AOBD-ML Project Final Report

Group No. - 4: The Beta House

Janki Trivedi 121018
Abhay Soni 131001
Priyansh Shah 131037
Sarthak Mehta 131043

April 23, 2016

1 Abstract

Face Recognition has been widely used in many applications. There are many methods for face recognition which has been discussed in the paper. Here, LDA based face recognition has been implemented considering Mahalanobis distance as a classifier. The algorithm and results have been discussed in this paper

2 Introduction

Face recognition is a task that humans perform routinely and effortlessly in our daily lives. Face recognition system can help in many ways such as checking for criminal records, enhancement of security by using surveillance cameras in conjunction with face recognition system, detection of a criminal at public place, pattern recognition.

Face recognition, as one of the major biometric technologies, has become increasingly important owing to rapid advances in image capture devices such as surveillance cameras, camera in mobile phones etc. Face recognition is one of the few biometric methods that possess the merits of both high accuracy and low intrusiveness.

Before face recognition, we obviously need to detect the faces in an input image. The detection stage is the first stage; it includes identifying and locating a face in an image. The recognition stage is the second stage; it includes feature extraction, where important information for discrimination is saved, and the matching, where the recognition result is given with the aid of a face database. Face detection and face recognition both are classification problems. Face detection is a binary classification problem: whether a face or not a face. whereas face recognition is a multi-class classification (one-vs-all) problem. After face detection is performed, we apply face recognition technique only on the Region Of Interest (ROI) obtained during face detection.

Face detection : Face detection is a computer technology that determines the locations and sizes of human faces in arbitrary images. It detects facial features and ignores anything else, such as buildings, trees and bodies. Face detection can be regarded as a specific case of object-class recognition, a major task in computer vision. Software is employed to detect the location of any faces is being detected in the video. Generalized patterns of what a face looks like are employed to pick out the faces.

3 Objective

To recognize a face in a real time video with posture and Illumination variation by different approaches.

4 Literature Survey

4.1 Face Detection

4.1.1 The Viola-Jones Algorithm

In methodology, the procedure classifies images based on the value of simple features because the feature based system operates much faster than a pixel-based system.

The Viola-Jones face detector contains three main ideas that can detect faces successfully which runs in real time:the integral image , classifier learning with AdaBoost and the attentional cascade structure.

Integral Image

Integral image is summed area table which is for quickly and efficiently computing the sum of values in a rectangle subset of a grid. The integral image can be used to compute simple rectangular features. The features aredefined as weighted intensity difference between two to four rectangles.

AdaBoost

It is classifier learning with adaboost so in this a feature set and a training set of positive and negative images are given. Any approach can be used to learn a classification function. Here, AdaBoost is used to select a set o features and train the classifier. Boosting is a method to find a highly accurate hypothesis by combining many weak hypothesis where each of these hypothesis has a moderate accuracy. AdaBoost is Adaptive Boosting.

Originally, this algorithm is used to boost the classification performance of a simple learning algorithm.The weak algorithm selects the single rectangle feature which separates the positive and negative examples. And for each feature, it determines the optimal threshold classification function such that the least features are misclassified.

Attentional cascade

Attentional cascade is very evaluative constituent in this algorithm. It is smaller but more efficient and boosted classifier which can reject most of the negative sub-windows while keeping almost all examples the positive examples. So majority sub-windows are rejected in early stage, which makes the detection faster. Overall process is to degenerate decision tree, which is called cascade. Each node will make a binary decision whether the window will be kept for the next round or rejected immediately. If the decision thresholds were set too aggressively, the final detector will be very fast, but the overall detection rate may be hurt. On the other hand, if the decision thresholds were set very conservatively, most sub-windows will need to pass through many nodes, making the detector very slow.

