Journal of Theoretical and Applied Information Technology
31st August 2017. Vol.95. No.16

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

KERNEL LINEAR COLLABORATIVE DISCRIMINANT REGRESSION CLASSIFICATION FOR FACE RECOGNITION USING LOCAL BINARY PATTERN

[1]P. KIRAN KUMAR REDDY
[1]Professor, MRIET-JNTUH, Department of CSE, Hyderabad, India
E-mail: [1]kiran.penubaka@gmail.com

ABSTRACT

Binary feature descriptors have been widely used in computer vision field due to their excellent discriminative power and strong robustness, and local binary patterns (LBP) and its variations have proven that they are effective face descriptors. However, the forms of such binary feature descriptors are predefined in the hand-crafted way, which requires strong domain knowledge to design them. In this paper, we propose a simple and efficient Kernel Linear Collaborative Discriminant Regression Classification (KLCDRC) feature learning method to learn a discriminative binary face descriptor in the data-driven way. Firstly, similar to traditional LBP method, we extract block based feature vectors by computing and concatenating the difference between center patch and its neighboring patches. Then learn a feature mapping to project these pixel difference vectors into low-dimensional binary vectors. Lastly, we cluster and pool these projected binary codes into a histogram-based feature that describes the co-occurrence of binary codes. And we consider the histogram-based feature as our final feature representation for each face image. We investigate the performance of our KLCDRC-LBP, KLCDRC and LCDRC on ORL and YALE databases. Extensive experimental results demonstrate that our KLCDRC descriptor outperforms other state-of-the-art face descriptors.

Keywords: *Binary feature descriptors, Histogram, Low dimensional binary vectors, Local Binary Patterns, Kernel LCDRC.*

1. INTRODUCTION

During the past decades, face recognition has been successfully applied in many fields, such as access control, ID authentication, and watch-list surveillance and so on. It has still attracted much attention due to its theoretical and practical challenges. Generally, there are two critical problems on conventional face recognition system: feature representation and classifier training. Recently, most of the existing works are focusing on these two aspects to improve the performance of face recognition methods when they faced with a variety of intra-class variability's. For face representation, the purpose is to extract discriminative features to make face images of different individual more separable. In other words, the further distance between feature with different individual is, the better feature representation method will be. For classifier training, the goal is to design an efficient (supervised/unsupervised) classifier to distinguish different face patterns. In this paper, we mainly focus on the issues of face feature Representation. Feature representation is an important problem, and it significantly affects the performance of face recognition system due to the large variations caused by the expression, pose, illumination, aging and so on. These intra-class variabilities' reduce the similarity of face images from the same individual, even the intra-class variability's often larger than the inter-class variability's in many benchmark datasets. Local features representation methods, as opposed to holistic features, describe the pattern of each local region of a face image and combine the described information of all regions into a final feature representation.

The representative examples of local features include Gabor wavelets and local binary pattern (LBP) [5]. Though, these features have achieved great success for some controlled scenarios through designing low-level features elaborately, they cannot achieve excellent performance when they faced with extreme intra-class variability and uncontrolled scenarios. Since most of these face representation methods are hand-crafted, they can achieve excellent performance only when strong

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

prior knowledge is provided. Therefore, it is a challenging problem in face recognition that how to extract robust and discriminative feature when the intra-class variabilities are large. Learning features from data itself instead of manually designing features is considered as a plausible way to overcome the limitation of hand-crafted features. In this paper, we have proposed a simple and efficient Kernel Linear collaborative discriminant regression classification for face recognition using Local Binary Pattern (KLCDRC-LBP) to learn a discriminative binary face descriptor in the data-driven way. Inspired by the research that binary codes are robust to local variations, we intend to learn discriminative and robust binary codes from raw pixels by Kernel methods, which connects the multi-class spectral clustering problem with the orthogonal Procrustes problem. Firstly, we have extracted Local Binary Pattern to improve the low level feature selection and improvisation. Then, we have performed the unsupervised data embedding method (such as LCDRC) to reduce the dimensionality of the LBP Representation.

Organization of the paper is as follows: Section II presents Literature Review about recent methodology. Section III shows proposed iris recognition system. Experimental results are discussed in section IV and finally section V gives conclusion.

2. LITERATURE REVIEW

A. Yao and S. Yu [12] have presented a new face representation method coined spatial feature interdependence matrix (SFIM). As per SFIM, the face image was projected onto an undirected connected graph in a manner that explicitly encodes feature interdependence-based relationships between local regions. They calculated the pair-wise interdependence strength as the weighted discrepancy between two feature sets extracted in a hybrid feature space fusing histograms of intensity, local binary pattern and oriented gradients. SFIM-based face descriptor is embedded in three different recognition frameworks, namely nearest neighbour search, subspace-based classification, and linear optimization-based classification, to attain the goal of face recognition. Extensive experimental results on four well-known face databases and comprehensive comparisons with the state-of-the-art results were provided to demonstrate the efficacy of the proposed SFIM-based descriptor.

