Facial Expression Recognition Using Fisher Face

Digital Singal Processing (DSP),
School Of Engineering And Applied Science, Ahmedabad University.
Akhil Vavadia(1401095), Akshat Doshi(1401119), Pinak Divecha(1401098), Vatsal Patel(1401123)
Group-25

***Abstract*—In this paper we developed a facial expression recognition algorithm using Fisher Face method. Taking a pattern classification approach, we consider each pixel in an image as a coordinate in a high-dimensional space. The images of a particular face are under varying illumination but fixed pose. Our projection method is based on Fishers Linear Discriminant and produces well separated classes in a low-dimensional subspace, even under severe variation in lighting and facial expressions.**

***Index Terms*—Fisherface, Eigenface, Haar cascade, Viola Jones, Linear Discriminant Analysis (LDA), Principal Component Analysis(PCA).**

I. Introduction

Within the last several years, many algorithms have been proposed for facial expression recognition. While much progress has been made toward recognizing expression under small variations in lighting, facial expression and pose, reliable techniques for recognition under more extreme variations have proven elusive.There are many methods use for facial expression recognition for example,Nueral network,suppor vector machine,Eigenface,Fisherface(LDA),Regression.

II. Our approach

In this paper we have used Linear Discriminant Analysis or Fisherface algorithm for facial expression recognition. maximizes the ratio of between-class scatter to that of within-class scatter and on other hand the Eigenface method is also based on linearly projecting the image space to a low dimensional feature space which uses PCA and gives projection directions that maximize the total scatter across all classes. The PCA finds a linear combination of features that maximizes the total variance in data but it doesn't consider any classes and so a lot of discriminative information may be lost when throwing some components away. This can yield bad results, especially when it comes to classification. In order to find a combination of features that separates best between classes the Linear Discriminant Analysis instead maximizes the ration of between-classes to within-classes scatter. The idea is that same classes should cluster tightly together.

We have Cohn-Kanade dataset which have labeled image for different facial expressions. In our approach we have used 6 expressions and neutral face which act as 7 classes in LDA.

Fig. 1: happy

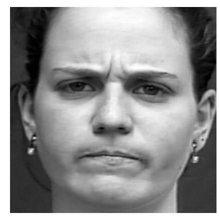

Fig. 2: anger

Fig. 3: fear

Fig. 4: disgust

Fig. 5: sad

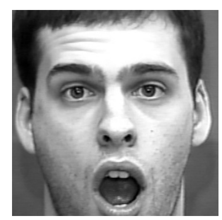

Fig. 6: surprise

General approach for facial expressoin recognitin is shown is Fig-7.The emotion sequences actually contain archetypical emotions. Each image sequence consists of the forming of an emotional expression, starting with a neutral face and ending with the emotion. So, from each image sequence we want to extract two images; one neutral and one with an emotional expression.

The classifier will work best if the training and classification images are all of the same size and have only a face on them. We need to find the face on each image, convert to grayscale, crop it and save the image to the dataset. We have use a HAAR filter for face finding.

Now we need to split the complete dataset into a training set and a classification set. We use the training set to teach the classifier to recognize the to-be-predicted labels, and use the classification set to estimate

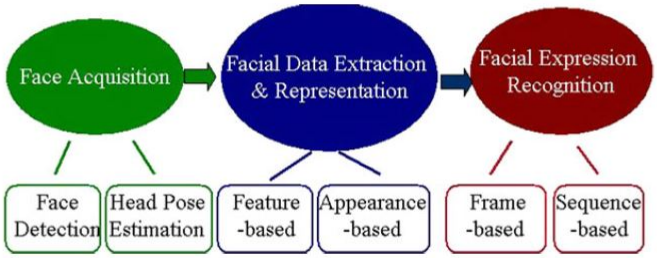

Fig. 7: General approach

the classifier performance. The reason for splitting the dataset is to estimating the classifier performance on the same set as it has been trained is unfair, because we are not interested in how well the classifier memorizes the training set. Rather, we are interested in how well the classifier generalizes its recognition capability to never-seen-before data.

For now lets create the training and classification set, we randomly sample and train on 80% of the data and classify the remaining 20%.

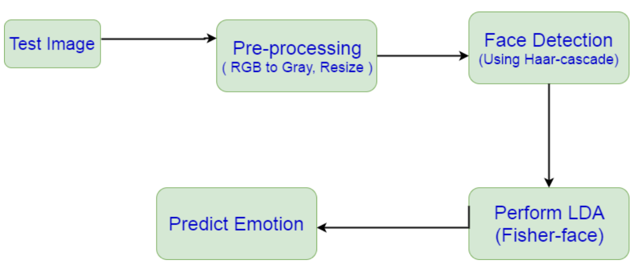

Fig. 8: Our approach

III. Viola-Jones algorithm

First step towards the facial expression recognition is to detect the faces from given set of images. So, for face detection and feature extraction we have use viola-jones methodology.

