A

DISSERTATION REPORT

ON

FACE AND EXPRESSION RECOGNITION USING PCA AND DISTANCE CLASSIFIERS

is submitted in partial fulfillment for the award of degree of

MASTER OF TECHNOLOGY

IN

EMBEDDED SYSTEMS

BY

LAVANYA MEKALA

(2015PEB5126)

UNDER THE GUIDANCE OF

Associate Prof. Rakesh Bairathi

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING

MALVIYA NATIONAL INSTITUTE OF TECHNOLOGY JAIPUR

JUNE 2017

© Malaviya National Institute of Technology Jaipur, 2016. All rights reserved.

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING

MALAVIYA NATIONAL INSTITUTE OF TECHNOLOGY

JAIPUR (RAJASTHAN)-302017

CERTIFICATE

This is to certify that the dissertation report entitled "**FACE AND EXPRESSION RECOGNITION USING PCA AND DISTANCE CLASSIFIERS"** submitted by **LAVANYA MEKALA** (2015PEB5126), in partial fulfillment of Degree **Master of Technology** in **EMBEDDED SYSTEMS** during the academic year 2016-2017. To best of my knowledge and belief that this work has not been submitted elsewhere for the award of any other degree.

The work carried out by her has been found satisfactory under my guidance and supervision in the department and is approved for submission.

Mr. Rakesh Bairathi

Associate Professor
Department of Electronics and Communication

MNIT Jaipur

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING

MALAVIYA NATIONAL INSTITUTE OF TECHNOLOGY

JAIPUR (RAJASTHAN)-302017

CANDIDATE DECLARATION

This is to certify that the dissertation report entitled "**FACE AND EXPRESSION RECOGNITION USING PCA AND DISTANCE CLASSIFIER**" is being submitted by me in partial fulfillment of degree of **Master of Technology** in **Embedded Systems** during **2016-2017** in a research work carried out by me under supervision of **Associate Prof. Rakesh Bairathi** and content of this dissertation work in full or in parts have not been submitted to any other institute or university for award of any degree or diploma. I also certify that no part of this dissertation work has been copied or borrowed from anywhere else. In case any type of plagiarism is found out. I will be solely and completely responsible for it.

LAVANYA MEKALA

Embedded Systems

M.TECH

2015PEB5126

Acknowledgement

I am grateful to my supervisor **Associate Prof. Rakesh Bairathi** for his constant guidance and encouragement and support to carry out this work. His excellent co-operation and suggestion provided me with an impetus to work and made the completion of work possible. He has been great source of inspiration to me, all through. I am very grateful to him for guiding me how to conduct research and how to clearly & effectively present the work done.

I would like to express my deepest sense of gratitude and humble regards to our Head of Department **Prof. K.K Sharma** for providing necessary facility in the Department. I am very thankful to **Associate Prof. Sanjeev Agarwal** and all other faculty members of ECE, MNIT for their valuable assistance and advice.

I would also like to thank my friends for their support in providing the database that is required for accomplishing this work and the discussions which proved valuable for me. I am indebted to my parents and family for their constant support, love, encouragement and sacrifices made by them so that I could grow up in a learning environment.

Finally, I express my sincere thanks to all those who helped me directly or indirectly to successfully complete this work.

(LAVANYA MEKALA)

Abstract

A face delivers genuine perception about expressions and emotions. The distortions of facial features and prompt variations provide more reliable cues than verbal communication. Thus, human communication is of 2 types: verbal and nonverbal. Here, nonverbal communication is interacting through facial expressions. Although human can identify faces and facial expressions of persons in real world with little effort, automatic recognition of face and expressions has been a true challenge to build an automated system which equals the human ability to recognise faces and different facial expressions. Because human-machine interaction is not as natural as interaction among humans. Automatic face recognition has been a exigent problem in image processing field and has gained a great attention over the last few eras because of its application in different fields like data privacy, different surveillance systems, security systems, authentication or verification etc. One difficulty in face recognition is how to handle variation in expressions, pose, and illumination conditions with limited training database.

The objective of my work is to recognise the face and expression of the person with different pose and with spectacles using Viola-Jones, principal component analysis and distance classifiers using Matlab. In the first stage to create database Images are taken in daytime in room light condition (similar illumination) with plain background using Samsung A7 mobile camera. Images of five persons with five different expressions including happy, sad, disgust, surprise, and neutral with different angle views (frontal, side, and top) and with spectacles are collected. Images are resized and pre-processed by median filtering to remove any salt and pepper noise present and illumination is compensated by Imadjust. Face database of persons with different expressions and views and Expression database of expressions of all persons and views are separately created. Viola Jones algorithm is used to detect the face part image, in this stage faces of all images are detected. in the third stage feature extraction is done by Principal component analysis. And features (Eigen vectors or eigen faces) of each person and each expression are stored respectively. Distance between projected train image and projected test image is calculated. Classification is done based on minimum distance by using Euclidean distance classifier and cosine angle distance classifier. And the performance of algorithm on face recognition and expression recognition using Euclidean distance classifier and cosine angle distance classifier is evaluated. Testing is done on 58 images which also include images with different angle views, and spectacles. Highest recognition accuracy of **86.2%** for face recognition and **84.48%** for expression recognition is obtained with Euclidean distance classifier while the training database has images of frontal, side, top views, and with spectacles. While the training database has only frontal faces the highest recognition accuracy of **68.96%** for face recognition and **63.79%** for expression recognition is obtained with Euclidean distance classifier.