H.S. Du et al. [13] have presented a new low-rank sparse representation-based classification (LRSRC) method for robust face recognition. Using the representation vector of a test sample, LRSRC classified the test sample into the class which generated minimal reconstruction error. Experimental results on Extended Yale B, CMU PIE, and AR databases confirmed that the proposed method was effective and robust, even when face images are corrupted by illumination variations, expression changes, disguises, or occlusions.

X. Qu et al. [14] have proposed a new face recognition method that improved Huang's linear discriminant regression classification (LDRC) algorithm. This research work found a discriminant subspace by maximizing the between-class reconstruction error and minimizing the within-class reconstruction error simultaneously, where the reconstruction error was obtained by utilizing Linear Regression Classification (LRC). They implemented a better between-class reconstruction error measurement which was obtained using the collaborative representation instead of class-specific representation and considered as the lower bound of all the class-specific between-class reconstruction errors. So, the maximization of the collaborative between-class reconstruction error maximized every class-specific between-class reconstruction and emphasized the small class-specific between-class reconstruction errors. This was beneficial for the following LRC. Many experiments were conducted and found that LRC has a much higher recognition accuracy on the subspace learned by LCDRC than that of LDRC.

H. Li and C.Y. Suen [15] have presented a novel approach for face recognition by extracting dynamic subspace of images and obtaining the discriminative parts in each individual. They used these parts to represent the characteristic of discriminative components and gave a recognition protocol to classify face images. The experiments carried on publicly available databases (i.e., AR, Extended Yale B, and ORL) validated its accuracy, robustness and speed. The proposed method required lower dimensions training samples but gained a higher recognition rate than other popular approaches.

Wang, et al. [16] have presented a minimum error entropy based atomic representation (MEEAR) framework for face recognition. The MEEAR produce discriminative representation vector by minimizing the atomic norm regularized Renyi's entropy of the reconstruction error. The experiments carried out on popular six real world face databases (i.e. AR, PIE, Yale, LFW, Extended Yale B, CMU Mobo) were employed. The proposed

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

MEEAR improved the both the recognition accuracy and reconstructed results.

Lei, et al. [17] have presented an Eigen Directional Bit-Plane (EDBP) robust scheme for face recognition. The proposed EDBP has overcome the shortcomings of the holistic features that are easily influenced by illumination change and occlusion. Proposed scheme encodes both local and global structures of the facial images by applying Principal Component Analysis (PCA) to the Directional bit-planes (DBP). The experiments carried on publicly available databases (i.e., AR, Extended Yale B) validated its robustness and computation time. The classifiers with EDBP are validated for practical conditions such as illumination changes, expression changes, occlusion and disguise. The proposed technique was efficient in computation time and effectively applied in the consumer applications.

Wang, et al. [18] have presented an Adaptive Sparse Representation based Classification (ASRC) for face recognition. The ASRC considers both sparsity and correlation in representation. In Sparsity ASRC selects the most discriminative samples for presentation. With the correlation information, ASRC rectify the occlusion, corruption or variations by training images of other subjects. In experimental section two datasets were employed real world face image databases (i.e. ORL, AR) and UCI repository databases. The experimental result showed that proposed method was more efficient in pattern recognition. Advanced method was also employed in feature selection and event detection in multimedia. The ASRC method represents accurate and efficient linear representation and good recognition performance.

3. PROBLEM STATEMENT

In previous face recognition experiment, report indicates that 70% of training, 30% of testing and 80% of training, 20% of testing was employed in face recognition research. So, the accuracy of face recognition rate was 67.77% average. But, our experimental result takes 70% of testing and 30% of training. We achieved higher recognition rate compare to existing experiments. Our proposed experimental analysis takes the effective recognition rate and the corresponding feature dimensions are utilizing in four various training classes.

4. PROPOSED SYSTEM

In proposed, the Face Recognition (FR) approach is utilized to analysis the human facial images by employing KLCDRC-LBP algorithm, it exploits the fisher criterion on the discriminant sub-space. Fisher criterion improves the KLCDRC algorithm as LCDRC algorithm in order to increase the proportion of Between Class Reconstruction Error (BCRE) over Within Class Reconstruction Error (WCRE) for calculating the projection matrix (U) of LDRC.