Stages in the cascade are constructed by training classifiers using AdaBoost and then adjusting the threshold to minimize false negatives. To define in more simpler way, Initially, the algorithm has many positive images and negative images to train the classifier. The positive images would be images with face and negative images would be without face as it is for face detection. After that, extract features from it. For this, haar features are used. Each feature is a single value obtained by subtracting sum of pixels under white rectangle from sum of pixels under black rectangle. The sum of pixels under white and black rectangles is done by integral images.

Now some features are good but not applicable on every part of face. Adoboost selects the best feature out of all. To do this, apply each and every feature on all training dataset and find best threshold which can classify the faces into positive and negative. To check if a window is not face, cascade of classifier is helpful. We group the features into different stage of classifiers and apply one-by-one. If a window fails in the first stage itself, then after discarding it we dont consider other features but if passes then we continue the process and the window which passes all stages is a face region.

4.2 Face Recognition

4.2.1 Principal Component Analysis

The goal of PCA is to reduce the dimensionality of the data by finding the eigenvectors that span the space formed by the data. Eigenvectors are used to quantify the variation between multiple faces in the data. In this technique, we construct a matrix (A) of all the samples of all the persons. Then we calculate its covariance matrix (AA^T) and its corresponding eigenvectors and represent each face as a linear combination of the eigen vectors by calculating the set of weights. Now for face recognition, we project an unknown face on the face space and calculate its corresponding set of weights. Then we find the minimum distance and correspondingly classify it to that class (person). Here covariance matrix will be huge. So computation cost for covariance matrix will be huge. To overcome this cost, calculate the eigenvectors of matrix A^TA and multiply it with A. We only need to select few eigenvectors which significantly define the data. But there are some disadvantages like PCA does not take care of intra-class variation, its less effective in lighting conditions or pose variation.

4.2.2 Linear Discriminant Analysis

LDA improves PCA by taking care of inter-class as well as intra-class variation. LDA maximizes inter-class variation while minimizing intra-class variation. LDA searches for the projection axes on which the data points of different classes are far from each other while requiring data points of the same class to be close to each other. 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 inter-class variations and minimum intra-class variation by applying the Eigen decomposition on the scatter matrices. LDA is time consuming and requires a lot of memory storage. A smart move will be to apply PCA first and then LDA.

4.2.3 Local Binary Pattern

Both LDA and PCA gives low performance in lighting conditions. This problem is solved by introducing LBP technique. Local Binary Pattern (LBP) is quite robust against face images with different facial expressions, different lighting conditions, image rotation and aging of persons.

The Local Binary Pattern (LBP) operator is a non-parametric 3x3 kernel which summarizes the local spacial structure of an image. At a given pixel position (x_c, y_c), this operator works with the eight neighbours of a pixel, using the value of this centre pixel as a threshold. If a neighbour pixel has a higher gray value than the centre pixel or the same gray value than a one is assigned to that pixel, else it gets a zero. The LBP code for the centre pixel is then produced by concatenating the eight ones or zeros to a binary code.