Contents

List of Figures

List of Table

Chapter 1

Introduction

1.1 Motivation

There are two ways of communication between humans: one is verbal communication and the other is nonverbal communication. It is reported that 80% of the communication is done through nonverbal communication [48]. And face expression is a way of nonverbal communication between humans.

Interaction among humans is a natural process. Every person has a unique face and every expression is unique. Hence it is quite easy for humans to recognise people's faces and their expressions as well even in different illumination conditions, with different angle views, and with spectacles etc. but the same task poses a great challenge for computers in human-machine interaction. Automatic recognition of face and expressions has been a real challenging problem to develop a computerised system which matches the human capability to recognise faces and different facial expressions. automatic face recognition and expression recognition has received a great consideration in image processing field over the last few eras because of their application in different fields like data privacy, different surveillance systems, security systems, authentication or verification etc. Lot of research has been conducted and still going on to develop many robust algorithms to efficiently recognise human faces and expressions with variations in illumination, pose, and background etc. face and expression recognition is done using face detection, feature extraction and classification algorithms using Matlab.

1.2 Background

Automatic face recognition and expression recognition has received a great consideration in image processing field over the last few eras because of their application in different fields like data privacy, different surveillance systems, security systems, authentication or verification etc. Lot of research has been conducted and still going on to develop many robust algorithms to efficiently recognise human faces and expressions.

Various algorithms have been instigated on different static and non-static conditions [1].Static situations comprise of static and even background, identical postures, similar brightness conditions, neutral frontal faces. Non-static conditions comprise of position, partial occlusion alignment, varying illumination conditions and hair on face due to which recognition process becomes a complex problem. All these factors influence face recognition process [1].The various algorithms developed are of two techniques in the face and expression recognition practice. One technique considers facial features and the other technique considers the recognition problem’s holistic view.

a) Feature based approach: In this method local features extraction local features of the face like nose, mouth, eyes of the person’s face are identified and segmented. These local features are fed to the structural classifier for training. Examples include dynamic link architecture, hidden Markov model (HMM), and pure geometry [1] etc.

b) Holistic approach: It is also called global feature extraction method. The statistical features are extracted from the entire training sample images. These techniques include Fisher faces, Eigen faces, support vector machine (SVM), nearest feature lines, probabilistic Eigen faces, and independent component analysis [1].

c) Hybrid approach: The fusion of feature based and holistic approach is hybrid approach. These techniques include modular Eigenfaces, component based, shape normalized and hybrid local feature methods.

The technique implemented here is Principal component analysis which is a global feature extraction method. Variation in the lightening conditions also affect the recognition accuracy, hence lot of illumination compensation methods have been also introduced.

1.3 Software and Hardware Used

1.3.1 Software Used-

1.3.1.1 MATLAB for Software Simulation.

MATLAB is used as a simulation tool for face and expression recognition of a person. MATLAB (matrix laboratory) is developed by Mathworks. It provides libraries and function whose code is written in programming languages including C, C++, Fortran, java and Python [27].

1.3.2 Hardware Used

1.3.2.1 14MP Mobile Camera

14MP Samsung A7 Mobile Camera is used to take images. Images of five persons with five different expressions with different pose and with spectacles are taken in day time in room light conditions. These images are separated as training database and testing database.

1.3.2.2 Monitor

Monitor is used as a display device. The simulation results and recognised face and expression of the person are displayed on screen.

1.4 Outline of the Thesis

This thesis is about my work on face and expression recognition of a person using PCA and distance classifiers. The work is demonstrated in six chapters.

1.4.1 Chapter 1

This chapter has shed some light on motivation of the work, the challenges faced to build a robust recognition system, and latest trends in this field. This chapter also provides the list of software and hardware used.

1.4.2 Chapter 2

In this chapter various methods and approaches presented by researchers in different papers for face and expression recognition are analysed.

1.4.3 Chapter 3

In this chapter the various stages involved in face and expression recognition algorithm and different methods for each stage are illustrated.

1.4.4 Chapter 4

In this chapter implementation of my work is explained in an algorithm and through a flow chart.

1.4.5 Chapter 5

In this chapter Matlab simulation results are included. The results include pre-processing results, features extracted and final results showing recognised person and expression.

1.4.6 Chapter 6

This chapter shows conclusion and future work.

Chapter 2

Literature survey

In this section, the literature review of existing face recognition and expression recognition algorithms is discussed.

S.Agrawal and P.Khatri[1], "Facial Expression Detection Techniques: Based on Viola and Jones Algorithm and Principal Component Analysis," ***2015 Fifth International Conference on Advanced Computing & Communication Technologies*****, Haryana, 2015, pp. 108-112.**

S.Agrawal and P.Khatri [1] presented an algorithm for the expression detection of a person. The steps involved in the algorithm are reading an image, RGB to grey conversion, detecting the feature points, lightening compensation of the image, skin segmentation of the image, and then PCA is applied to extract global features of the faces. Euclidean distance is used to measure the nearest match to the test image. They have tested it in frontal faces of four different expressions namely sad, happy, neutral, and surprise and achieved an accuracy of 99.84% [1].