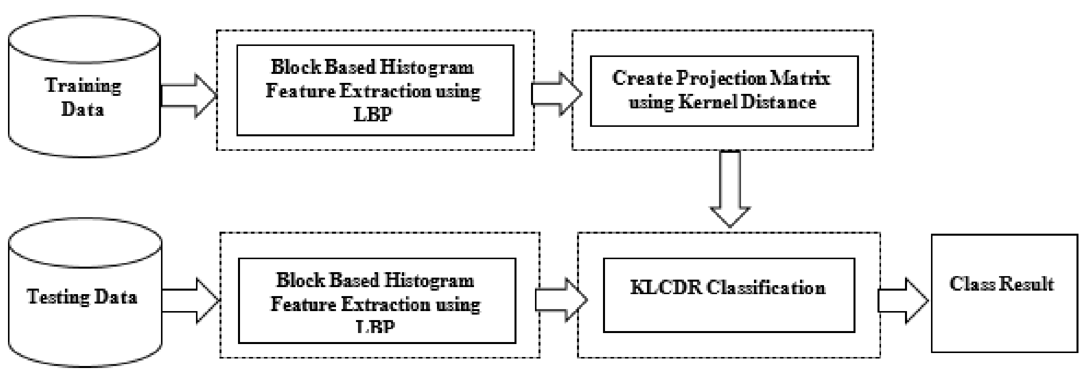

Figure 1: Block Diagram

4.1 Block Based Local Binary Pattern for Histogram Calculation

The procedure consists of using the texture descriptor to build several local descriptions of the face and combining them into a global description. Instead of striving for a holistic description this approach was motivated by two reasons: the local feature based or hybrid approaches to face recognition have been gaining interest lately[9], [11], which is understandable given the limitations of the holistic representations. These local feature based and hybrid methods seem to be more robust against variations in pose or illumination than holistic methods.

Journal of Theoretical and Applied Information Technology
31st August 2017. Vol.95. No.16

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

Another reason for selecting the local feature based approach is that trying to build a holistic description of a face using texture methods is not reasonable since texture descriptors tend to average over the image area. The facial image is divided into local regions and texture descriptors are extracted from each region independently. The descriptors are then concatenated to form a global description of the face. See Figure 2 for an example of a facial image divided into rectangular regions. The basic histogram can be extended into a spatially enhanced histogram which encodes both the appearance and the spatial relations of facial regions. These histograms are further used to classify the facial portion using KLCDRC.

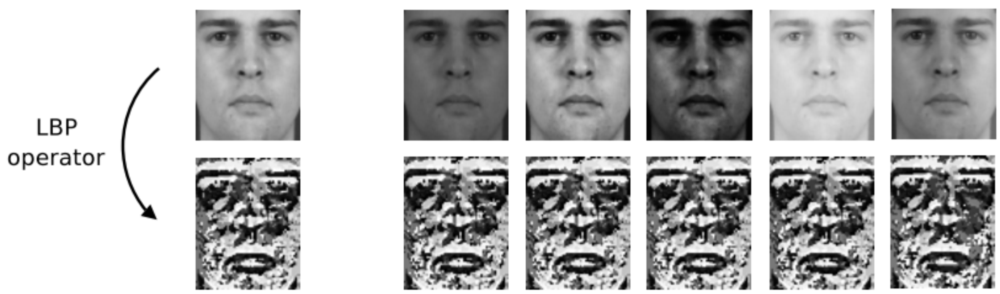

Figure 2: LBP - Histogram Calculation

4.2 Linear Discriminant Regression Classification (LDRC)

Training facial images of the $i-th$ class are stated as $C_i \in \Re^{S \times n_i}$, each column C_i is S dimensional to the facial images of class i. In which, the training images n_i are represented in vector as $i = 0,1,2,....d$, where d is declared as the total number of classes. Considering, the probe face images P, which is denoted by utilizing C_i,

$$P = C_i \beta_i, i = 0,1,2...d \qquad (1)$$

Where, $\beta_i \in \Re^{n_i \times 1}$ represented as regression parameter, β_i is determined by applying the least square estimation. In mathematically, it is specified as,

$$\hat{\beta}_i = \left(C_i^T C_i\right)^{-1} C_i^T P, i = 0,1,2,...d \qquad (2)$$

Hence, the projected vector of parameters $\hat{\beta}_i$ with the predictor C_i is employed to determine the response vector of each class i, by equating the equations (2) and (1),

$$\hat{P}_i = C_i \hat{\beta}_i = C_i \left(C_i^T C_i\right)^{-1} C_i^T P = H_i P, i = 0,1,2,...d \qquad (3)$$

