternatively, the LRC makes the classification of the test images on the basis of an assumption of the face images fitting to the similar class in the particular space. With the help of the linear regression, the test face picture is arranged into a class. In the last couple of years, numerous upgrades to SRC, CRC & LRC have been proposed [6–11] [12].

Face recognition systems are known not basically reliant on complex learning strategies. Be that as it may, any endeavor at recognition in such a high dimensional space is powerless against an assortment of issues regularly alluded to as the course of dimensionality. Since the images are converted into low-dimensional vectors in the face space at the stage of face extraction. The primary goal is to discover such a basis function for this transformation which could noticeably speak to faces in the face space. Various methodologies for dimension reduction have been accounted in the literature such as Principle Component Analysis (PCA) [13], [14], Linear Discriminant Analysis (LDA) [15], & Independent Component Analysis (ICA) [5] [16], [17], [18]. With the help of dimension reduction methodology, the classification algorithm extracts the efficient features to make the better classification accuracy. The selected dimension reduced feature makes the classification algorithm become fast & accurate manner.

In any case, the scheme of the dimension reduction & classifier is normally isolated from one

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

another that may reduce the general performance of the face recognition system. Even though, selected feature of LDA does not fundamentally ensure a better performance of LRC for example, it is ideal to develop a dimension reduction system that is suitable for the classifier which will be utilized. According to the classification rule generated by the LRC, LDRC [2] selects low dimensional features that support to LRC for better fast & accurate. LRC utilizes the Fisher criterion to learn a discriminant subspace. The fisher criterion is fraction between the between-class reconstruction error (BCRE) over the within-class reconstruction error (WCRE) is maximized [12].

2. LITERATURE REVIEW

S. M. Huang & J. F. Yang [2] presented linear regression classification methodology with the help of class-specific representation where it was distinguished by Between-Class Reconstruction Error (BCRE) & Within-Class Reconstruction Error (WCRE) to discover a discriminant subspace by increasing the value of BCRE & decreasing the value of WCRE simultaneously. Main drawbacks of the LDRC was maximization of the overall BCRE was simply dominated by some large class-specific BCRE. This thing makes the subsequent LRC incorrect.

Xiaochao Qu et al [12] have presented a method for face recognition named as Linear Collaborative Discriminant Regression Classification (LCDRC). This method enhanced Huang's LDRC algorithm. Additionally, this research paper adopted a better BCRE measurement that was achieved by utilizing the collaborative representation rather than class-specific representation. The main drawback of LCDRC was that it was used the single linear regression model which was consisting of one predictor that leads to anomalous results in accuracy.

Xiaochao Qu et al [18] have presented an enhanced discriminant linear regression classification (EDLRC) algorithm to further improve the discriminant power of LDRC. They haven't used all those classes for calculating BCRE rather than they have only considered about the classes with small reconstruction error. Through maximizing the construction error of the true class's similar classes, their EDLRC increased the discriminatory power of LDRC. Their experiment showed that EDLRC performed better than LRC & LDRC for ORL & AR database.

Yan-li Liu et al [19] presented a linear regression based method by generating an extended set for a probe image. They produced the low dimension features for a probe. Additionally, they created virtual samples through summing up randomness into down sampling. The next phase was to classify the probe through utilizing canonical correlation analysis.

3. PROPOSED METHODOLOGY

The LRC algorithm is improved by a LDRC algorithm which is done by using the Fisher criterion. The LDRC used to increase the proportion of the BCRE over the WCRE to discover an optimum projection matrix (U) for the LRC. There are some drawbacks in LDRC. The problem is the maximization of the overall BCRE is easily interrupt by some large class-specific between-class reconstruction errors. In the past, LDRC was improved by the LCDRC with the aid of collaborative representation as a replacement for class-specific representation to obtain better BCRE measurement. The LCDRC utilized the collaborative between-class reconstruction error (CBCRE) rather than BCRE. The obtained CBCRE was smaller than each class-specific between-class reconstruction error. The maximizing of CBCRE tends to better separate the WCRE & the small class-specific BCRE than conventional BCRE. The problem in the LCDRC is to improve the BCRE with the collaborative method, which is not used much for the WCRE. In LCDRC, the classification error happens when the true class & false class have similar small reconstruction errors. In this paper, we have planned to improve the LCDRC with our proposed methodology that can help to reduce the WCRC into as small as possible. After that WCRC is used to calculate CBCRE that helps to classify the analysis of the image accurately.

Figure 1 represents the overall block diagram of the proposed WH-PLCDRC methodology. Initially the facial database is put into the feature extraction process which leads to convert the face image into appropriate format that is convenient for classification processing. Next stage is to derivation of projection matrix. The hat matrix is derived with the aid of the projection matrix to find the reconstruction matrix for each class. Finally, the classification is done by the calculation of reconstruction error of each class.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Subsequently face image is classified in the class that has the minimum reconstruction error.