Meijing Li et al. [2] "Face recognition technology development with Gabor, PCA and SVM methodology under illumination normalization condition" Cluster Computing (2017).

In this paper they proposed a technique to compensate illumination variation of face images for face recognition process by fusing histogram equalization and Gaussian low pass filter. Gabor filter and principal component analysis are used for feature extraction. Further the classification is done using support vector machine (SVM)

D. Reney and N. Tripathi [3], "An Efficient Method to Face and Emotion Detection," ***2015 Fifth International Conference on Communication Systems and Network Technologies*****, Gwalior, 2015, pp. 493-497.doi: 10.1109/CSNT.2015.155.**

In this paper they implemented an algorithm for face and emotion detection of a person [3].separate databases have been created for face and emotion recognition. Viola Jones algorithm is used for face detection, and facial features are extracted and stored. Voice features are stored for emotion recognition by evaluating Mel frequency components [3]. Further emotion classification is done by using knn (k-nearest neighbour) classifier. They obtained a recognition accuracy of 94.5% to 97% [3].

Ekman et al. [4] developed a renowned and successful facial action coding system. "The Facial Action Coding System (FACS) identifies the facial muscles that cause changes in the facial expression thus enabling facial expression analysis" [4]. FACS comprises 46 Action Units (AU) describing the facial behaviours. They have used LEM (line edge map) descriptor and Hausdorff distance (calculates distance between a metric space's two subsets).They concluded that average recognition rate of females exceeds that of males by 7.8%. Zhang and Tjondronegoro [5] proposed Gabor feature extraction which is patch based from the automatically cropped faces. Support vector machine of four different kernels is used to carry the classification, and distance metrics to compare the patches of input image with train images. They obtained a recognition accuracy of 94.8% for CK database and 92.93% for JAFFE database.

Chapter 3

Face and expression recognition using image preprocessing

“The field of digital image processing refers to processing digital images by means of a digital computer using a set of precise algorithmic rules and mathematical methods. A digital image is composed of a finite number of elements, each of which possesses a particular value and location. These elements are signified as picture elements, image elements, pels, and pixels. Pixel is the term which refers to the elements of a digital image” [42].It permits us to exert different preprocessing algorithms to the input image making the image noise free. Abstraction of significant data from digital images to analyse the characteristics of the image is also image processing. “It can involve simple applications such as reading bar coded tags or more complex application such as face recognition. The applications of image processing are continually expanding in all areas of science such as- Robotics, machine vision, filtering, Defense, optical character recognition, remote sensing, Security, face detection, finger print detection in offices, medicine etc.” [38].

Face recognition and expression recognition algorithms basically involves three major stages namely face detection, feature extraction, and classification followed by image preprocessing steps to remove noise and to enhance image quality.

3.1 Image preprocessing

RGB to gray

“When converting an RGB image to grayscale, we have to take the RGB values for each pixel and make output as a single value reflecting the brightness of that pixel. One such approach is to take the average of the contribution from each channel: (R+B+C)/3. However, since the perceived brightness is often dominated by the green component, a different, more "human-oriented", method is to take a weighted average, e.g.: 0.3R + 0.59G + 0.11B" [44]. This conversion is done by retaining the luminance of the image while saturation is eliminated along with hue data.

Figure 3.1 RGB to Gray [43]

3.2 Image filtering and enhancement

Images are often corrupted by noise, variation in illumination and intensity. Hence, it is essential to remove noise by smoothening filters and enhance image contrast by image enhancement techniques.

3.2.1 Median Filter

Median filtering being a nonlinear process plays a great role in diminishing impulsive noise, which is also referred to as salt-and-pepper noise. Unlike other smoothening filters it has the competency to restore the discontinuities or edges in an image. In median filtering a filter window of a particular size is moved along the image. As the name suggests it replaces each pixel intensity value by the median value of the pixels surrounding it. The level of smoothening depends on the window's size that has been chosen.

For example, suppose the pixel values within a window are 5, 6, 26 and 15 whereas the pixel that is to be replaced has a value of 14 then the intensity value of that pixel is replaced by 14. The following figure shows the image which has noise in the form of white and black dots(pixels) and the smoothened image after noise removal by median filtering .

Figure 3.2 Median filter effects [28]

3.2.2 Gaussian filter for Smoothing

Gaussian filter has its impulse response as Gaussian function. The exertion of Gaussian filter on an input image which results in Gaussian blur or Gaussian smoothening on that image is accomplished by convolving the input digital image with the 2D Gaussian function which is given below.

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}$$

Figure 3.3 Gaussian filter effects (σ is standard deviation)[31]

3.2.3 Homomorphic filter

Homomorphic filter is a technique to perform image enhancement. It simultaneously performs brightness normalization and contrast stretching. Homomorphic filter eliminates unwanted random signal which is a multiplicativenoise [30]. Illumination disparities are considered as a multiplicative noise and can be reduced by filtering in the log domain i.e. frequency domain [29].