Where, H_i is stated as hat matrix, that plots P into $\hat{P}_i$. Finally, the RE of each class is determined and then LRC allocates the class P with lowest RE.

$$e_i = \left\| P - \hat{P}_i \right\|_2^2, i = 0,1,2,...d \qquad (4)$$

Feature extraction technique LDRC implements discriminant analysis in the LRC to provide effective discrimination, by employing labelled training data. Assuming, all the facial images from the matrix are denoted as $C = \left[C_1,.....C_i,...,C_n\right] \in \Re^{S \times n}$, where n is characterized as the number of images and S is denoted as the dimension of images. Hence, the class label of C_i is declared as $l(C_i) \in \{0,1,2,...d\}$. Considering, the sub-space projection matrix $U \in \Re^{S \times n}$ and each face images can be projected into the sub-space as,

$$U^* = \max_U \frac{tr(U^T BCREU)}{tr(U^T (WCRE + \in l)U)} \qquad (5)$$

Where, $\in$ is symbolized as a positive number, l is an identity matrix and $tr(.)$ represents trace operator, In case $n < S$, the label of P_i is similar to label of C_i and it is characterized as $l(P_i) \in l(C_i)$. The sub-space of projection matrix U is obtained by increasing BCRE and decreasing WCRE simultaneously, WCRE and BCRE are mathematically expressed as follows,

$$WCRE = \frac{1}{n}\sum_{i=1}^{n}(C_i - C_i^{\text{intra}})(C_i - C_i^{\text{intra}})^T \quad (6)$$

$$BCRE = \frac{1}{n(d-1)}\sum_{i=1}^{n}\sum_{j=1, j\neq l(C_i)}^{d}(C_i - C_{ij}^{\text{inter}})(C_i - C_{ij}^{\text{inter}})^T \quad (7)$$

Where, C_{ij}^{inter} is denoted as the RE of C_i and $l(C_i) \neq j$, C_i^{intra} is characterized as the RE of C_i (C_i is emitted from the training matrix, while determining reconstruction).

4.3 Linear Collaborative Discriminant Regression Classification (LCDRC)

The following section describes, how the large class-specific BCRE domination issue can be diminished by the collaborative representation. In LCDRC, the WC features are compared with the total number of class's d features. While comparing, the ratio of distance between the classes are extremely maximized and also significantly reduce the distance of with in class features. In WCRE, individual features of the class are compensate with the d number of class features. Finally, the association between the WCRE and CBCRE can be denoted as,

$$WCRE = \sum_{i=1}^{d}\sum_{j=1}^{n}\left\|U^T C_{ij} - U^T C_{ij}^{\text{intra}}\beta_{ij}^{\text{intra}}\right\|_2^2 \quad (8)$$

$$CBCRE = \sum_{i=1}^{d}\sum_{j=1}^{n}\left\|U^T C_{ij} - U^T C_{ij}^{\text{inter}}\beta_{ij}^{\text{inter}}\right\|_2^2 \quad (9)$$

The respective equations (8) and (9) can be further re-written as follows,

$$WCRE = \sum_{i=1}^{d}\sum_{j=1}^{n}\left(C_{ij} - C_{ij}^{\text{intra}}\beta_{ij}^{\text{intra}}\right)^T UU^T\left(C_{ij} - C_{ij}^{\text{intra}}\beta_{ij}^{\text{intra}}\right)^T \quad (10)$$

$$CBCRE = \sum_{i=1}^{d}\sum_{j=1}^{n}\left(C_{ij} - C_{ij}^{\text{inter}}\beta_{ij}^{\text{inter}}\right)^T UU^T\left(C_{ij} - C_{ij}^{\text{inter}}\beta_{ij}^{\text{inter}}\right)^T \quad (11)$$

In the both CBCRE & WCRE have the factor of $1/n$, therefore, it is safe to eliminate $1/n$ from CBCRE & WCRE simultaneously without disturbing the value of CBCRE over WCRE. Under some algebraic deduction, CBCRE & WCRE can be denoted as follows,

$$WCRE = \sum_{i=1}^{d}\sum_{j=1}^{n} tr\left(U^T\left(C_{ij} - C_{ij}^{\text{intra}}\beta_{ij}^{\text{intra}}\right)\left(C_{ij} - C_{ij}^{\text{intra}}\beta_{ij}^{\text{intra}}\right)^T U\right) \quad (12)$$

$$CBCRE = \sum_{i=1}^{d}\sum_{j=1}^{n} tr\left(U^T\left(C_{ij} - C_{ij}^{\text{inter}}\beta_{ij}^{\text{inter}}\right)\left(C_{ij} - C_{ij}^{\text{inter}}\beta_{ij}^{\text{inter}}\right)^T U\right) \quad (13)$$