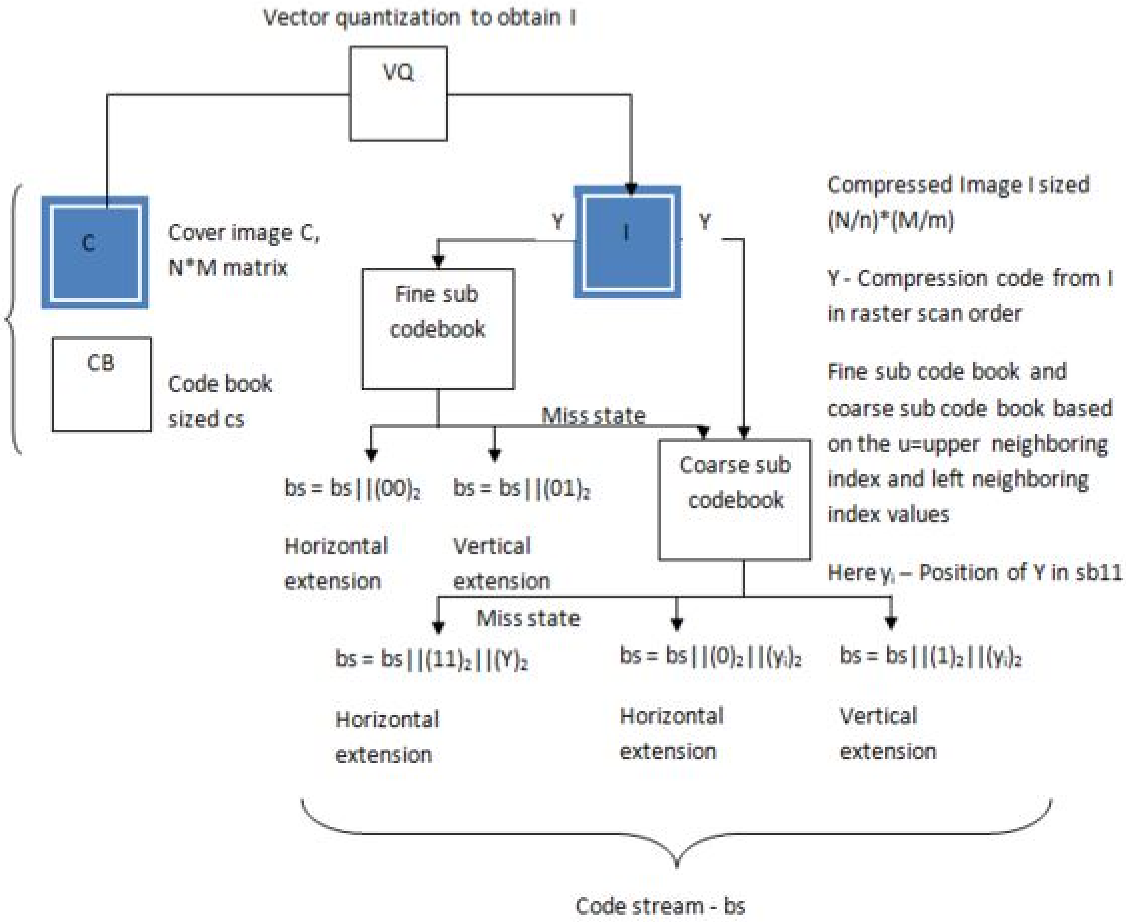


Figure 1: Local Binary Pattern

4.2.4 LDA Based Algorithms

The performance of classical LDA is often degraded by the fact that its classification accuracy is not directly related to the inter-class variation. A solution to this problem is to introduce weighting functions into LDA. There comes Fractional-step Linear Discriminant Analysis algorithm (F-LDA) where the dimensionality reduction is implemented in a few small fractional steps allowing for the relevant distances to be more accurately weighted. Classes that are closer together in the space can result in mis-classification so such classes should be more heavily weighted in the input space. But this method can not be applied to high dimension data due to high computational cost. As discussed earlier, smart move is to apply PCA and then LDA. While doing this, we may discard some significant eigenvectors while performing PCA. To avoid this, Direct LDA (D-LDA) method is used in place of a seperate PCA step where data is processed directly in the original high-dimensional input space avoiding the loss of significant information due to the PCA pre-processing step. Both these variants of LDA can be used to get accurate classification.

5 Motivation

Simple solution for face recognition will be to compare the input image vector with all the database image vectors. But the vector size is huge. So the computation cost is huge and the solution is also time consuming. To avoid such problems, we need to use dimension reduction methods (PCA and LDA). Here we have used only LDA because using PCA will only help in increasing fps (frames per second) and won't be responsible for an increase in accuracy. LDA is preferred over PCA as it takes care of inter-class as well as intra-class variations while the latter does not. PCA is preferred while for large database LDA is preferred.