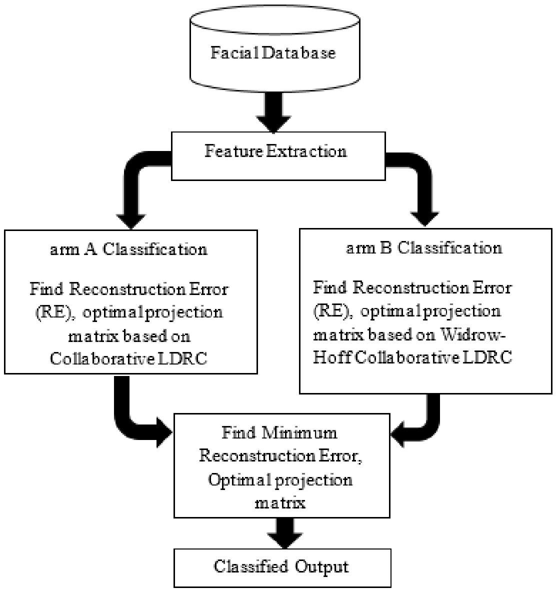

Figure 1: Block Diagram of proposed WH-PLCDRC

LDRC finds the optimal projection matrix "U" by minimizing the value of WCRC & increase the value of LCDRC. The optimal projection matrix "U" helps to increase the reconstruction error of false class & decrease the reconstruction error of true class, thus the discriminating ability of LDRC is better than the LRC. However, LDRC does not distinguish the similar classes & the dissimilar classes. So, all of those classes' reconstruction errors are maximized. The main objective of this research is to improve the accuracy of probe image classification by minimizing the value of WCRC with the help of our proposed methodology. Proposed methodology used parallel process to find an optimal projection matrix with the help of the LCDRC method (arm A) & implement Widrow-Hoff learning method into the LCDRC (arm B). Both arms A & arm B are varying with the calculation of WCRC. At the time, when arm A returns the current value of WCRC then arm B derives the future value of WCRC. After that, the optimal projection matrix is derived by the proposed methodology after the selection of suitable value of WCRC derived from the arm A & arm B method.

3.1 Linear Collaborative Discriminant Regression Classification (LDRC) by arm A

We denote the training face images of the ith class as $X_i \in \Re^{m\times n_i}$. Each column of X_i **Error! Bookmark not defined.**is a dimensional face image of class i in which there are n_i training face images, & $1 \le i \ge c$ where c is the total number of classes.

Consider y is the probe face image that can be represented using X_i according to

$$y = X_i \alpha_i \text{, where } 1 \le i \ge c \qquad (1)$$

$\alpha_i \in \Re^{n_i \times 1}$ is the regression parameters; α_i can be calculated using the least-square estimation as,

$$\hat{\alpha}_i = \left(X_i^T X_i\right)^{-1} X_i^T y \text{, } 1 \le i \ge c \qquad (2)$$

The reconstruction of y by each class can be obtained as,

$$\hat{y}_i = X_i \hat{\alpha}_i = X_i\left(X_i^T X_i\right)^{-1} X_i^T y = H_i y \text{, } 1 \le i \ge c \qquad (3)$$

Where H_i is called hat matrix that maps y into $\hat{y}$ the reconstruction error of each class is calculated as

$$e_i = \left\|y - \hat{y}_i\right\|_2^2 \text{, } \ldots\ 1 \le i \ge c \qquad (4)$$

LRC then assigns the y to the class that has the smallest reconstruction error

Consider the all training image matrix $X = [X_1, X_2, \ldots, X_c] \in \Re^{m\times n}$ where $X_i = [x_{i1}, x_{i2}, \ldots x_{in_i}] \in \Re^{m\times n_i}$, where m is the dimension of each training face image, n_i is the number of training face image from class i, & $n = \sum_{i=1}^{c} n_i$. Consider the subspace projection matrix to be learned is $U \in \Re^{m\times d}$ & $d < m$. Each x_{ij} could be mapped to the learned subspace by $y_{ij} = U^T x_{ij}$, where $1 \le j \le n_i$. The entire training face image matrix is thus mapped as $Y = U^T X \in \Re^{m\times d}$, and, for each class.