Where $tr(\cdot)$ is symbolized as the trace operator, eventually the following WCRE and BCRE can be denoted as follows,

$$WCRE = \frac{1}{n}\sum_{i=1}^{d}\sum_{j=1}^{n}\left(C_{ij} - C_{ij}^{\text{intra}}\beta_{ij}^{\text{intra}}\right)\left(C_{ij} - C_{ij}^{\text{intra}}\beta_{ij}^{\text{intra}}\right)^T \quad (14)$$

$$CBCRE = \frac{1}{n}\sum_{i=1}^{d}\sum_{j=1}^{n}\left(C_{ij} - C_{ij}^{\text{inter}}\beta_{ij}^{\text{inter}}\right)\left(C_{ij} - C_{ij}^{\text{inter}}\beta_{ij}^{\text{inter}}\right)^T \quad (15)$$

4.4 Kernel Linear cumulative discriminant regression

From the equation of (1) and (2) projection of similarity matrix its depends on regression line. Instead of using regression line our proposed system replace line into plane. The kernel LDRC equation is as follows,

$$P_i = K_i^r \beta_i + e \quad (16)$$

Then, the kernel linear regression aims to minimize the residual errors as

$$\hat{\beta}_i = \arg\min_{\beta_i}\left\|K_i^r \beta_i - P_i\right\|_2^2 \quad (17)$$

We first perform singular value decomposition (SVD) on the kernel matrix K_i as

$$K_i = CS\,V^T \quad (18)$$

We propose a constrained rank-r approximation of K_i defined as

$$K_i^r = CS^r V^T \quad (19)$$

Here r represents number of elements SVD .The above solution can be also solved by the least square estimation since it has the same form as stated in (2). After the low-rank approximation, we can use the pseudo-inverse of K_r^i to obtain the least-square solution as

$$\beta_i = \left(K_i^r\right)^- P \quad (20)$$

Journal of Theoretical and Applied Information Technology
31st August 2017. Vol.95. No.16

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

Where the pseudo-inverse of K_i^r is expressed by

$$\left(K_i^r\right)^- = C\left(S^r\right)^- V^T \quad and \quad \left(S^r\right)^- = diag\left\{\lambda_1^{-1}, \lambda_1^{-1}, ..., \lambda_1^{-1}, ..., 0, 0, ...0\right\} \tag{21}$$

Since $\left(K_i^r\right)\left(K_i^r\right)^- \neq I$, it will be feasible to compute the minimum reconstruction error between the original vector and projected vector for determining the classification results. In the classification phase, the response vector i for the i^{th} class can be predicted by

$$\tilde{P}_i = K_i^r \tilde{\beta}_i \tag{21}$$

Apply (20) in (21)

$$\tilde{P}_i = X_i \ P_i \, here \ X_i = K_i^r\left(K_i^r\right)^- \tag{22}$$

This experimentation gives a kernel plane changes in terms of feature subspace. Excremental setup and its performance is discussed in next section.

5. RESULTS AND DISCUSSION

In this section, the experimental outcome is described in detailed, which is implemented in PC with 1.8GHz Pentium IV processor using MATLAB (version 6.5). To evaluate the effectiveness of proposed algorithm, the performance of LCDRC is compared with RLCDRC on the reputed face database sets like (ORL and YALE B). In our experiment, all the facial images are cropped at the size of 32×32 .

5.1 Result for ORL Database

The ORL facial database set holds 400 face images with 40 individuals, each individual contains 10 face images respectively. Here, the following face images are taken under numerous facial expressions and altered lightening conditions, the sample face images of ORL database is given below in figure 1.

Figure 3: ORL Dataset Sample Images

The performance of LCDRC and the proposed RLCDRC in ORL database is determined and compared by referring the following figures 4, 5, 6, and 7. The effective recognition rate and the corresponding feature dimensions are given in four various training classes. All the training classes confirms that the proposed scheme is very effective in nature.