6 Algorithm

```
Data: X , L
// data input and class categorising label vector ;
Result: W
// eigen vector W ;
initialization;;
unique(l)' ;
Sb=0;
Sw=0;
Mean(X) ;
while number of classes do
    calculate Sw;
    calculate Sb;
    // calculating sb and sw for each class and we append it;
end
[W,lambda]=eig(Sb^-1 Sw) ;
Y=W*X ;
```

Algorithm 1: Algorithm of LDA

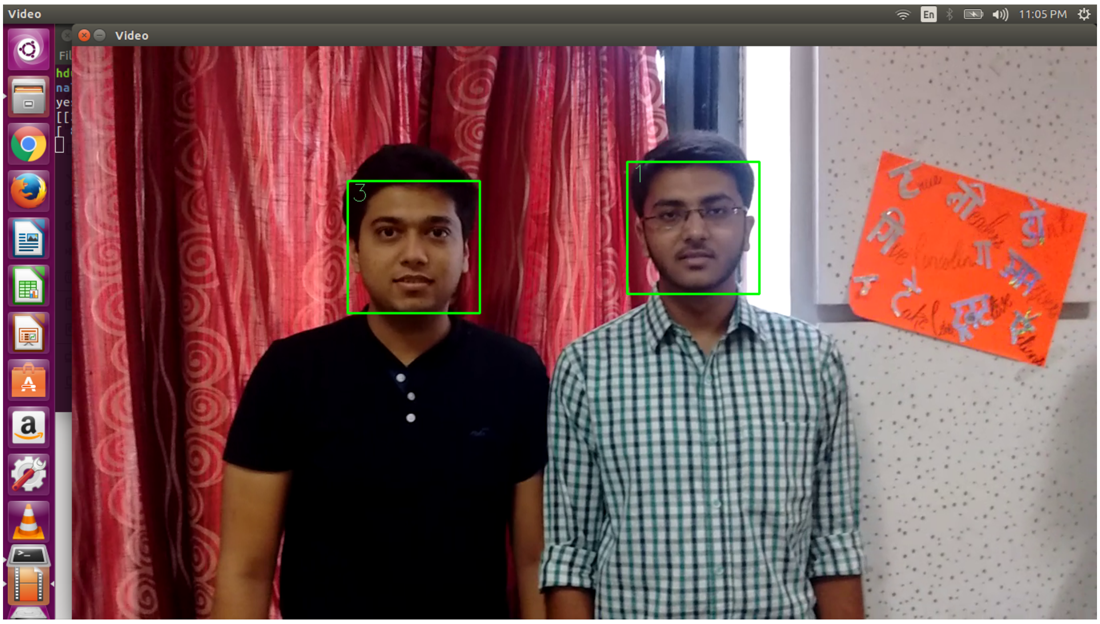
Data: Video Read
Result: Real time Image Output with face detection and recognition
initialization;
frame=0;
counter=0;
ROIPTS=0;
ROIframe=0;
X= train() // train faces for recognition in database;
L // label vector for Differentiating ;
W = LDA(X,L) // calculate eigenvector and eigenvalues via LDA;
while *true* **do**
    read current frame;
    counter++;
    if *counter%10==0* **then**
        Viola jones;
        ROIPTS=returns set 4 point for ROI;
        roi = frame[y,y+h, x:x+w];
        Calculate and Normalise HSV Histograme;
        tz = reshape(roi) // reshape 100 x 100 to 1 x 10000 proj = W*tz ;
        // project query image
        Mahalanobis distance between database and query projected
        image to classify;
        show frame with label having predicted person;
    end
    Thread 2;
    Read ROIPTS;
    Backprojection;
    Meanshifting;
    Track;
    show frame;
end

Algorithm 2: Algorithm for Face detection and Recognition by using LDA and CAMSHIFT (for tracking)