Illumination of the image is made more even by increasing the high-frequency components and decreasing the low-frequency component because the illumination in an image is represented by low-frequency components and reflectance in an image is represented by high-frequency components.

Figure 3.4 Homomorphic filtering effects [29]

3.2.4Wiener filter

The Wiener filter aims at computing an unknown signal's statistical estimate with the help of an associated signal which is an input and produces the estimate as an output by filtering the known signal [32]. The Wiener filter eliminates the noise from the corrupted signal providing an estimate of the unknown signal. The Wiener filter is used to restore the original image especially when the image is corrupted with noise and blurred.

Figure 3.5 wiener filter effects [33]

3.2.4 Histogram equalization

Histogram equalization is an image enhancement technique which improves the contrast in an image by adjusting the intensity levels (stretching the intensity range). Especially when the image is represented by close contrast values histogram equalization improves contrast by spreading out the most frequent intensity values [34].

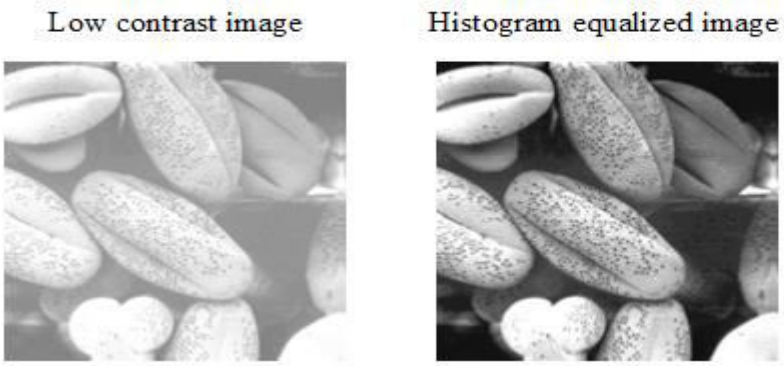


Figure 3.6 Histogram equalization effects[34]

3.2.5 Gamma correction

Gamma correction describes the relationship between numerical value of a pixel and its luminance. Without gamma, scenes captured by digital cameras will differ from how they appear to our eyes (on a standard monitor) [35].the overall brightness of an image is controlled by Gamma correction. Images with improper gamma correction look either too dark, or bleached out. Changing the extent of gamma correction changes both the brightness and the ratios of red to green to blue [36].gamma correction is also called gamma encoding or gamma compression [35].

Figure 3.7 Gamma correction effects[35]

3.2.6 Adjust Image

Adjust is a significant procedure that performs the mapping of intensity values of one image into another. The mapping is performed in such a way that a saturation level of 1 percent of data is maintained at high and low intensities there by the contrast of the output image is increased [12]. It enhances the recognition accuracy and gives better results.

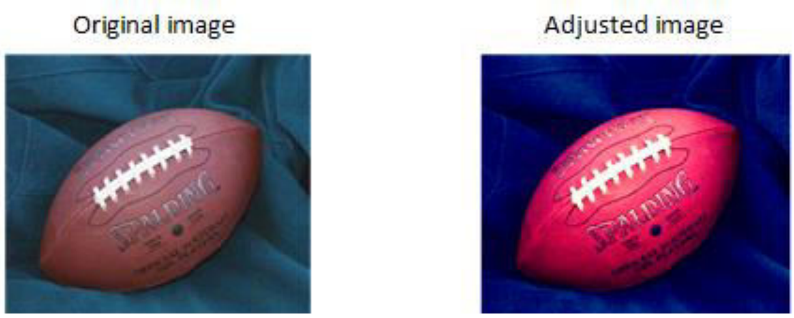


Figure 3.8 Adjust image[37]

3.3 Face And Expression Recognition

There are three major stages involved in face and expression recognition system.

1.Face detection

2.Feature extraction

3.classification

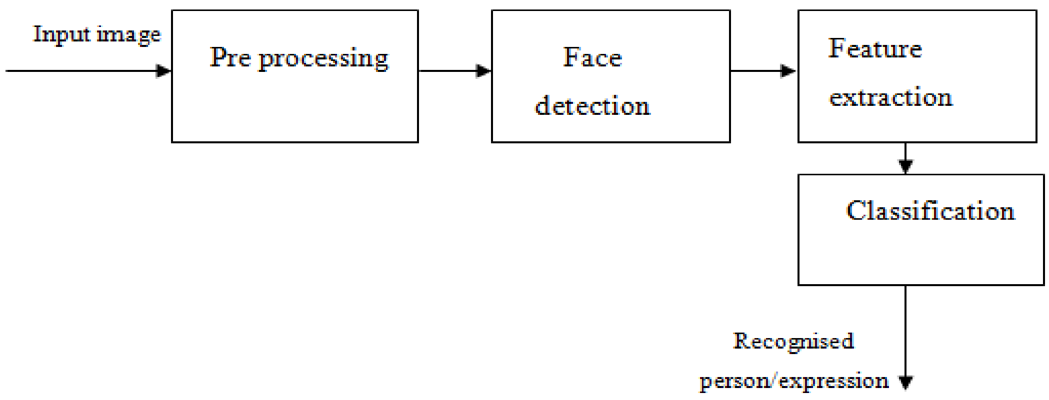


Figure 3.9 Architecture of face and expression recognition system

3.3.1 Face detection

Face detection is identifying the human face in an arbitrary scene eliminating the background. Face detection is the first stage in face and expression recognition based systems because all we need to analyze here is the face information and facial features. And face detection is done prior to feature extraction.