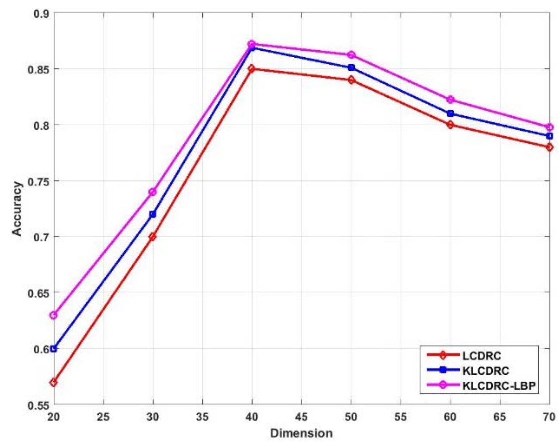

Figure 4: Accuracy calculation for 2 Train

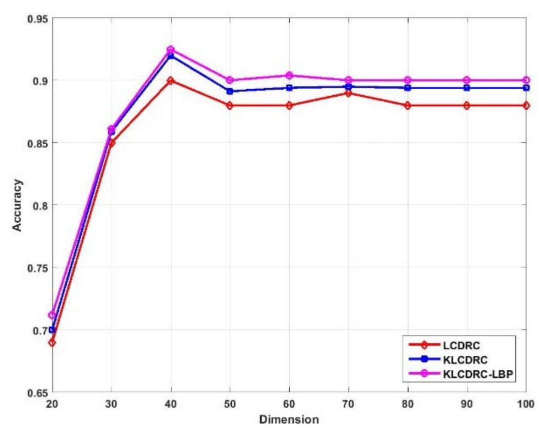

Figure 5: Accuracy Calculation for 3 Train

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

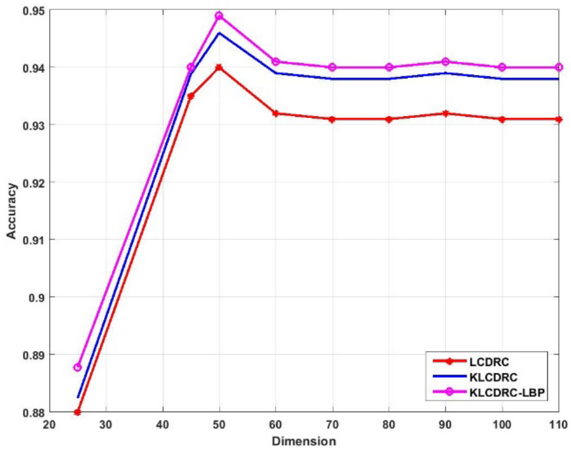

Figure 6: Accuracy Calculation for 4 Train

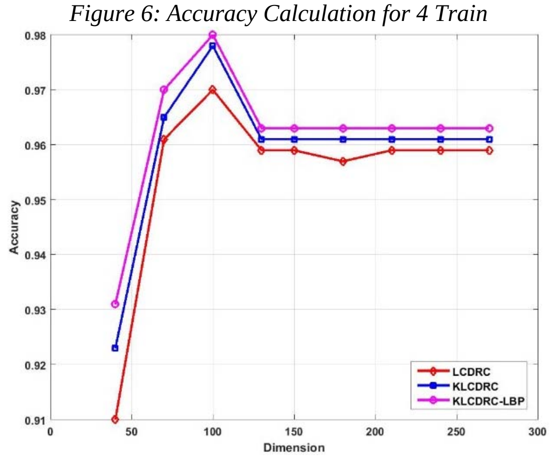

Figure 7: Accuracy Calculation for 8 Train

5.2 Result for YALE B Database

Normally, YALE B face database contains 15 individuals with 165 face images, each individuals holds 11 facial images under altered configurations and with different facial expressions, the sample face images of YALE B database is mentioned in figure 6.


Figure 8: YALE B Dataset Sample Images

The performance of LCDRC and the proposed RLCDRC in YALE B database is determined and compared by referring the following figures 9, 10, 11, and 12. For example, the significant recognition rate and the corresponding feature dimensions are mentioned in four various training classes. By analyzing all the training classes, the proposed approach shows a significant outcome in FR.