Journal of Theoretical and Applied Information Technology
31st July 2016. Vol.89. No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

$$Y_i = U^T X_i \in \Re^{d \times n_i} \tag{5}$$

$$e_i = \left| Y_i - \tilde{Y}_i \right|^2 \tag{6}$$

$$CBCRE = \frac{1}{n} \sum_{i=1}^{c} \sum_{j=1}^{n_i} \left\| y_{ij} - \hat{y}_{ij}^{\,inter} \right\|_2^2 ,$$

$$WCRE_{armA} = \frac{1}{n} \sum_{i=1}^{c} \sum_{j=1}^{n_i} \left\| y_{ij} - \hat{y}_{ij}^{\,intra} \right\|_2^2 \tag{7}$$

where $\hat{y}_{ij}^{\,inter} = Y_{ij}^{inter} \alpha_{ij}^{inter}$ & $\hat{y}_{ij}^{\,intra} = Y_{ij}^{inter} \alpha_{ij}^{inter}$. The value of represents Y_{ij}^{inter} the Y with Y_i eliminated & Y_{ij}^{intra} is the Y_i with y_{ij} eliminated. The value of α_{ij}^{inter} and α_{ij}^{intra} can be obtained from $\hat{\alpha}_i = \left(X_i^T X_i \right)^{-1} X_i^T y$ where $1 \le i \ge c$. Notice that prior to obtaining U, the value of α in the learned subspace is unknown to us. However, we can calculate $\hat{\alpha}$ in the original space & use $\hat{\alpha}$ as an approximation of α. It can be seen from the difference between CBCRE & BCRE is that CBCRE uses cross-class collaborative representation & BCRE uses class-specific representation from the definition of CBCRE in Equation 7.

According to the relationships between X & Y, CBCRE & WCRE can be rewriter as follows.

$$CBCRE = \sum_{i=1}^{c} \sum_{j=1}^{n_i} \left\| U^T x_{ij} - U^T X_{ij}^{inter} \alpha_{ij}^{inter} \right\|_2^2 ,$$

$$WCRE_{armA} = \sum_{i=1}^{c} \sum_{j=1}^{n_i} \left\| U^T x_{ij} - U^T X_{ij}^{intra} \alpha_{ij}^{intra} \right\|_2^2 \tag{8}$$

which can be further rewritten as

$$CBCRE = \sum_{i=1}^{c} \sum_{j=1}^{n_i} \left(x_{ij} - X_{ij}^{inter} \alpha_{ij}^{inter} \right)^T U U^T \left(x_{ij} - X_{ij}^{inter} \alpha_{ij}^{inter} \right)^T ,$$

$$WCRE_{armA} = \sum_{i=1}^{c} \sum_{j=1}^{n_i} \left(x_{ij} - X_{ij}^{intra} \alpha_{ij}^{intra} \right)^T U U^T \left(x_{ij} - X_{ij}^{intra} \alpha_{ij}^{intra} \right)^T \tag{9}$$

Notice that both CBCRE & WCRE have the factor of $1/n$, therefore, it is safe to eliminate $1/n$ from CBCRE & WCRE simultaneously without affecting the value of the ratio of CBCRE over WCRE. Under some algebraic deduction, CBCRE & WCRE can be written as:

$$CBCRE = \sum_{i=1}^{c} \sum_{j=1}^{n_i} tr\left(U^T \left(x_{ij} - X_{ij}^{inter} \alpha_{ij}^{inter} \right) \left(x_{ij} - X_{ij}^{inter} \alpha_{ij}^{inter} \right)^T U \right),$$

$$WCRE_{armA} = \sum_{i=1}^{c} \sum_{j=1}^{n_i} tr\left(U^T \left(x_{ij} - X_{ij}^{intra} \alpha_{ij}^{intra} \right) \left(x_{ij} - X_{ij}^{intra} \alpha_{ij}^{intra} \right)^T U \right) \tag{10}$$

Where $tr(\cdot)$ is the trace operator. E_b and E_w are defined as follows

$$E_b = \frac{1}{n} \sum_{i=1}^{c} \sum_{j=1}^{n_i} \left(x_{ij} - X_{ij}^{inter} \alpha_{ij}^{inter} \right) \left(x_{ij} - X_{ij}^{inter} \alpha_{ij}^{inter} \right)^T ,$$

$$E_w = \frac{1}{n} \sum_{i=1}^{c} \sum_{j=1}^{n_i} \left(x_{ij} - X_{ij}^{intra} \alpha_{ij}^{intra} \right) \left(x_{ij} - X_{ij}^{intra} \alpha_{ij}^{intra} \right)^T , \tag{11}$$

Eventually, CBCRE & WCRE can be rewritten as,

$$CBCRE = tr\left(U^T E_b U \right),$$

$$WCRE_{armA} = tr\left(U^T E_w U \right) \tag{12}$$

3.2 Widrow-Hoff Learning Linear Collaborative Discriminant Regression Classification (LCDRC) by arm B

The basic principle of Linear Collaborative Discriminant Regression Classification (LCDRC) in parallel arm B can be simply explained as follows: LCDRC adopts Widrow-Hoff learning rule; it needs a set of given input vectors & their corresponding expected output vectors; each input has a corresponding network output; the projection matrix weights are adjusted as per the difference between the quantities of input & the expected output; making sure the quadratic sum of training error is minimum.