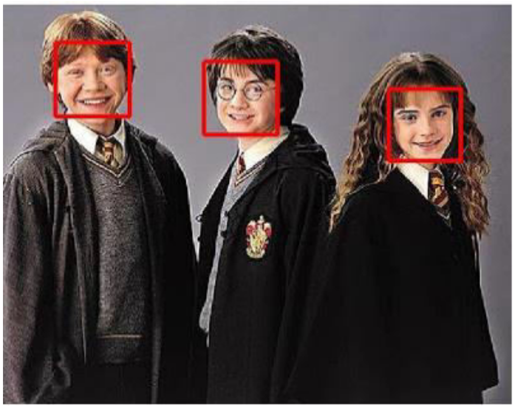


Figure 3.10 face detection in an image [47]

There are two ways of detecting faces one id from still images the other is detecting faces from videos. Different face detection techniques are reviewed below

3.3.1.1 knowledge based:

Yu-Buhee et.al [8] proposed a method in which knowledge of human face like shape of the face, skin colour, and face shape is directly set to a window size and colour signature for calculating the centre of difference with an aim to detect the face region with a rectangular window [8]. A detection rate of 93.4% was achieved with this method [8]

3.3.1.2 Feature based:

Kin Choong Yow et.al [10] proposed an algorithm which uses spatial filtering techniques to detect feature points from the image and then these face feature points are grouped into faces by perceptual grouping principles. The detection rate of 85% was achieved with this method.

3.3.1.3 Texture Based

In texture based face detection the textural features (properties) of the face are used for face detection. These texture features can be statistical or multi resolutional or combination of both.

3.3.1.4 Skin colour Based:

Sanjay Kr.Singh et.al [15] proposed an algorithm in which algorithm comparison is done by three colour spaces HIS, RGB, and YCbCr. a new skin colour was obtained by combining these three colours based on face detection algorithm.when the skin region is identified background colours are darkened to extract face features like eyes, nose, mouth. And now with a relevant threshold value this image is converted into binary image and grayscale image. An accuracy of 95.18% detection rate was achieved.

3.3.1.5 Multiple features

Jiaming Li et al. [16] proposed AdaBoost algorithm which selects features from multiple face region. Adaboostdivides the face into multiple sub regions based on multiple region orthogonal component principal component analysis features like eyes, mouth and nose. This region combination becomes input to AdaBoost classifier, it chooses best such combination for each stage and changes the weights for next iteration [16]. AdaBoost is a machine learning algorithm, it produces a strong classifier wich is a result of combination of only fewer visual features in a simple classifiers family.

3.3.1.6 Viola Jones Algorithm

Viola Jones algorithm is the oldest and most commonly used for face detection from the image. The basic concept of the Viola-Jones algorithm is scanning a sub-window that can identify faces in a given scene. In the conventional image processing approach input image is rescaled to different sizes and then a fixed size detector is run through these images. But the calculation of different size images is rather time-consuming. Viola Jones algorithm rescale the detector and image is not rescaled here. Though each time a different size detector is run through the image it is not time-consuming as the conventional standard approach. Because viola Jones detector is scale invariant which needs same number of calculations irrespective of the size of the detector. This detector is built using an integral image and few simple rectangular features which are evocative of Haar wavelets [17].

In the first step an input image is converted into an integral image. Each pixel in the integral image is equal to sum of all the pixels to the left and above of respective pixel in the input image [17]. It is explained in the below figure.

1	1	1
1	1	1
1	1	1

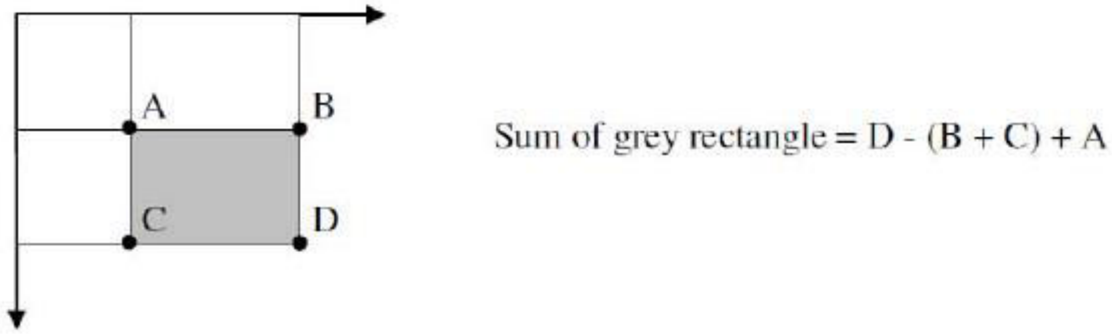
Input image

1	2	3
2	4	6
3	6	9

Integral image

Figure 3.11 integral image [3]

Integral image makes it easy to estimates of all pixels within a rectangle with the help of only four values. These are the pixel values which are part of the integral image that overlap with the rectangle’s corners in the input image. It is explained in the below figure.



