International Journal of Trend in Research and Development, Volume 6(6), ISSN: 2394-9333
www.ijtrd.com

Reduction of Intrinsic Dimension of Feature Space Inherited By Subspace Methods for Expression Recognition

[1]G.P.Hegde, [2]Deekshith K and [3]Nagaratna Hegde,
[1,2]Department of Information Science and Engineering, SDM Institute of Technology, Ujire, Karnataka, India
[3]Department of Computer Science and Engineering, Vasavi College of Engineering, Hyderabad, India

Abstract: This paper mainly focuses on study of proper discrimination of facial expression classes by optimizing the singularity metric problems. Several earlier LDA based subspace approaches have been surveyed are all related to solution criteria for singularity matrix problem. Intrinsic dimension of linear discriminant feature space has been reduced into lower space by the subspace approaches. Some of the methods have seen which preserves the null space and also eliminates null space of scatter matrix. This study also emphasizes on comparison of LDA based subspace methods.

Keywords: *Linear Discriminant Analysis (LDA), Small Sample Size, Face Recognition.*

I. INTRODUCTION

A variety of subspace approaches are defined in [1] for efficient facial expression recognition (FER) by solving the problem of dimensionality reduction of higher feature dataset. These approaches find maximum applications in the research of artificial intelligence and pattern recognition domains. Facial expression recognition has many potential applications to human-mobile interaction, data-driven animation, image retrieval, human emotional analysis, medical patient treatment etc. Gunes et al. [3] described about effect of automatic facial expression recognition. Shan et al. [4] worked on discriminative feature based facial expression scheme using Local Binary Pattern and concluded that facial features are more dominant key elements for expression recognition. Face detection is a primary step for further task while recognizing expressions as described in [5].

II. SUBSPACE APPROACHES

Subspace method finds many application in dimension reduction of high dimensional datasets. Belhumeur et al. [7] 1997 presented a Fisherface subspace method and addressed two stage PCA +LDA approach, using PCA high dimensional face data is projected into a low dimensional feature space and then LDA is performed in the low dimensional PCA subspace . Only high eigen values are considered in eigenface and lower eigen values are removed in the PCA subspace. Due to removal of small eigen values introduces a loss of discrimination information. In this work preservation of local and global discrimination features in PCA subspace was carried out by CEGKLSWFDA approach. Due to removal of lower eigen values causes the loss of local and global discrimination features of face images. In our earlier work [1][2]] PCA subspace depends on total number of training samples and its size. When the PCA subspace dimension is relatively high. Singularity and instability properties of within class scatter matrix of Linear Discrimiant Analysis (LDA) yields larger number of pixel values or more number of face features after projection of high dimensional dataset into subspace for efficient facial expression recognition. But total number of training samples is limited or lesser than the dimension of the feature space. In this paper LDA based method is compared with other methods with respect to singularity problems in within scatter matrix and curse of dimensionality of feature dataset. Figure 1 shows under sampling problem of LDA.

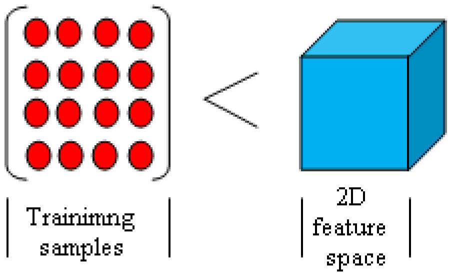

Figure 1 Under sampling problem of LDA

III. BRIEF OVERVIEW OF LDA

Li et al. [14] illustrated about Linear Discriminant Analysis (LDA) method encodes discriminatory information in a linear separable space of which bases are not necessarily orthogonal. Researchers have demonstrated that the LDA based algorithms outperform the PCA algorithm for many different tasks . Traditional linear discriminant analysis (LDA) based methods suffer from the disadvantage that their optimality criteria are not directly related to the classification ability of the obtained feature representation. Moreover, their classification accuracy is affected by the "small sample size" (SSS) problem which is often encountered in face recognition tasks. Linear Discriminant Analysis is a well-known method which projects the data onto a lower dimensional vector space such that the ratio of the between-class distances to the within-class distance is maximized, thus achieving maximum discrimination. The optimal projection can be readily computed by applying the eigen objective functions, such as face recognition, all scatter matrices in question can be singular since the data is from a very high-dimensional space, and in general, the dimension exceeds the number of data points. This is known as the under sampled or singularity problem. Let us consider 2D image of f(x,y) having x rows and y columns forming a matrix x and y. Total dimension of the image is given by n=xy. Then convert the image vector of size nx1 by cascading each column placed one after the other. Then we get training set as Ii=[I1, I2, ……. Im] of image vectors of size nxm. Where m is the number of training images. Then compute the mean value of the training set image vectors at each pixel level called mean face. Size of the mean face is given as