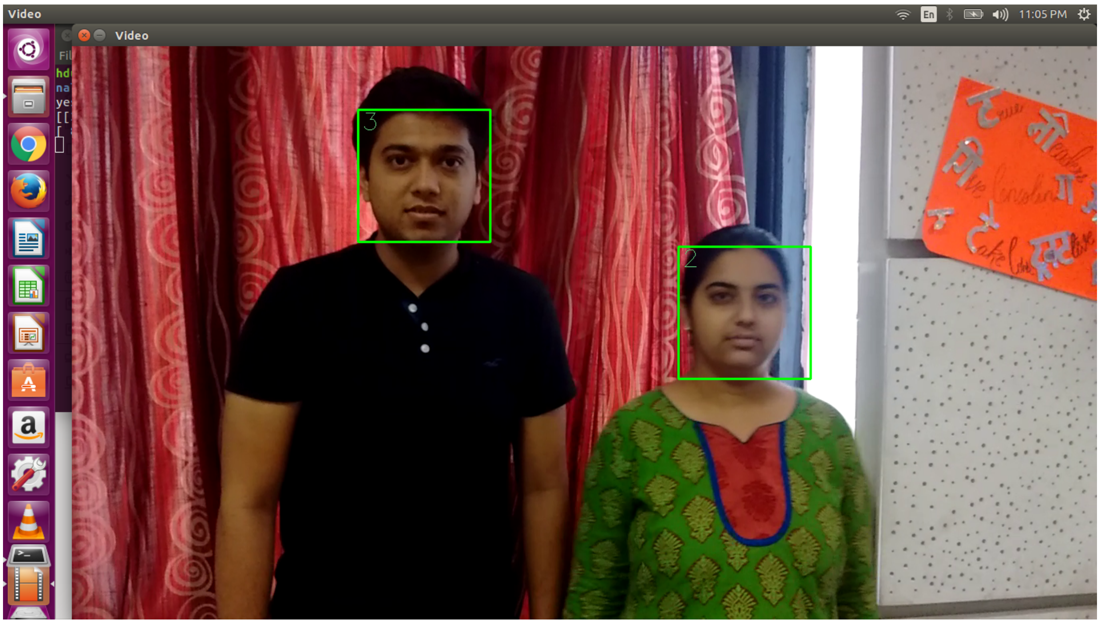
7 Result Interpretation

In this section, we discuss about the experimental results. We have taken a database of our mugshots. Our database has an total of 33 images, 11 images of each person (three in total). We take all these images and convert it in gray-scale and then normalize it. Then all the images are resized to dimension 100X100. Then the images are converted to 1D column vector. Now all such vectors of same class are appended column wise i.e. each vector to the right of another. Then the vectors of another class is appended row wise i.e. each vector of one class is appended just below the vector of another class. In this way, X is generated whose dimension here is 30000X11. Then we generate a label matrix which helps us to identify as to which class a particular vector belongs to. The label matrix's dimension will be 30000X1 i.e. for each feature of an image of a class there will be a label which depicts its class. Now, X and L both will be used in calculating projection vector matrix w. Now from matrix w, we select