F **Figure 3.12 sum of pixels calculation in a rectangle [3]**

In the above figure rectangle A is included in both the rectangles B and C. hence A has to be added to the calculation. This is how the sum of pixels inside rectangles of different sizes is evaluated in a fixed time. Few of the rectangle features are shown below.

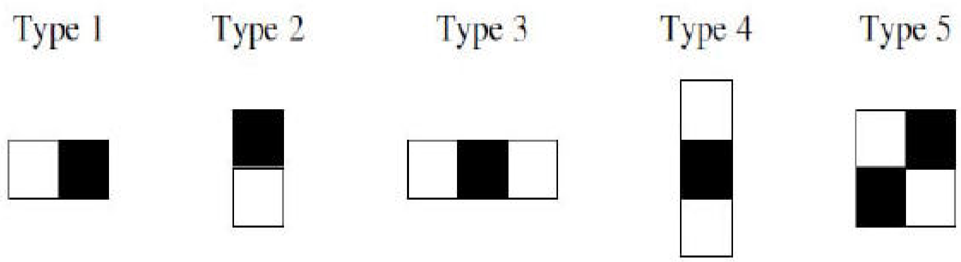


Figure 3.13 rectangle features (haar features) [3]

sum of the white rectangle(s) is subtracted from the sum of the black rectangle(s) and thus every feature yield in a single value [17].Viola-Jones has factually found in his experiments that satisfactory results are found with a detector whose base resolution is 24*24 pixels [3]. In the first step integral image is created and in the second stage a simple yet efficient classifier is constructed using AdaBoost algorithm which is a machine learning algorithm. this classifier selects only a few number of crucial visual features out of a very huge set of potential features [17].Then these classifiers are cascaded allowing to discard the background regions in the image and more computation is performed on likely face-like regions.

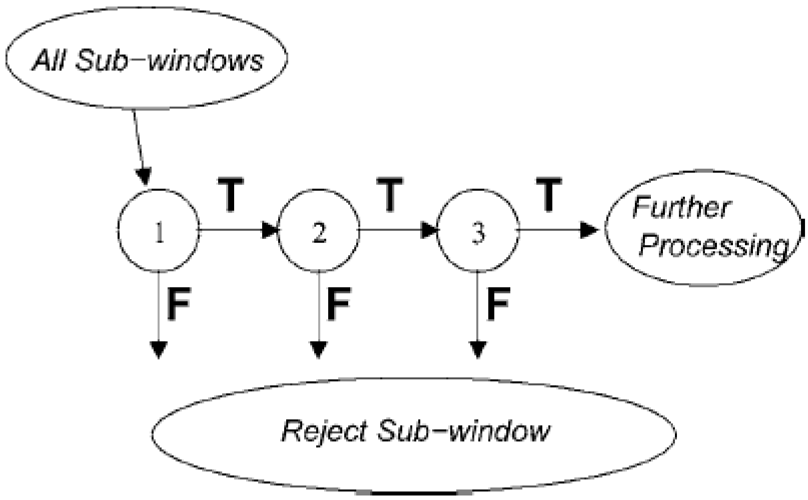


Figure 3.14 schematic diagram of detection cascade [17]

A sequence of classifiers is employed to each and every sub-window. A large number of negative samples are discarded by first classifier by minimum processing. Few more negative samples are discarded in the subsequent stage classifiers but take more computations. The task is more likely a decision tree making the task difficult for following classifiers.

3.3.2 Feature Extraction

when we need to process a large amount of data in the process of implementing any recognition algorithm there is a possibility that some of the data might be redundant (eg. Pixel repetition in images [18]). But computations on large amount of data requires more computational time and huge amount of memory. so there is a need for this data to be transformed into reduced set of features without losing the crucial information.Feature extraction is decreasing the quantity of data that is desired to describe a large volume of data. The features that can represent the entire initial data set with adequate accuracy are selected using different feature extraction or feature selection methods. Feature extraction is a stage involved in machine learning, image processing and pattern recognition algorithms [18].

Few important feature extraction methods in face and expression recognition are explained below.

3.3.2.1 Feature Extraction using Discrete Cosine Transform (DCT)

According to S. Dabbaghchian et al [19] dct based feature extraction consists of two stages. In the first stage dct is applied on an image and in the second stage few dct coefficients are chosen. When dct is applied on a a 2D image it results in a 2D matrix consisting of dct coefficients hence called coefficient matrix [19].each coefficient indicates frequency of the image. Low-frequency components i.e. the coefficients on top left corner of the matrix possess most critical information and the high-frequency components i.e. the coefficients on the bottom right possess redundant information. Here, a static Coefficient selection Method is used [19]. The most dominant coefficients are chosen using a fix mask in a zig zag manner [14].

