Project Report

Face Recognition using Fisher Linear Discriminant Analysis (LDA)

Big Data Analytics

(CS 696)

Fall 2018

By

Reetu Hooda

A25244657

Table of Contents

1. Introduction

(a) Summary of Project 1

An image processing methodology for face detection and PCA on the detected images was implemented in project 1. The image processing involved steps to clean, color balance, and appropriately segmentation of the data. Certain features were detected, such as the eyes, lips, and teeth. Those features are isolated with segmentation, and inter-feature distances, angles, and curvatures were measured. The completed training implementation used Viola-Jones's Haar-like feature cascade detector to detect faces as well as eyes and mouths. Detected faces were cropped, resized, and mean subtracted, then PCA is performed to obtain the eigenfaces.

(b) Project 2

This project includes the Fisher linear discriminant analysis (LDA) in a reduced dimensionality. For each emotion that we wish to train a predictor for, we performed Fisher LDA, in which the goal is to optimize the objective function that minimizes within class variance and maximizes between class variance to gain clear class separation between the class of interest and the other classes.

2. Implementation and Results

The workflow of the algorithm is summed up by the flowchart shown in Figure 1.

One way to represent the input data is by finding a subspace which represents most of the data variance. This can be obtained with the use of Principal Components Analysis (PCA). When applied to face images, PCA yields a set of eigenfaces. These eigenfaces are the eigenvectors associated to the largest eigenvalues of the covariance matrix of the training data. The eigenvectors thus found correspond to the least-squares (LS) solution. This is indeed a powerful way to represent the data because it ensures the data variance is maintained while eliminating unnecessary existing correlations among the original features (dimensions) in the sample vectors.

When the goal is classification rather than representation, the LS solution may not yield the most desirable results. In such cases, one wishes to find a subspace that maps the sample vectors of the same class in a single spot of the feature representation and those of different classes as far apart from each other as possible. The techniques derived to achieve this goal are known as discriminant analysis (DA).

The most known DA is Linear Discriminant Analysis (LDA), which can be derived from an idea suggested by R.A. Fisher in 1936. When LDA is used to find the subspace representation of a set of face images, the resulting basis vectors defining that space are known as Fisherfaces.

The objective function is given by:

$$argmax_w J(w) = \frac{w^T S_B w}{w^T S_W w}$$

where S_B is the between-class scatter matrix defined as:

$$S_B = (m_2 - m_1)(m_2 - m_1)^T$$

and S_W is the within-class scatter matrix defined as:

$$S_W = \sum_j^2 \sum_{x \in C_j} (x - m_j)(x - m_j)^T$$

And m_j is the mean of class j.

When performing the LDA, we will proceed with the one versus-all (OVA) approach for each emotion, where all nontarget emotion training samples will be grouped. Then, we perform PCA once again on $({S_W}^{-1} S_B)$. The eigenvector corresponding to the largest eigenvalue is the known as the Fisherface for the emotion in-training, some of which are shown in Figure 2(a)-2(e). Figure 2(a)-(e) are top eigenvectors reshaped to 100x100 images (Fisherfaces) after Fisher LDA for Anger, Fear, Happy, Sad and Surprise.

We then project all the training data used to calculate the Fisherface for each emotion onto that particular Fisherface. Binning the projection values into histograms to examine the distribution allows us to determine thresholds for each Fisherface's projection values. The Fisherfaces do reasonably in separating the classes for each emotion, as shown in Figure 3(a)-3(e). Figure 3(a)-3(e) corresponds to Distributions of training data projected back onto calculated Fisherfaces for Anger, Fear, Happy, Sad and Surprise. Distributions of within-class shown in red and outside-class shown in blue are relatively well separated. These Fisherfaces thresholds can then be used to classify test data that we have. We will detect and crop the test images in the same manner in which we did for the training images, and then project the test image onto each Fisherface. Then a classification prediction can be made based on the projection coefficient and the threshold we have established.