The linear machine learning method adopts the Widrow-Hoff learning rule, the projection matrix weights & threshold values that are revised by learn W-H function [11]. The Widrow-Hoff learning rule can be used to train the weights in one parallel arm of our system. So as to make their linearity approximates a functional expression. Considering the equation 6, it can be seen from that the linear regression has a linear error curve & therefore, it must have only one

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

minimum error amount. Since e (w, b) is only dependent upon the projection matrix weights & feature vectors, the minimum error value can be reached by adjusting the weights. The Widrow-Hoff learning rule continuously adjusts the weights by fastest descending. According to the gradient descent, the variation of weight vectors is directly proportionate to the current position's grad of e (w, b). As for the input point

$$\frac{\partial e}{\partial U^T} = -2(y_i - U^T x_i)x_i^T \tag{13}$$

$$\frac{\partial e}{\partial U^T} = -\frac{\partial U^T}{\partial t},$$

$$U^T{}_{i+1} = U^T{}_i + \eta_i (y_i - U^T x_i)x_i^T \tag{14}$$

The above two Equation 13 and 14 are the Widrow-Hoff learning rule. They are also called the Minimum Mean Square Error Method. The change of the weights of the Widrow-Hoff learning rule is in direct proportion to the within the class reconstruction error (WCRE) & feature vectors. In this algorithm, it is not needed to get a derivative, so it is quite simple & has the benefits of a high speed of convergence & a high accuracy. The "η_i" in the above formula is the learning rate. When the learning rate is higher than the learning & convergence speeds are faster. But when "ŋ" is excessively big then the learning process is unstable & the error is bigger. Therefore, a proper learning rate should be selected. We got equation 15 after substituting equation 14 using equation 5.

$$Y_{i+1} = U_{i+1}{}^T X_i \in \Re^{d \times n_i} \tag{15}$$

Eq (7-12) rewritten for arm B,

$$WCRE_{armB} = \frac{1}{n}\sum_{i=1}^{c}\sum_{j=1}^{n_i}\left\| y_{ij} - \hat{y}_{ij}^{\text{int}\,ra} \right\|_2^2 \tag{16}$$

$$WCRE_{armB} = \sum_{i=1}^{c}\sum_{j=1}^{n_i}\left\| U_{i+1}{}^T x_{ij} - U_{i+1}{}^T X_{ij}^{\text{int}\,ra} \alpha_{ij}^{\text{int}\,ra} \right\|_2^2 \tag{17}$$

$$WCRE_{armB} = \sum_{i=1}^{c}\sum_{j=1}^{n_i} tr\left(U_{i+1}{}^T \left(x_{ij} - X_{ij}^{\text{intra}}\alpha_{ij}^{\text{intra}}\right)\left(x_{ij} - X_{ij}^{\text{intra}}\alpha_{ij}^{\text{intra}}\right)^T U_{i+1}\right) \tag{18}$$

$$E_w = \frac{1}{n}\sum_{i=1}^{c}\sum_{j=1}^{n_i}\left(x_{ij} - X_{ij}^{\text{int}\,ra}\alpha_{ij}^{\text{int}\,ra}\right)\left(x_{ij} - X_{ij}^{\text{int}\,ra}\alpha_{ij}^{\text{int}\,ra}\right)^T, \tag{19}$$

$$WCRE_{armB} = tr\left(U^T E_w U\right) \tag{20}$$

Find effective WCRE Arm value,

$$WCRE = \begin{cases} WCRE_{armA} \le WCRE_{armB} \; WCRE = WCRE_{armA} \;\&\; U = U_i \\ WCRE_{armA} \ge WCRE_{armB} \; WCRE = WCRE_{armB} \;\&\; U = U_{i+1} \end{cases} \tag{21}$$

The derivation of CBCRE is same as LCDRC as followed.

To maximize CBCRE & minimize WCRE simultaneously, the maximum margin criterion (MMC) is adopted to maximize the following criterion.

$$\max_U \; J(U) = tr(CBCRE - WCRE) \tag{22}$$

$$\max_U \; J(U) = \max_U \left(tr\left(U^T (E_w - E_w) U\right)\right) \tag{23}$$

The equation 23 can be solved through deriving the maximum d eigenvalues, based on that eigenvectors are as following

$$(E_b - E_W) u_k = \lambda_k u_k, \; 1 \le k \ge d \tag{24}$$

$$\lambda_k = \eta_F \lambda_k \; 1 \le k \ge d$$

where η_F is final learn rate of Widrow-Hoff method which helps to optimize the Lagrange multipliers λ_k.