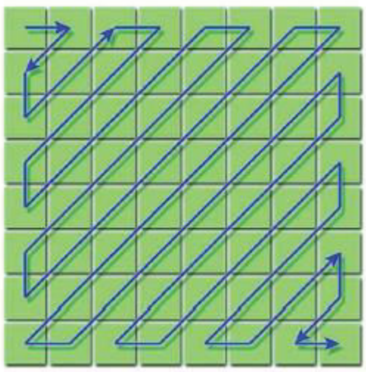


Figure 3.15 DCT coefficients with a zig-zag scan [14]

3.3.2.2 Facial Feature Extraction using Gabor Filter

Gabor filter is a linear filter [21] which is used for the analysis of texture, feature extraction and edge detection [22]. It analyses the presence of specific frequency in specific direction [21].The Gabor filter is denoted by a 2D mask. Here, mask is a 2D matrix of pixels where a value (weight) is allocated to every pixel. This mask is run through the image and a convolution procedure is executed. When a Gabor filter is applied to an image, it gives the peak response at edges and at points where there is a change in texture. A Gabor filter retorts to changes in texture and edges. A filter retorts to a specific feature, implies that the filter has a unique value at the spatial locality of that feature [22]. This is how features are extracted using Gabor filter.

Figure 3.16 original(left), gabor filtered image (right) [22].

3.3.2.3 Facial Feature Extraction using Principal Component Analysis:

Principal Components Analysis is a statistical procedure that can identify the patterns in data [23], and it can represent the entire data set emphasising the variances and likenesses. Once the patterns are found among the data set dimensionality reduction is done by reducing the number of data variables without incurring loss of information. in face and expression recognition algorithms the entire image data set is denoted by using few number of features which are Eigen faces extracted y principal component analysis. These Eigen faces are nothing but the Eigen vectors [24].

The steps involved in Principal component analysis are described below [25] [26].

Step 1: Obtaining face images I_1, I_2, I_3,…I_M (training faces)

Step 2: Represent every image Ii as vector.

N X N Image as N^2X1

Step 3: Average face vector of training faces is given by:

$\psi = 1/M (\sum Ii)$ (i=1 to M)

Step 4: Normalization of training faces

$\Phi = Ii - \psi$

Step 5: Covariance matrix C calculation

$C = P.P^T$, $P = [\Phi1, \Phi2, \Phi3......]$

Step6: Compute the Eigen vectors of $C=P.P^T$

The matrix is $P.P^T$ is very huge, so we consider $P^T.P$ matrix.

Step7: Compute the M best Eigen vectors of $P^T.P$

Step8: Select the top K Eigen vectors that can denote the entire training set of images

Step9: convert lower dimensional 'K' Eigen vectors to original face dimensionality.

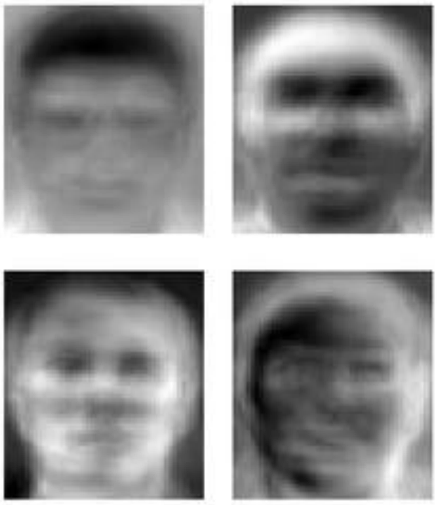


Figure 3.17 Few eigen faces of AT&T database [24]

3.3.2.4 Facial Feature Extraction using Independent Component Analysis:

Marian Stewart Bartlett et.al [7] proposed Independent Component Analysis "ICA which is derived from the principle of optimal information transfer through sigmoidal neurons. ICA was performed on FERET database face images using two different architectures, in one architecture the images were considered as random variables and the outcomes are pixels, and in the second architecture the pixels were treated as random variables and the outcomes are images. Spatially native base images for the faces were found in first architecture. A factorial face code was produced by second architecture. A classifier that combined the two ICA depictions gave the best performance" [13].

3.3.2.5 Facial Feature Extraction using Linear Discriminant Analysis

"LDA is alike PCA but linear discriminant analysis is supervised where as principal component analysis is unsupervised.PCA focuses on maximising the variance in a dataset by finding the component axes directions (principal components). LDA focuses on maximising the separation between multiple classes by computing the directions ("linear discriminants")" [42].

3.3.3 Classification

After the feature extraction of facial features, classifying these features conferring to our requirement is the final stage in any face and expression recognition algorithms.in face recognition task classification is made with respect to persons, here the numbers of persons become the number of classes. And in face expression recognition task classification is made with respect to expressions, here the number of expressions become the number of classes. Various classification techniques are described below.

3.3.3.1 Hidden Markov Model (HMM):

Hidden markov models are mostly used in expression recognition process. In a *hidden* Markov model the output is state dependant and it is visible. But the state is not visible as it is hidden. Each state is associated with a probability distribution to determine the output prediction. Hence, the state sequence associated information is provided by the generation of outputs by the HMM model [44].

3.3.3.2 Support Vector Machine

An SVM design demonstrates how the examples are portrayed as feature points in space. And how they are mapped making it possible to categorise examples of different classes with adequate gap in between them. When we need to predict the class of new examples we accomplish this task by considering the side of the gap in which these new examples are mapped [45].this is how SVM predicts the class of new test pattern by maximising the gap between categorised layers.