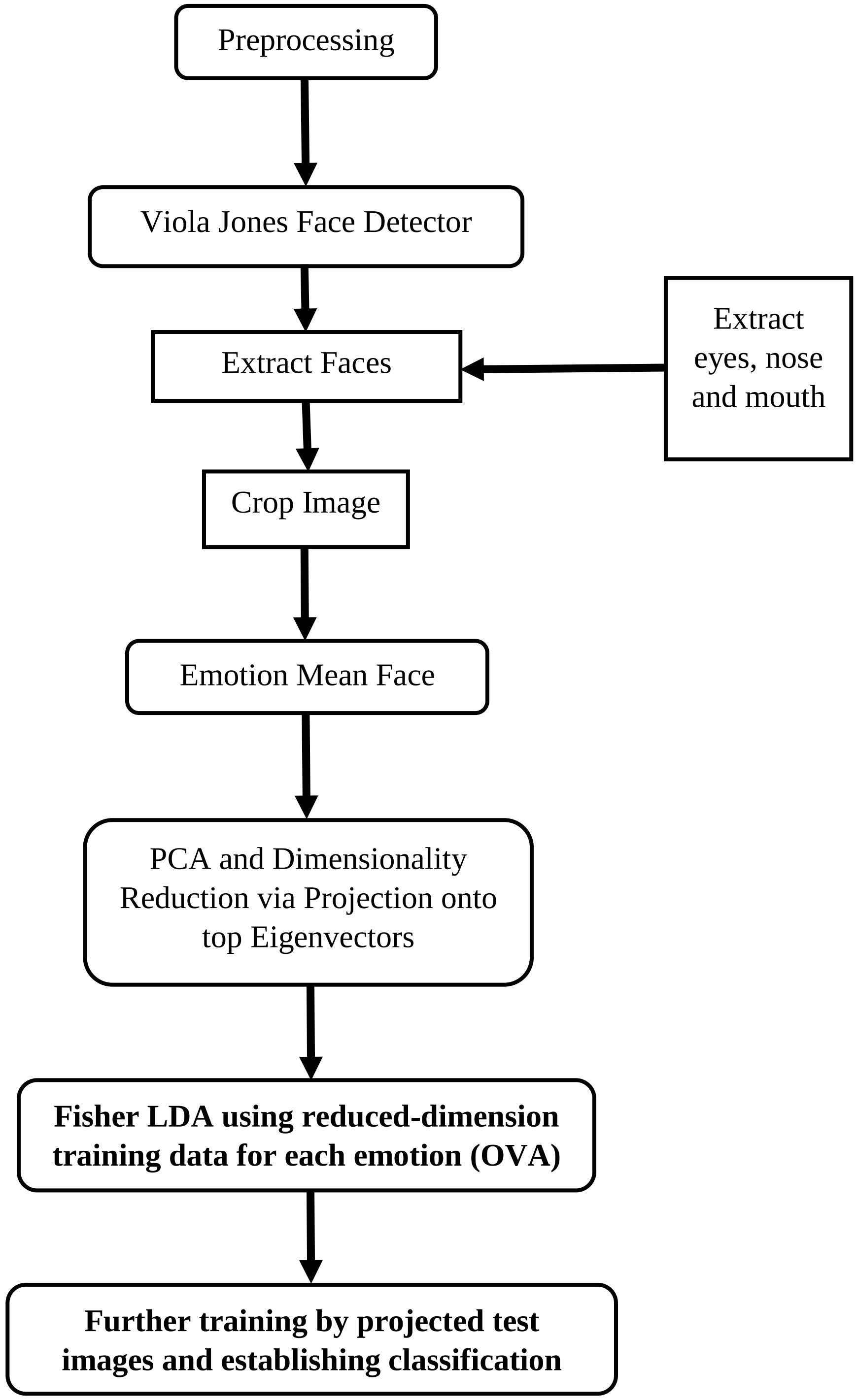

Figure 1: Flowchart of the algorithm.

.

Figure 2(a): Happy Fisherface.

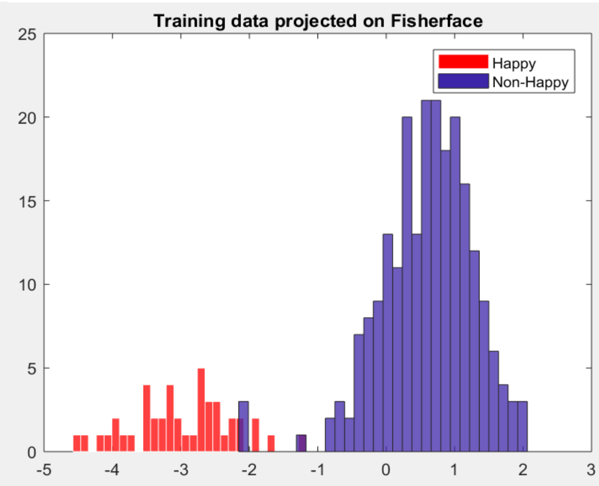

Figure 3(a): Threshold for Happy.

