International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 04 Issue: 08 | Aug -2017 www.irjet.net p-ISSN: 2395-0072

Linear Discriminant Analysis for Human Face Recognition

Mr. Rahul M. Ohol[1], Mrs. Shilpa R. Ohol[2]

[1]*Lecturer, Dept. Of Medical Electronics Govt. Polytechnic Miraj, Maharashtra, India*
[2] *Asst. Professor, Dept. Of Mathematics P.V.P.I.Technology, Sangli, Maharashtra, India*

---***---

Abstract – *There are many possible techniques for classification of data. Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) are two commonly used techniques for data classification and dimensionality reduction. Linear Discriminant Analysis easily handles the case where the within-class frequencies are unequal and their performance has been examined on randomly generated test data. This method maximizes the ratio of between-class variance to the within-class variance in any particular data set thereby guaranteeing maximal separability. The use of Linear Discriminant Analysis for data classification is applied to classification problem in face recognition and speech recognition.*

***Key Words*: Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), face recognition, data classification**

1. INTRODUCTION

Biometrics is an old Greek word, "Bio", meaning life and "Metric" the measure of, so Biometrics is in essence, the measure of life. Simply defined, Biometrics is the automated use of physiological or behavioural characteristics to determine or verify identity of a person.A concise definition of biometrics is "The automatic recognition of a person using distinguishing traits." A more expansive definition of biometrics is "Any automatically measurable, robust and distinctive physical characteristic or personal trait that can be used to identify an individual or verify the claimed identity of an individual." This definition requires elaboration. Biometrics is the science of verifying the identity of an individual through physiological measurements or behavioural traits. Since biometric identifiers are associated permanently with the user and they are more reliable than token or knowledge based authentication methods. Biometrics offers several advantages over traditional security measures [1].

1.1 Why Face Recognition?

Among the biometrics, the face is the most natural physiological characteristic to recognize each other. Hence, people consider face a "good" biometric for automatic identity recognition systems.There are a number of reasons to choose face recognition. These are as follows

- It is non-intrusive and requires no physical interaction on behalf of the user.The system captures faces of people in public areas, which minimizes legal concerns for reasons explained below. Moreover, since faces can be captured from some distance away, facial recognition can be done without any physical contact.
- The acquisition process can be performed with a limited person cooperation.
- It is accurate and allows for high enrolment and verification rates.
- It does not require an expert to interpret the comparisons.
- It can use the existing hardware infrastructure i.e. existing cameras and image capture devices.
- It is the only biometric technology that allows you to perform passive identification in a one-to-many environment.

2 Linear Discriminant Analysis(LDA)

Originally developed in 1936 by R.A. Fisher, discriminant analysis is a classic method of classification that has stood the test of time. Discriminant analysis often produces models whose accuracy approaches (and occasionally exceeds) more complex modern methods. Discriminant analysis can be used only for classification (i.e., with a categorical target variable), not for regression. The target variable may have two or more categories. It is also known as Fisher Discriminant Analysis (FDA).

2.1 Defination of groups

The groups to be discriminated can be defined either naturally by the problem under investigation, or by some preceding analysis, such as a cluster analysis. The number of groups is not restricted to two, although the discrimination between two groups is the most common approach. Note that the number of groups must not exceed the number of variables describing the data set. Another prerequisite is that the groups have the same covariance structure (i.e. they must be comparable).

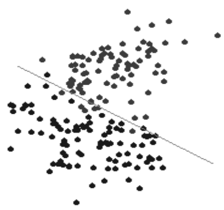

Fig -1: Cluster analysis

2.2 Estimation of the parameters of the discriminating function

There is only one direction of the discriminating line which yields the best separation results. The determination of the coefficients of the discriminating function is quite simple. In principle, the discriminating function is formed in such a way that the separation (=distance) between the groups is maximized, and the distance within the groups is minimized. A transformation function is found that maximizes the ratio of between-class variance to within-class variance as illustrated by this figure 4.4

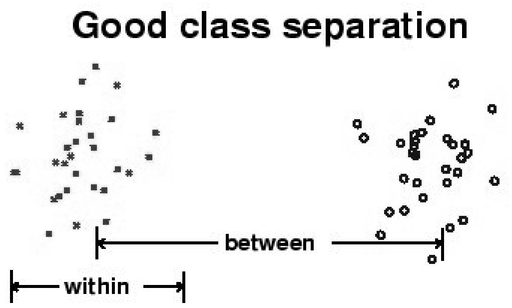

Fig -2: Good class seperation

The transformation seeks to rotate the axes so that when the categories are projected on the new axes, the differences between the groups are maximized. The following figure shows two rotates axes. Projection to the lower right axis achieves the maximum separation between the categories; projection to the lower left axis yields the worst separation.