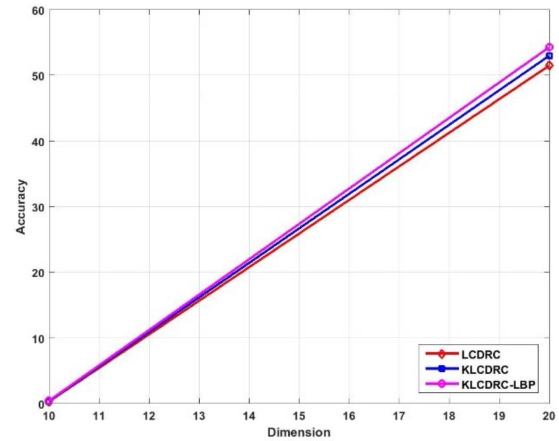

Figure 9: Accuracy Calculation for 2 Train

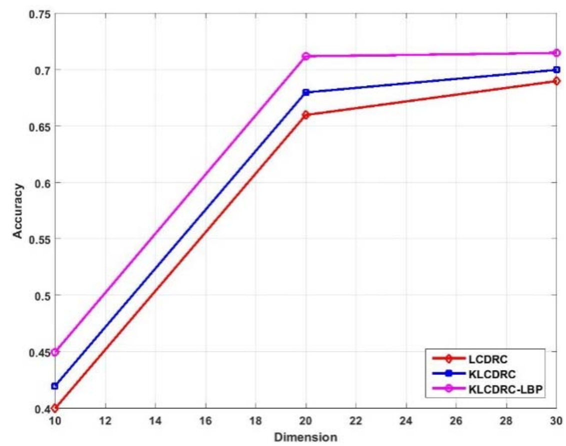

Figure 10: Accuracy Calculation for 3 Train

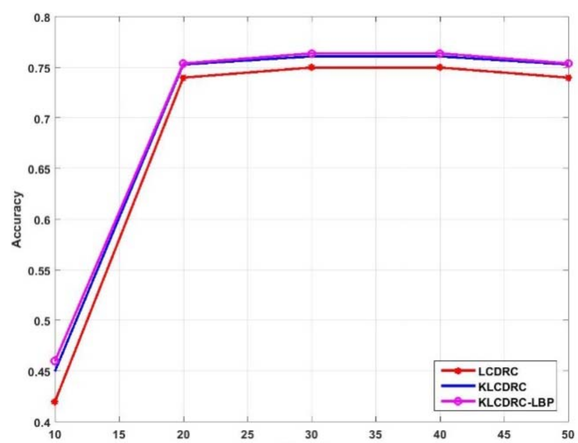

Figure 11: Accuracy Calculation for 4 Train

Journal of Theoretical and Applied Information Technology
31st August 2017. Vol.95. No.16

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

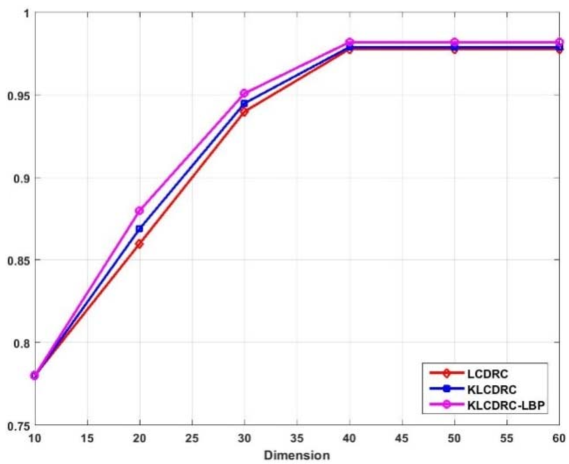

Figure 12: Accuracy Calculation for 8 Train

The following Table.1 indicates the performance analysis of KLCDRC-LBP over KLCDRC and LCDRC for three different database sets.

We randomly select five images from each subject as the training images and one face image from the remaining face images as the probe face image. The class-specific between-class reconstruction errors of the probe face image are calculated. It is obvious that some class-specific between-class reconstruction errors are very similar to the within-class reconstruction error, where the LCDRC easily makes wrong classification. Meanwhile, some class-specific between-class reconstruction errors are much larger than the within-class reconstruction error. The calculation of BCRE adds all the class-specific between-class reconstruction errors without considering their magnitude relationships. As a result, the obtained value of BCRE is mainly determined by those large class-specific between-class reconstruction errors. The maximization of BCRE tends to increase the values of those large class-specific between-class reconstruction errors. Therefore, the learned subspace is unable to well separate the within-class reconstruction error with those small class-specific between-class reconstruction errors.

Table 1: Performance Evaluation Table For ORL

Facial Database	Methodology	Two training Accuracy	Three training Accuracy	Four training Accuracy	Eight training Accuracy
ORL	LCDRC	85	90.1	94	97
	KLCDRC	86.88	92.0	94.6	97.8
	KLCDRC-LBP	87.22	92.25	94.9	98
YALE B	LCDRC	51.5	69	75	97.78
	KLCDRC	52	70	76.11	97.9
	KLCDRC-LBP	52.3	71.5	76.4	98.2

The proposed CBCRE is similar with weighted LDA in where small weights are assigned to those class pairs with large distances and large weights to those class pairs with small distances. The purpose of those weights is to restrain the dominant role of those class pairs with large distances. The proposed CBCRE actually imposes weighs in an implicit way that the small class-specific between-class reconstruction errors are emphasized and large class-specific between-class reconstruction errors are restrained. The small class-specific between-class reconstruction errors are the main reasons that LRC makes wrong clarifications, therefore, the emphasize of them improves the classification accuracy.

The major limitation of our work is in less number of iteration we can't get the accurate output.