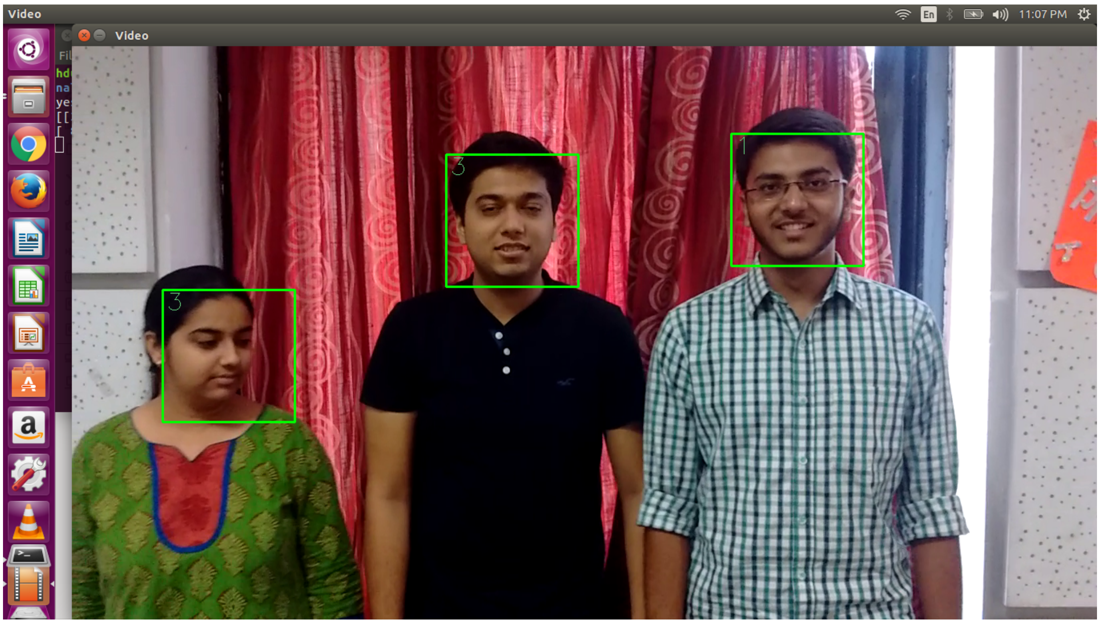
C - 1 significant columns, where C is the number of classes. Then we calculate projections Y = wt * x. Now we take each frame one by one from the video and detect the faces and apply recognition on those faces. Now each face detected will be reduced to dimension 100X100 and then converted to 1D column vector. Then we find projection of the face using the same projection vector matrix w. Then we use Mahalanobis distance as a classifier. So the detected face is classified to the class corresponding to the minimum distance. In this way all the detected faces are classified



[a] Accurate Classification



[b] Accurate Classification



[c] Misclassification

Class 1 is Abhay, Class 2 is Janki, Class 3 is Sarthak. Here, in Figure 2, classification is being done accurate but we are getting some misclassification as shown in figure 3

8 Conclusion

When using LDA and PCA both, there is a chance of loss of information. This is when F-LDA and D-LDA is used to overcome this problem. For extreme lighting conditions, pose variations, where LDA and PCA does not work good, LBP is used.

Above algorithm works well in case of face recognition of a single face at a time. The accuracy drops when there are multiple faces. This algorithm is still sensitive to the type of training data. More fps can be achieved by adding techniques like PCA, cam-shift and tracking for face detection. Classification accuracy can still be improved by taking the projection of the whole training set instead of the mean method Also, the distance metrics play an important role in accuracy of the classification

References

[1] P. Turk and A. Pentland, *Eigenfaces for Recognition*, Journal of Cognitive Neuroscience, March 1991.

[2] K. Etemad and R. Chellappa *Discriminant analysis for recognition of human face images*, Journal of Optical society of America, Vol.14, No.8, 1997.

[3] Md. Abdur Rahim, Md. Najmul Hossain,Tanzillah Wahid, Md. Shafiul Azam,, *Face Recognition using Binary Local Patterns (LBP)*, Global Journal of Computer Science and Technology Graphics and Vision, Volume 13 Issue 4 Version 1.0 Year 2013.

[4] S. Marcel, Y. Rodriguez, G. Heusch, *[2]On the Recent Use of Local Binary Patterns for Face Authentication*, International Journal of image and video processing, special issue on facial image processing

[5] T. Ahonen, A. Hadid, and M. Pietikainen, *Face Recognition with Local Binary Patterns.*

[6] J. Lu, K.N. Plataniotis, and A.N. Venetsanopoulos, *Face Recognition Using LDA Based Algorithms.*

[7] An Introduction to Face Recognition Technology, Shang-Hung Lin, Ph.D, IC Media Corporation, http://www.inform.nu/Articles/Vol3/v3n1p01-07.pdf

[8] http://what-when-how.com/face-recognition/introduction-to-face-recognition-part-1/

[9] http://disp.ee.ntu.edu.tw/ pujols/Face