Facial expressions are generated by contractions of facial muscles, which results in temporally deformed facial features such as eye lids, eye brows, nose, lips and skin texture, often revealed by wrinkles and bulges. The most important thing is the location of facial actions, their intensity as well as their dynamics. Facial expression intensities is measured by determining the geometric deformation of facial features. The reasons behind using viola jones algorithm is it is, Robust very high detection rate very low false-positive rate always, Face detection - The goal is to distinguish faces from non-faces. Viola Jones Algorithm has 4 main steps.

[1] Haar Feature Selection
[2] Creating an Integral Image
[3] Adaboost Training
[4] Cascading Classifiers.

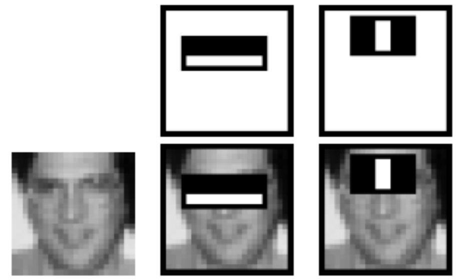

Fig. 9: Haar Feature

In the above image we can see that all human faces share some similar properties. These regularities may be matched using Haar Features. Where The eye region is darker than the upper-cheeks. The nose bridge region is brighter than the eyes.

As we can see in the above image the first and second features selected by AdaBoost.The two features are shown in the top row and then overlayed on a typical training face in the bottom row. The first feature measures the difference in intensity between the region of the eyes and a region across the upper cheeks. The feature capitalizes on the observation that the eye region is often darker than the cheeks. The second feature compares the intensities in the eye regions to the intensity across the bridge of the nose. Thus, the object detection framework employs a variant of the learning algorithm AdaBoost to both select the best features and to train classifiers that use them.

The overall form of the detection process is that of a degenerate decision tree, what we call a cascade. A positive result from the first classifier triggers the evaluation of a second classifier, which has also been adjusted to achieve very high detection rates. A positive result from the second classifier triggers a third classifier, and so on. A negative outcome at any point leads to the immediate rejection of the sub-window. Stages in the cascade are constructed by training classifiers using AdaBoost and then adjusting the threshold to minimize false negatives. Note that the default AdaBoost threshold is designed to yield a low error rate on the training data. In general a lower threshold yields higher detection rates and higher false positive rates.

IV. EigenFaces vs. FisherFaces

The Eigenface method is also based on linearly projecting the image space to a low dimensional feature space. However, the Eigenface method, which uses principal components analysis (PCA) for dimensionality reduction, yields projection directions that maximize the total scatter across all classes. In choosing the projection which maximizes total scatter, PCA retains unwanted variations due to lighting and facial expression. As illustrated stated by Moses- The variations between the images of the same face due to illumination and viewing direction are almost always larger than image variations due to change in face identity. Thus, while the PCA projections are optimal for reconstruction from a low dimensional basis, they may not be optimal from a discrimination standpoint. We should point out that Fishers Linear Discriminant is a classical technique in pattern recognition.

A. EigenFaces

The main idea of eigenface is to get the features in mathematical sense instead of physical face feature by using mathematical transform for recognition.

There are two phases for face recognition using eigenfaces. The first phase is the training phase. In this phase, a large group of individual faces is acted as the training set. These training images should be a good representation of all the faces that one might encounter. The size, orientation and light intensity should be standardized. For example, all images are of size 125 x 125 pixels and all are frontal faces. Each of the images in the training set is represented by a vector of size N by N, with N representing the size of the image. With the training images, a set of eigenvectors is found by using Principal Component Analysis (PCA).

Advantage of this mathod is recognition is simple and efficient compared to other matching approaches. Data compression is achieved by the low dimensional subspace representation. Raw intensity data are used directly for learning and recognition without any significant low-level or mid-level processing low-level or mid-level processing. No knowledge of geometry and reflectance of faces is required

A drawback of this approach is that the scatter being maximized is due not only to the between-class scatter that is useful for classification, but also to the within-class scatter that, for classification purposes, is unwanted information. Recall the comment by Moses: Much of the variation from one image to the next is due to illumination changes. Thus if PCA is presented with images of faces under varying illumination, the projection matrix will contain principal components (i.e., Eigenfaces) which retain, in the projected feature space, the variation due lighting. Consequently, the points in the projected space will not be well clustered, and worse, the classes may be smeared together.