4. EXPERIMENTAL RESULT AND DISCUSSION

4.1 Experiments on ORL, YALE, Extended YALE B

To verify the effectiveness of the proposed algorithm, we compared the performances of the proposed LCDRC to those of WH-PLCDRC on the famous face databases: ORL. The sample face images are shown in Figure 2. All the experiments randomly select part of the database for training & the remainder for test. Each experiment is repeated 10 times & the average result is reported. PCA is used to extract features.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Figure 2: ORL Face Data

The regularization parameter used in CRC is set to 0.001. Thel1 l s [26] package is used in SRC with the tolerance set to 0.01. Before implementing proposed & LCDRC, PCA is also used to reduce the dimension of the face images. MMC [21] is used both in LDRC & LCDRC such that the performance difference comes from BCRE & CBCRE.

The ORL face database contains 400 face images of 40 individuals with 10 face images for each subject. The face images were taken under different light conditions & with different facial expressions. All the face images are cropped to be 32x32 in our experiment. Fig. 3 shows the performance of the proposed & LCDRC methods given two, & four for each class.

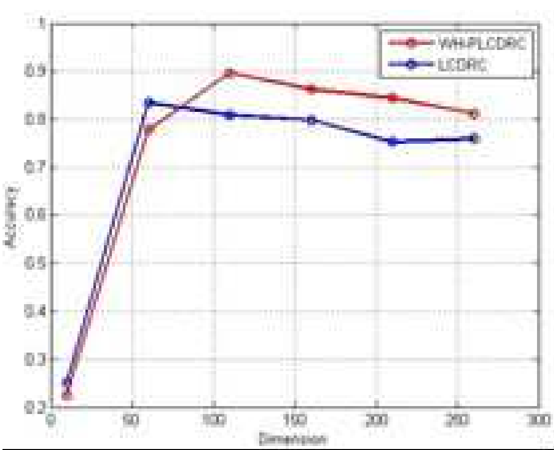

Figure 3: Two Train ORL Database Accuracy

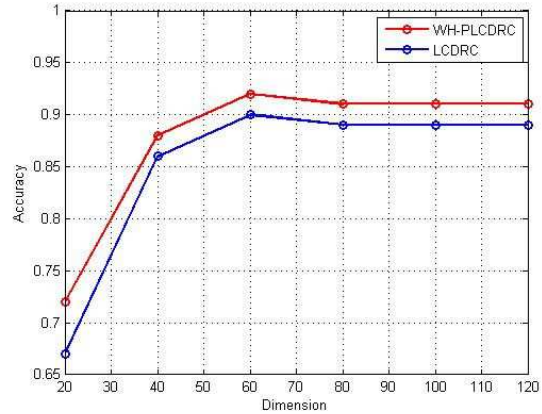

Figure 4: Three Train ORL Database Accuracy

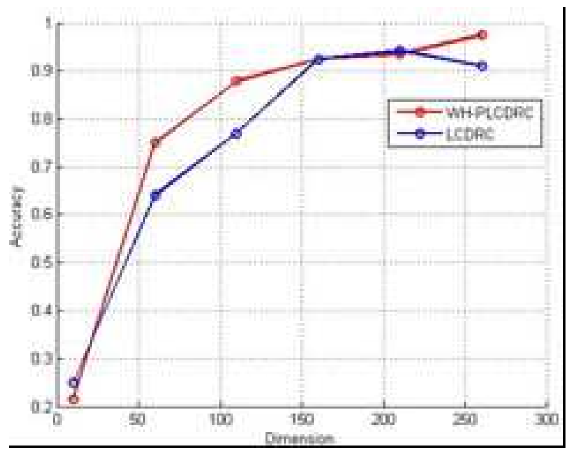

Figure 5:Four Train ORL Database Accuracy

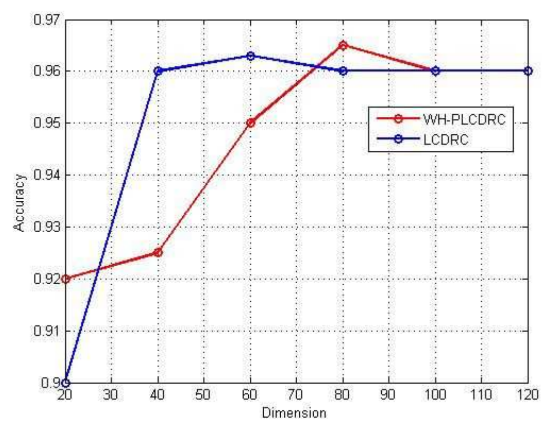

Figure 6: Five Train ORL Database Accuracy

Journal of Theoretical and Applied Information Technology
31st July 2016. Vol.89. No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