$$\overline{I_i} = \frac{1}{m}\sum_{i=1}^{m} I_i$$

Then subtract the mean value from image I_i and compute mean subtracted image such as

$$\Psi_i = \left(I_i - \overline{I_i}\right)$$

A=[Ψ_1 Ψ_2 Ψ_3 - - - - Ψ_m] is the matrix of all the mean subtracted training image vectors and its size is *nxm*. Then covariance matrix is computed which represents the scatter degree of all feature vectors related to the mean vector.

$$C = A.A^T = \frac{1}{m}\sum_{i=1}^{m}\Psi_i\Psi_i^T$$

From the *nxn* covariance matrix eigenvectors are calculated, the size of the eigenvectors is *mxm*. Covariance matrix can be considered as

$$L = A.A^T = \frac{1}{m}\sum_{i=1}^{m}\Psi_i\Psi_i^T$$

Using *L* eigenvectors and respective eigenvalues can be calculates as $LV=\lambda V$, where V is the set of eigenvectors associated with its eigenvalue λ. By projecting the matrix A the eigenface can be computed and normalized, hence new eigenvectors w_k are formed. Each mean subtracted image project into eigenspace using,

$$w_k = V_k^T\Psi = V_k^T\left(I_i - \overline{I_i}\right)$$

Where $k = 1,2,3, \ldots \ldots . . m'$ Finally, weight matrix can be computed as

$$\Omega = \left[w_i, w_2, w_3 \ldots\ldots w_m'\right]^T$$

All the training images are projected into the eigenface space, The transformation from n dimensional space into m' dimensional space hence dimensional reduction is achieved. Actual LDA can be computed by considering the eigenface projection of PCA called

$$J(W) = \left|\frac{W^T.S_b.W}{W^T.S_w.W}\right|$$

It uses calculate S_b and S_w original image pixels. Foe subspace LDA method we have to consider eigenface projection of PCA weight matrix. So between class scatter matrix S_b is computed by considering Ω and mean value. That is a_i , a_o

Eigenface space class mean is

$$a_i = \frac{1}{b_i}\sum_{k=1}^{bi_i}\Omega_k$$

Where b_i is training images in class c m_i is the arithmetical average of the eigenface projected training image vector corresponding to the same individual, i=1,2,------c and its size is (m'x1) . similarly a_o

$$a_o = \frac{1}{m_i}\sum_{k=1}^{m_i}\Omega_k$$

IV. OPTIMIZATIOM OF SINGULARITY PROBLEMS

Hythem et al. proposed [9] Relevance-Weighted Two Dimensional Linear Discriminant Analysis (RW2DLDA). Its over comes the singularity problem implicitly, while achieving efficiency. Moreover, a weight discriminant hyper plane is used in the between class scatter matrix, and RW method is used in the within class scatter matrix to weight the information to resolve confusable data in these classes. This method attempts to model the difference between the classes of pixel data. It is a powerful face recognition method that overcomes the limitation of Principle Component Analysis method by applying the linear discriminant criterion. This criterion tries to maximize the ratio of the determinant of the between class scatter matrix of the projected samples to the determinant of the within class scatter matrix of the projected samples Statistical discrimination methods are more suitable for discrimination of classes and its classification. Fisher discriminant Analysis is a modified method of linear Discriminant Analysis has been used on statistical feature dimensional reduction method. Within the class scatter matrix either will be singular, if its rank is less than or it is unstable if the total number training samples are not significantly larger than dimension of feature space. It is also called under samples problem or singularity problem.