B. Linear Discriminant Analysis

Linear Discriminant Analysis is a well-known scheme for feature extraction and dimension reduction. It has been used widely in many applications such as face recognition, image retrieval, microarray data classification, etc. Classical LDA 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 (transformation) can be readily computed by applying the Eigen decomposition on the scatter matrices.When LDA is applied to Face reconition application the obtained eigenvectors are called fisherfaces.

The advantage of fisherface projection approach is to solve the illumination problem by maximizing the ratio of between-class scatter to within-class scatter. However, finding an optimum way of projection that is able to simultaneously separate multiple face classes is almost impossible. LDA based algorithms outperform PCA based ones, since the former optimizes the low dimensional representation of the objects with focus on the most discriminant feature extraction while the latter achieves simply object reconstruction.

V. Algorithm Of LDA

[1] Construct the Image matrix X with each column representing an image. Each image is a assigned to a class in the corresponding class vector C.

[2] Project X into the $(N-c)$-dimensional subspace as P with the rotation matrix W_p identified by a Principal Component Analysis,
where,
N is the number of samples in X and
C is unique number of classes - length(unique(C)).

[3] Calculate the between-classes scatter of the projection P as,

$$S_b = \sum_{i=1}^{c} N_i * (\mu_i - \mu) * (\mu_i - \mu)^T$$

where,
μ is the total mean of P
μ_i is the mean of class i in P
N_i is the number of samples for class i.

[4] Calculate the within-classes scatter of P as

$$S_w = \sum_{i=1}^{c} \sum_{x_k \in X_i} (x_k - \mu_i) * (x_k - \mu_i)^T,$$

where,
X_i are the samples of class i
x_k is a sample of X_i
μ_i is the mean of class i in P.

[5] Apply a standard Linear Discriminant Analysis and maximize the ratio of the determinant of between-class scatter and within-class scatter. The solution is given by the set of generalized eigenvectors W_f of S_b and S_w corresponding to their eigenvalue. The rank of S_b is at most $(c-1)$, so there are only $(c-1)$ non-zero eigenvalues, cut off the rest.

[6] Finally obtain the Fisherfaces by $W = W_p * W_f$.

VI. Data Set

Here, we have taken Cohn-Kanade data set for your approach. There are total 593 images used in our approach. Out of the total images 80:20 of images are used for training and testing.

VII. Result

We have implemented two datasets to test accuracy of our approach.First we tried with standard cohn-kanade database.In this database we get accuracy near 80%.Also there is different accuracies at different run due to small dataset size.Using a larger dataset will probably enhance the detection quite a bit.

Secondly we tried our approach with random images downloaded from google,this gave us accuracy near 50%. So here are some of the mistake,

Fig. 10: Surprise as Happy

Fig. 11: Disgust as Sadness

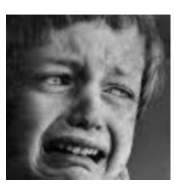

Fig. 12: Sadness as Disgust

Fig. 13: Anger as Happy

Fig. 14: Happy as Neutral

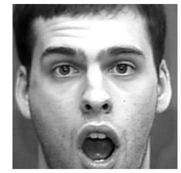

Fig. 15: Sad is detected as surprise.

VIII. Problem Faced

An intrinsic limitation of classical LDA is the so called singularity problem, that is, it fails when all scatter matrices are singular.

However, a critical issue using LDA, particularly in face recognition area, is the Small Sample Size (SSS) Problem. This problem is encountered in practice since there are often a large number of pixels available, but the total number of training samples is less than the dimension of the feature space. This implies that all scatter matrices are singular and thus the traditional LDA algorithm fails to use.

IX. Conclusion

Its clear that emotion recognition is a complex task, more so when only using images. Even for us humans this is difficult because the correct recognition of a facial emotion often depends on the context within which the emotion originates and is expressed.

References

[1] Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection by, Peter N. Belhumeur, Joao P. Hespanha, and David J. Kriegman.
[2] Y. Moses, Y. Adini, and S. Ullman, Face Recognition: The Problem of Compensating for Changes in Illumination Direction, European Conf. Computer Vision, 1994.
[3] https://en.wikipedia.org/wiki/ViolaJonesobjectdetectionframework
[4] http://disp.ee.ntu.edu.tw/ pujols/Eigenfaces%20and%20Fisherfaces.pdf
[5] http://csjournals.com/IJITKM/PDF%205-2/26SanjeevKumar.pdf