4.2 Experiments on YALE

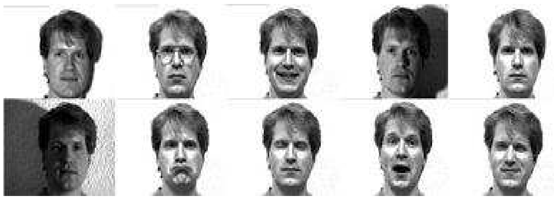

Figure 7: YALE Face Data

The YALE face database contains 165 faces of 15 individuals. For each individual, there are 11 images under different facial expressions or configurations. All of the face images are cropped into a size of 32x32. The performance comparison of different methods is shown in Fig. 5. WH-PLCDRC has much higher classification accuracy than the other method. For example, the best recognition rates & the corresponding feature dimensions of CRC, SRC, LRC, LDRC, LCDRC, WH-PLCDRC given four training face images for each class are 70.86%(50), 69.24%(50), 61.62%(50), 66.48% (50), 77:05%(30) & 86.6% respectively.

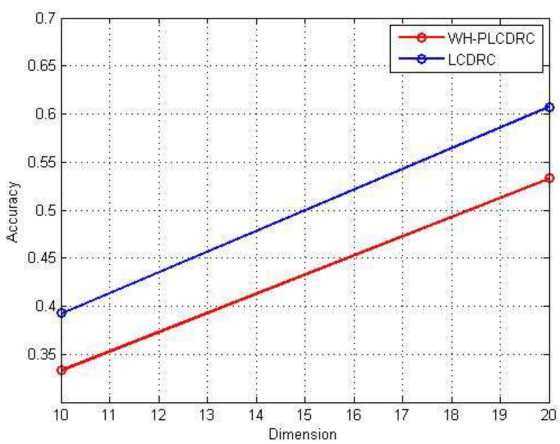

Figure 8: Two Train YALE Database Accuracy

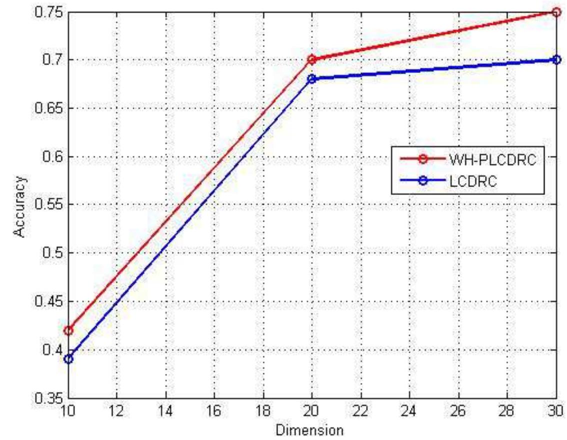

Figure 9: Three Train YALE Database Accuracy

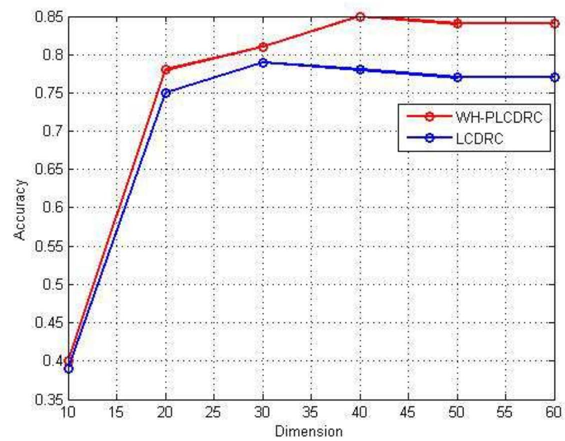

Figure 10: Four Train YALE Database Accuracy

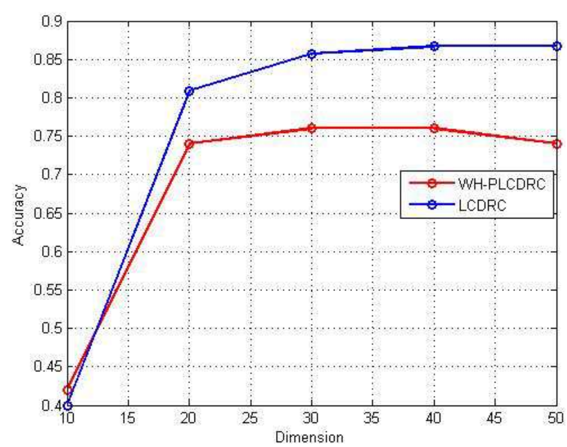

Figure 11: Five Train YALE Database Accuracy

Journal of Theoretical and Applied Information Technology
31st July 2016. Vol.89. No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