One of the solution to achieve dimensional reduction is implementiong intermediate dimensional reduction stage using PCA. Yu et al. [10] presented direct LDA which diagonalizes the between class scatter matrix to obtain low dimensional space and forming discriminant vectors which minimizes the within class scatter matrix. Chen introduced LDA based method to solve the problem of singularity matrix called CLDA (Chen LDA) when direct LDA is used still there is singularity matrix problem occurs. The basic idea is use discriminative informationn of the null space of the within class matrix to maximize the between class matrix. YLDA discriminant method was proposed by Yang and Yang [12], their method is capable of deriving discrimination information of LDA criterion in singular cases. Regularized Complete Linear Discriminant Analysis (RCLDA) proposed by Yang [13] utilized two regularized criterion to derive "regular" discriminant vectors in the range space of the within-class scatter matrix and "irregular" discriminant vectors in the null space of the within-class scatter matrix.

CONCLUSIONS

Subspace formaion and discimnation of facial expression class features is a challenging task in pattern recognition. Singularity problems of scatter matrix causes poor discrimination of classes. Different methods like CLDA, DLDA, YLDA, RCLDA are defined by several reserachers and tried to increase the discrimination ratio of within class matrix to maximize the between class matrix. In some methods null space has been removed. Facial expression classification is determined using discrimination ofr u no gular olocal and global features. This study concludeds singularity matrix may be singular or nonsingular depends on number of sample of trained data.

References

[1] G.P. Hegde, M. Seetha, Nagaratna Hegde, (2016). "Kernel Locality Preserving Symmetrical Weighted Fisher Discriminant Analysis based sun bspace approach for expression recognition", Engineering Science and Technology, an International Journal. Elsevier Publishers.

[2] G. P. Hegde , Shubhada V. P. , M. Seetha , Nagaratna P. Hegde, (2016). "Recognition of Expressions Based on Kernel Global and Local Symmetrical Weighted Fisher Discriminant Nonlinear Subspace Approach," International Journal of Applied Engineering Research ISSN 0973-4562 Volume 11, Number 5. pp 3558-3569.

[3] Hatice Gunes, Hayley Hung , "Is automatic facial expression recognition of emotions coming to a dead

end? The rise of the new kids on the block," Image and Vision Computing, 2016

[4] Caifeng Shan , Shaogang Gong, Peter W. McOwan, Facial expression recognition based on Local Binary Patterns: A comprehensive study Image and Vision Computing, Volume 27, Issue 6, 4 May 2009, Pages 803–816.

[5] Daniel McDuff, Rana El Kaliouby, Thibaud Senechal, David Demirdjian, Rosalind Picard, "Automatic measurement of ad preferences from facial responses gathered over the Internet," Image and Vision Computing Volume 32, Issue 10, Pages 630–640, October 2014.

[6] Liangyue Li, Sheng Li, Yun Fu, "Learning low-rank and discriminative dictionary for image classification," Image and Vision Computing, Volume 32, Issue 10, , Pages 814–823, October 2014.

[7] P.Belhumeur, J.Hespanha, and D.Kriegman. Using discriminant eigen features for image retrieval. PAMI, 19(7):711–720, 1997. 14.

[8] W. Zhao, R. Chellappa, and N. Nandhakumar. Empirical performance analysis of linear discriminant classifiers. In: Proc. Computer Vision and Pattern Recognition, (Santa Bar- bara, CA, 1998) 164–169.

[9] Hythem Ahmed, Jedra Mohamed, Zahid Noureddine, "Face Recognition Systems Using Relevance Weighted Two Dimensional Linear Discriminant Analysis Algorithm," Journal of Signal and Information Processing Vol.3 No.1 PP. 130-135

[10] H. Yu and J. Yang, "A direct LDA algorithm for high dimensional data - with application to face recognition",*Pattern Recognition*, vol. 34, 2067-2070, 2001.

[11] L. Chen, H. Liao, M. Ko, J. Lin, and G. Yu, "A new LDA-based face recognition system which can solve the small sample size problem", *Pattern Recognition*, 33 (10), pp. 1713-1726, 2000.

[12] J. Yang and J. Yang, "Why can LDA be performed in PCA transformed space? ", *Pattern Recognition*, vol. 36, pp. 563-566, 2003.

[13] Wuyi Yang, "Regularized Complete Linear Discriminant Analysis for Small Sample Size Problems, "Emerging Intelligent Computing Technology and Applications Volume 304 of the series Communications in Computer and Information Science pp 67-73