Figure 2(b): Surprise Fisherface.

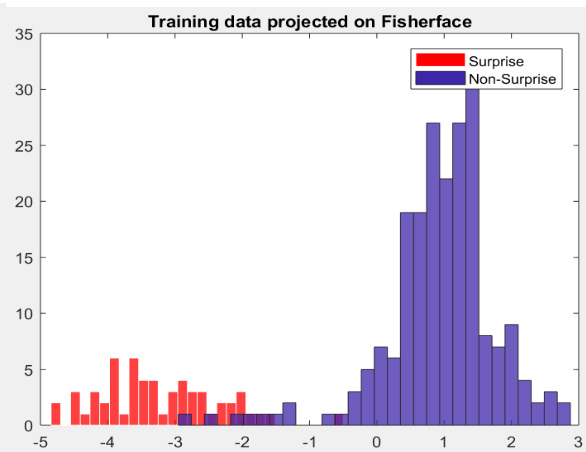

Figure 3(b): Threshold for Surprise.

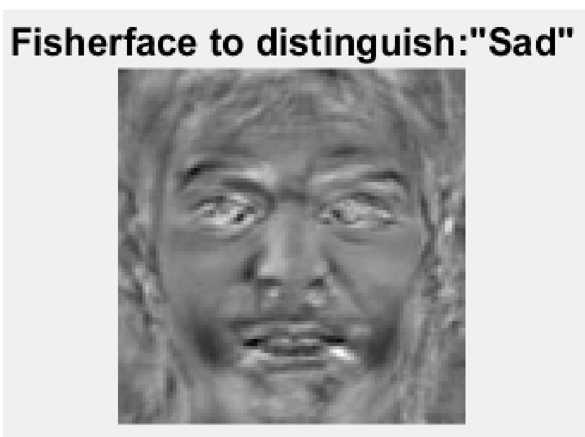

Figure 2(c): Sad Fisherface.

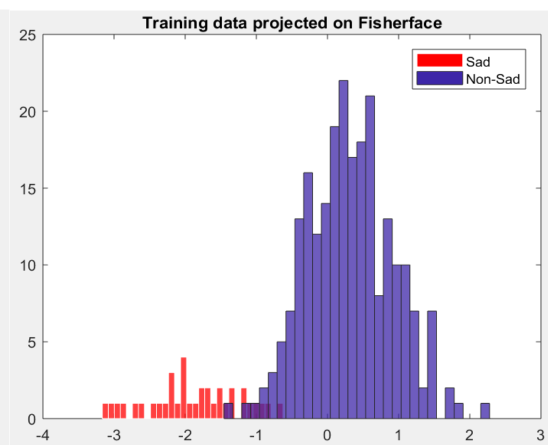

Figure 3(c): Threshold for Sad.

Figure 2(d): Fear Fisherface.

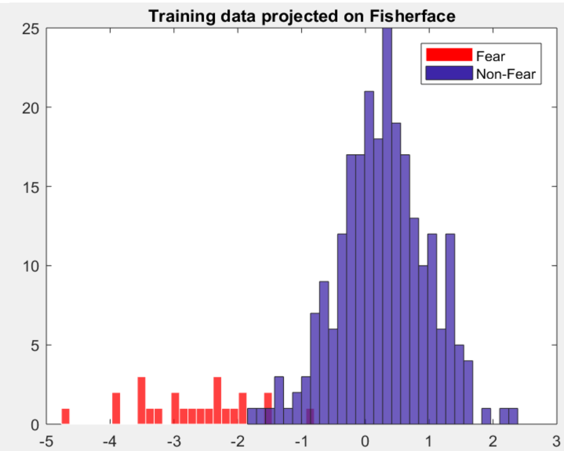

Figure 3(d): Threshold for Fear.

Figure 2(e): Angry Fisherface.