So, more number of iterations are required in this experimental analysis and it will increase the time complexity.

6. CONCLUSION

This paper proposes a local binary pattern based kernel linear collaborative discriminant regression classification for face representation and recognition in the data-driven way. Different from the previous face classification, our KLCDRC-LBP method learn feature representation from raw data and our proposed methods are binary feature descriptor. Therefore, it can extract more discriminative and robust feature than previous methods from raw pixels. According to different data embedding way, we propose unsupervised KLCDRC-LBP algorithm for analyzing the human facial images. It exploits the fisher criterion on the discriminant sub-space. Extensive experiments are conducted on constrained face datasets, such as ORL, YALE and YALE Extended B. The experiment results show that the proposed KLCDRC-LBP can extract discriminative feature and have good generalization under different scenarios. There are two possible research directions for our future work: 1. Our KLCDRC methods show excellent generalized power under unconstrained scenarios. Therefore, we can train them in a large scale dataset that collects sufficiently inter-class and intra-class variations, and apply them to solve some problems of unconstrained face recognition scenarios in real life. 2. Since our KLCDRC is an unsupervised method there are few general learning-based methods, like deep learning and extreme learning will used to learn discriminative feature representations of various visual tasks as long as extensive training samples.

REFRENCES:

[1] M. Turk, A. Pentland, "Face recognition using eigenfaces", *Proceedings of 1991 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1991, pp. 586–591

[2] 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, pp. 711–720, 1997.

[3] 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, pp. 1450-464, 2002.

[4] C. Liu, and H. Wechsler, "Gabor feature based classification using the enhanced Fisher linear discriminant model for face recognition", *IEEE Transactions on Image processing,* Vol. 11, No. 4, pp. 467–476, 2002.

[5] T. Ahonen, A. Hadid, and M. Pietikainen, "Face description with local binary patterns: application to face recognition", *IEEE transactions on pattern analysis and machine intelligence*, Vol. 28, No. 12, pp. 2037-2041, 2006.

[6] J. Lu, V. Liong, X. Zhou, and J. Zhou, "Learning compact binary face descriptor for face recognition", *IEEE transactions on pattern analysis and machine intelligence*, Vol. 37, No. 10, pp.2041-2056, 2015.

[7] S.M. Huang, and J.F. Yang, "Linear discriminant regression classification for face recognition", *IEEE Signal Processing Letters*, Vol. 20, No. 1, pp. 91-94, 2013.

[8] S.M. Huang, and J.F. Yang, "Unitary regression classification with total minimum projection error for face recognition", *IEEE Signal Processing Letters*, Vol. 20, No. 5, pp. 443-446, 2013.

[9] T. Ahonen, A. Hadid, and M. Pietikäinen, "Face recognition with local binary patterns", *Proceedings of Computer vision-eccv (Springer, Berlin* Heidelberg, 2004, pp. 469-481.

[10] J. Qian, J. Yang, and X. Yong, "General regression and representation model for classification", *PLoS One*, Vol. 9, No. 12 pp. e115214, 2014.

[11] B. Heisele, P. Ho, J. Wu, and T. Poggio, "Face recognition: component-based versus global approaches," *Compter Vision and Image Understanding*, Vol. 91, No. 1, pp. 6-21, 2003.

[12] A. Yao, and S. Yu, "Robust face representation using hybrid spatial feature interdependence matrix", *IEEE Transactions on Image Processing*, Vol. 22, No. 8, pp. 3247-3259, 2013.

[13] H.S. Du, Q.P. Hu, D.F. Qiao, and I. Pitas, "Robust face recognition via low-rank sparse representation-based classification", *International Journal of Automation and Computing*, Vol. 12, No. 6, pp. 579-587, 2015.

[14] 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, pp.312-319, 2015.

[15] H. Li, and C.Y. Suen, "Robust face recognition based on dynamic rank representation", *Pattern Recognition*, Vol. 60, pp. 13-24, 2016.

[16] Y. Wang, Y.Y. Tang, and L. Li, "Robust face recognition via minimum error entropy-based atomic representation", *IEEE Transactions on Image Processing*, Vol. 24, No. 12, pp. 5868-5878, 2015.

[17] L. Lei, S.W. Kim, W.J. Park, D.M. Kim, and S.J. Ko, "Eigen directional bit-planes for robust face recognition", *IEEE transactions on Consumer Electronics*, Vol. 60, No. 4, pp. 702-709, 2014.

[18] J. Wang, C. Lu, M. Wang, P. Li, S. Yan, and X. Hu, "Robust face recognition via adaptive sparse representation", *IEEE transactions on cybernetics*, Vol. 44, No. 12, pp. 2368-2378, 2014.