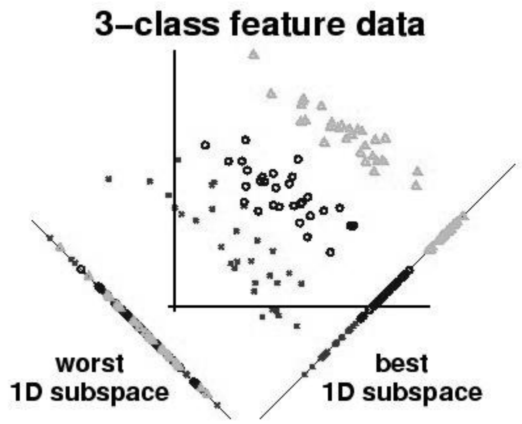

Fig -3: Class feature data

2.3 Mathematical Operations

Fisher discriminates group images of the same class and separates images of different classes. Images are projected from N-dimensional space (where N is the number of pixels in the image) to C-1 dimensional space (where C is the number of classes of images). For example, consider two sets of points in 2-dimensional space that are projected onto a single line. Depending on the direction of the line, the points can either be mixed together or separated. Fisher discriminants find the line that best separates the points. To identify a test image, the projected test image is compared to each projected training image, and the test image is identified as the closest training image.

As with eigenspace projection, training images are projected into a subspace. The test images are projected into the same subspace and identified using a similarity measure. What differs is how the subspace is calculated. Following are the steps to follow to find the Fisher discriminants for a set of images.

2.3.1 Calculate the within class scatter matrix

The within class scatter matrix measures the amount of scatter between items in the same class. For the i^{th}class, a scattermatrix (Si) is calculated as the sum of the covariance matrices of the centered images in that class.

$$Si = \sum_{x \in xi} (x - mi)(x - mi)^T$$

mi is the mean of images in the class. The within class scatter matrix (S_W) is the sum of all scatter matrices.

$$S_W = \sum_{i=1}^{C} S_i$$

C is the number of classes.

2.3.2 Calculate the between class scatter matrix

The between class scatter matrix (S_B) measures the amount of scatter between classes. It is calculated as the sum of the covariance matrices of the difference between the total mean and the mean of each class.

$$S_B = \sum_{i=1}^{c} ni(mi - m)(mi - m)^T$$

where ni is the number of images in the class, *mi* is the mean of the images in the class and *m* is the mean of all the images.

2.3.3 Solve the generalized eigenvalue problem

Solve for the generalized eigenvectors (V) and eigenvalues (λ) of the within class and between class scatter matrices.

$$S_B V = \lambda S_W V$$

2.3.4 Keep first C-l eigenvectors

Sort the eigenvectors by their associated eigenvalues from high to low and keep the first C 1 eigenvectors. These eigenvectors form the Fisher basis vectors.

2.3.5 Project images onto Fisher basis vectors

Project all the original (i.e. not centered)

images onto the Fisher basis vectors by calculating the dot product of the image with each of the Fisher basis vectors. The original images are projected onto this line because these are the points that the line has been created to discriminate, not the centered images.

Following are the steps to follow to find the Fisher discriminants of a set of images by first projecting the images into any orthonormal basis.

1. Compute means Compute the mean of the images in each class(mi)and the total mean of all images (m).

2. Center the images in each class Subtract the mean of each class from the images in that class.

3.Center the class means Subtract the total mean from the class means.

4. Create a data matrix Combine the all images, side-by-side, into one data matrix.

5.Find an orthonormal basis for this data matrix This can be accomplished by using a QR Orthogonal-triangular decomposition or by calculating the full set of eigenvectors of the covariance matrix of the training data. Let the orthonormal basis be U

6. Project all centered images into the orthonormal basis Create vectors that are the dot product of the image and the vectors in the orthonormal basis.

7. Project the centered means into the orthonormal basis

8. Calculate the within class scatter matrix The within class scatter matrix measures the amount of scatter between items within the same class. For the i^{th} class a scatter matrix (Si) is calculated as the sum of the covariance matrices of the projected centered images for that class.

9. Calculate the between class scatter matrix The between class scatter matrix (S_B)

measures the amount of scatter between classes. It is calculated as the sum of the covariance matrices of the projected centered means of the classes, weighted by the number of images in each class.

10. Solve the generalized eigenvalue problem

11. Keep the first C-l eigenvectors Sort the eigenvectors by their associated eigenvalues from high to low and keep the first C -1 eigenvectors. These are the Fisher basis vectors.

12. Project images onto eigenvectors: Project all the rotated original (i.e. not centered)

images onto the Fisher basis vectors. First project the original images into the orthonormal basis, and then project these projected images onto the Fisher basis vectors. The original rotated images are projected onto this line because these are the points that the line has been created to discriminate, not the centered images.

3. CONCLUSIONS

LDA attempt to maximize the between class scatter, while minimizing the within class scatter. In other words, moves images of the same class closer together, while moving images of different classes further apart. Lighting conditions, Image quality, Pose orientation plays an important role in face recognition system

REFERENCES

[1] John D. Woodward, Jr., Christopher Horn, Julius Gatune and Aryn Thomas, Biometrics A Look at Facial Recognition, Prepared for the Virginia State Crime Commission, Published by RAND,(2003).

[2]Gian Luca Marcialis and Fabio Roli, "Decision - level fusion of PCA and LDA-based face recognition algorithms," International Journal of Image and Graphics, Vol. 6, No.2, pp 293-311, (2006).

[3]Juwei Lu, Kostantinos N. Plataniotis, and Anastasios N. Venetsanopoulos, "Face Recognition Using LDA-Based Algorithms," IEEE Transactions on Neural Networks, Vol. 14, No.1, January 2003.

[4] A tutorial on Principal Component Analysis by Lindsay I Smith February 26,2002