4.3 Experiments on Extended YALE B

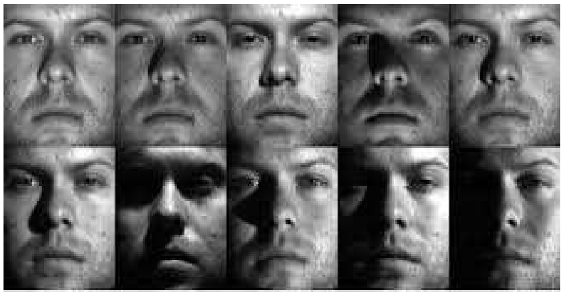

Figure 12: Extended YALE-B Face Data

The Extended YALE_B face database contains 16128 faces of 11 individuals. For each individual, there are 10 images under different facial expressions or configurations. All of the face images are cropped into a size of 32x32. The performance comparison of different methods is shown in Fig. 7. WH-PLCDRC has much higher classification accuracy than the other method. For example, the best recognition rates & the corresponding feature dimensions of LCDRC, WH-PLCDRC given four training face images for each class are 85.45%(50),88.70%(50) respectively.

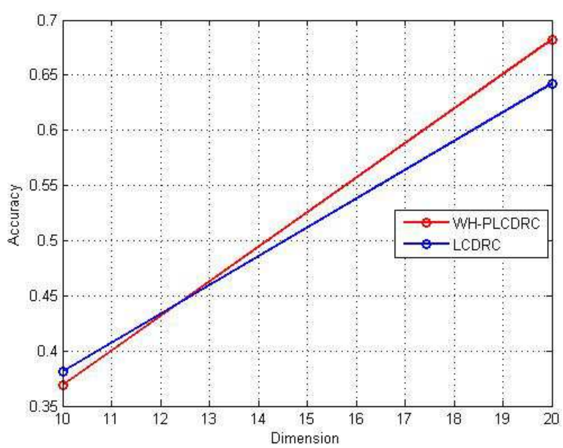

Figure 13: Two Train Extended Yale-B Database Accuracy

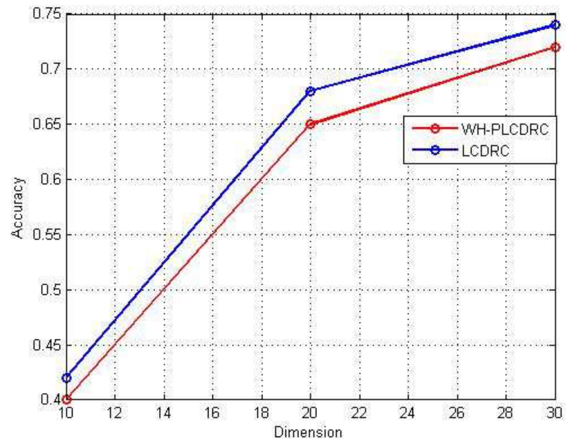

Figure 14: Three Train Extended Yale-B Database Accuracy

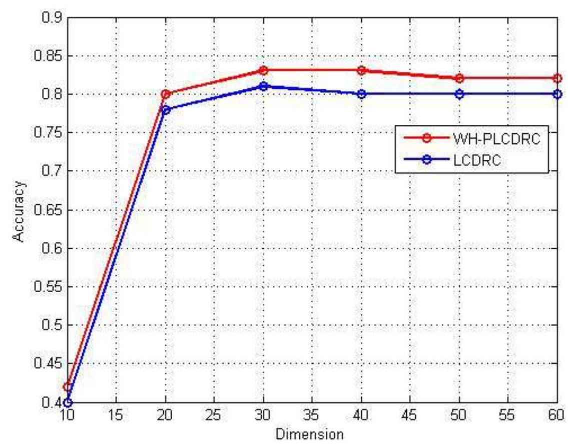

Figure 15: Four Train Extended Yale-B Database Accuracy

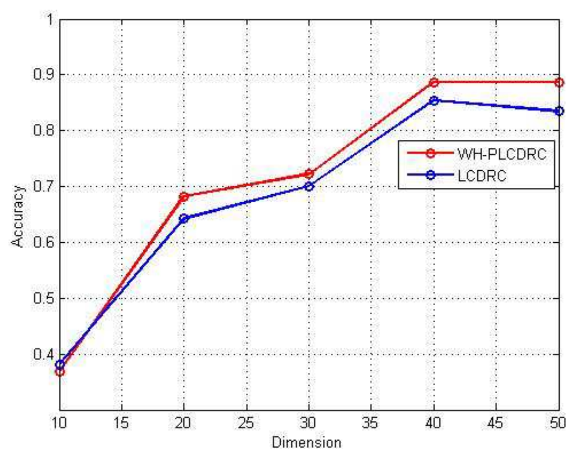

Figure 16: Five Train Extended Yale-B Database Accuracy

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 1: Performance Evolution Table