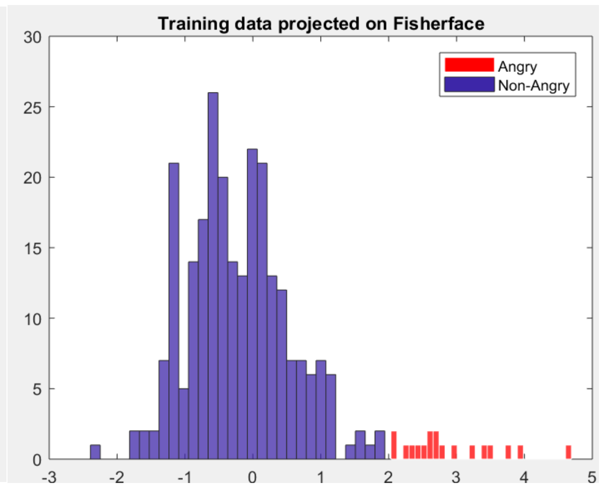

Figure 3(e): Threshold for Anger.

The testing results on test images from Cohn-Kanade using MATLAB are given in Table 1.

Algorithm	Accuracy
Fisherface only (Angry, Fear, Sad, Surprise, Happy)	90%
Fisherface only (All 7 emotions)	56%

Table 1: Testing results for classifier.

3. Discussion and Conclusion

Using the reduced dimensionality training dataset Fisher LDA was performed to extract Fisherfaces on which we can project test data. When we use test images from the Extended Cohn-Kanade dataset and project those images onto our Fisherfaces for classification based on our established thresholds, we have an accuracy of 56%. This is a poor result, as it is only marginally better than random- guessing. Upon further investigation, this is due to the Fisherface-approach's inability to effectively detect the expressions corresponding to disgust and contempt. However, when only detecting expressions of test images that correspond to anger, fear, happiness, sadness, and surprise, the Fisherface approach is more than 90% accurate.

This leads up to consider that anger, fear, happiness, sadness, and surprise are "easy-to-distinguish" emotions. This is likely attributable to the fact that the "easy-to-distinguish" emotions have very distinct and blatant feature manipulations associated with them. For example, the happiness expression has a very strong turning up of the mouth, which is seen to be strongly emphasized in the Fisherface for discerning happiness. The expression associated with fear has a very strong lateral pulling of the edges of the mouth, also evident in the associated Fisherface. The anger expression involves a downwards furrowing of the brow, the sad expression involves an obvious turning-down of the mouth, and surprise involves a very- obvious open mouth.

Contempt and disgust on the other hand, are much more difficult to detect, for potentially different reasons. It is possible that disgust is difficult to detect because it has feature orientations that are similar to those in several other emotions, such as an opening of the mouth that could be confused with happiness, fear, or surprise. The brows during a display of disgust is also furrowed similarly to the anger expression. The most tell-tale sign of disgust is an upward pulling of the nose, leading to wrinkling around the bridge of the nose. However, this detail is much more nuanced than the other more obvious expression characteristics, and can be lost during resolution reduction, mean-subtraction, and image misalignment. Contempt on the other hand, is difficult to detect since its characteristics are very faint in intensity. The expression for contempt is characterized by a neutral expression overall, with a unilateral pulling up of the mouth. This can be very difficult to distinguish as a human, so incorrect labeling of training data, as well as a methodological inability to capture the faint characteristics of the expression make contempt very difficult to detect.

4. Future Work

I wish to develop another classifier in addition to our Fisherface based classifier since, as we find out experimentally, the Fisherface approach is limited in success by itself. I leverage the fact that most expression information is encoded within the inner facial features, specifically the regions around the eyes, nose, and mouth. As is detailed in FACS, the inner facial features will move in certain distinct combinations with the exhibition of each emotion, as is described by Action Units.

References

[1] Lucey, P., Cohn, J.F., Kanade, T., Saragih, J. Ambadar, Z. The Extended Cohn-Kanade Dataset (CK+): A complete dataset for action unit and emotion-specified expression. IEEE Computer Society Conference CVPRW (2010).

[2] T. Mita, T. Kaneko, O. Hori, Joint Haar-like Features for Face Detection, “Proceedings of the Tenth IEEE International Conference on Computer Vision”, 1550- 5499/05 ©2005 IEEE.

[3] Computer Vision”, 2004. [9] M. A. Turk and A.P. Pentland, Face recognition using eigenfaces, “Proceedings of the IEEE”, 586-591, 1991.