	Database	Two training	Four training
		Accuracy	Accuracy
Proposed WH-PLCDRC	ORL	0.8968	0.9650
	YALE	0.5333	0.8666
	YALE-B	0.6820	0.8870
LCDRC	ORL	0.8343	0.9421
	YALE	0.3925	0.7603
	YALE-B	0.6425	0.8545

5. CONCLUSION

In this research, Widrow-Hoff Learning Parallel Linear Collaborative Discriminant Regression presented for Face Classification. Our proposed WH-PLCDRC minimized the values of WCRE that aids to maximize the value BCRE automatically which leads to derive the optimal projection matrix. Extensive experiments conveyed that classification accuracy of our proposed WH-PLCDRC is better than LCDRC than that of LDRC.

REFRENCES:

[1] W. Zhao, R. Chellappa, A. Rosenfeld, and P. J. Phillips, “Face recognition: A literature survey”, *ACM Computer. Surv.*, Vol. 35, No. 4, 2003, pp. 399-458.

[2] S. M. Huang and J. F. Yang, "Linear Discriminant Regression Classification for Face Recognition”, *IEEE Transactions On Signal Processing Letters*, Vol. 20, No. 1, 2013.

[3] J. Wright, A.Y. Yang, A. Ganesh, S.S. Sastry, and Y. Ma, “Robust face recognition via sparse representation”, *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol. 31, 2009, pp. 210-227.

[4] D. Zhang, M. Yang, and X. Feng, “Sparse representation or collaborative representation: which helps face recognition?”, *Proceedings of IEEE International Conference on Computer Vision (ICCV)*, 2011, pp. 471-478.

[5] I. Naseem, R. Togneri, and M. Bennamoun, “Linear regression for face recognition”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 32, No. 11, pp. 2106-2112, 2010.

[6] M. Yang, D. Zhang, and J. Yang, “Robust sparse coding for face recognition”, *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2011, pp. 625-632.

[7] R. He, W.-S. Zheng, and B. G. Hu, “Maximum correntropy criterion for robust face recognition”, *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol. 33, 2011, pp. 1561-1576.

[8] X. Tang, G. Feng, and J. Cai, “Weighted group sparse representation for undersampled face recognition”, *Neurocomputing*, Vol. 145, 2014, pp. 402-415.

[9] R. Timofte, and L. Van Gool, “Adaptive and weighted collaborative representations for image classification”, *Pattern Recog. Lett.*, Vol. 43, 2014, pp. 127-135.

[10] M. Yang, D. Zhang, and S. Wang, “Relaxed collaborative representation for pattern classification”, *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2012, pp. 2224-2231.

[11] Y. Lu, X. Fang, and B. Xie, “Kernel linear regression for face recognition”, *Neural Comput. Appl.*, Vol. 24, 2014, pp. 1843-1849.

[12] X. Qu, S. Kim, R. Cui, and H. J. Kim, “Linear collaborative discriminant regression classification for face recognition”, *Journal of Visual Communication and Image Representation*, Vol. 31, 2015, pp. 312-319.

[13] M. S. Bartlett, J. R. Movellan, and T. J. Sejnowski, “Face recognition by independent component analysis”, *IEEE Transactions on Neural Networks*, Vol. 13, No. 6, 2002, pp. 1450-1464.

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

[15] S. K. Bhattacharyya, and K. Rahul, “Face Recognition by Linear Discriminant Analysis”, *International Journal of Communication Network Security*, Vol.2, No.2, 2013.

[16] A. M. Martinez, “The ar face database,” *CVC Technical Report*, Vol. 24, 1998.

[17] I. Naseem, R. Togneri, and M. Bennamoun, “Face identification using linear regression”, *Proceedings of 16th IEEE International Conference on Image Processing*, 2009, pp. 4161-4164.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[18] X. Qu, and H. J. Kim, "Enhanced Discriminant Linear Regression Classification for Face Recognition", *Proceedings of IEEE Ninth International Conference on Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP) Symposium on Cognitive Computing in Information Processing*, 2014, pp.1-5.

[19] Y. L. Liu, D. R. Zhu, D. X. Zhang, and F. Liu, "A linear regression based face recognition method by extending probe images", *Optik - International Journal for Light and Electron Optics*, Vol. 126, No. 22, 2015, pp.3335-3339.