3.3.3.3 Neural Networks

Neural network is one of the trending classifier mechanisms which have solved many pattern recognition challenges including face recognition, character recognition, and expression recognition.

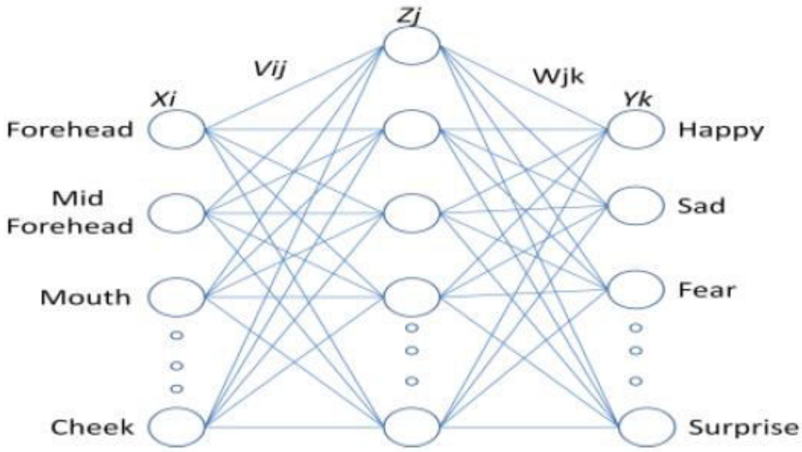


Figure 3.18 Architecture of feed-forward back propagation neural

.

The above figure shows the feed-forward back propagation of neural network [11].it has three different layers first layer is the input layer, the second layer is the hidden layer and the third layer is the output layer. Neurons present in the first layer deliver the information about forehead, mouth, cheek, and mid forehead etc to the second layer [11]. The second hidden layer performs computations required for the third layer to provide the output. The third layer delivers the expression as the result.

3.3.3.4 Distance classifiers

Distance classifiers are the simplest yet efficient classifiers in recognition algorithms. Similarity or dissimilarity between two images is measured by distance metrics, based on the minimum distance the test image is classified. Various distance measurements between images are explained below:

a) Euclidean distance:

It is the simplest and fastest distance calculation for similarity measurement. Euclidean distance computes the root of square differences amid the coordinates of a pair of entities [6]. It is a dissimilarity measurement so minimum the distance maximum is the similarity.

The Euclidean distance between m, n $\in R^T$ is computed by [9],

$$\mathbf{T\ (m,\ n)} = \|\boldsymbol{m}-\boldsymbol{n}\|_2 = \sqrt{\sum_{l=1}^{T}(\boldsymbol{m}_l-\boldsymbol{n}_l)^2}$$

Euclidean distance is calculated between the reconstructed test image and train images using the extracted Eigen space. The test image belongs to the class with smallest distance.

b) Cosine angle distance:

Angular distance or cosine similarity is measured between two non zero vectors. To find the cosine angle similarity between two images first the images are converted into vectors and the computation is performed. It is a similarity measurement hence higher the value higher the similarity. Hence it is subtracted from one to represent maximum similarity with minimum distance. The formula for cosine distance is given below [9]:

$$\mathbf{T\ (m,\ n) = 1\text{-}|\cos(\sigma)|}, \quad \cos(\sigma) = \frac{\boldsymbol{m}.\boldsymbol{n}}{\|\boldsymbol{m}_2\|\|\boldsymbol{n}_2\|}$$

c) Mahalanobis distance:

P.C. Mahalanobis introduced the Mahalanobis distance in 1936.For the identification and analysis of different patterns correlation amid the data variables is considered. The formula for Mahalanobis distance between two vectors is given below [9].

$$\mathbf{T\ (m,\ n)} = \sqrt{\left(((\boldsymbol{m}-\boldsymbol{n})^{T})\boldsymbol{X}^{-1}\left((\boldsymbol{m}-\boldsymbol{n})\right)\right)}$$

Where $\mathbf{x}$ is the covariance matrix, $\mathbf{m}$ and $\mathbf{n}$ are vectors.

d) City Block distance:

The absolute differences evaluated between two vectors are summed up to give city-block distance. It is also called L1 distance or Manhattan distance [9]. The formula to calculate Manhattan distance is given below [49]:

$$\mathbf{T\ (m, n)} = \|\boldsymbol{m} - \boldsymbol{n}\|_{\mathbf{1}} = \sum_{\boldsymbol{p}=\mathbf{1}}^{\boldsymbol{q}} \left|\boldsymbol{m}_{\boldsymbol{p}} - \boldsymbol{n}_{\boldsymbol{p}}\right|$$

Where $\mathbf{m}$ and $\mathbf{n}$ are vectors.

Chapter 4

Implementation

In the previous chapter we discussed about the various stages of the algorithm and different methods to implement those stages. In this chapter we are going to discuss the implementation of face and expression recognition algorithm.

The mechanism to recognise face and expression of a person involves three major stages including face detection, feature extraction and classification followed by pre-processing steps. The implementation is done on MATLAB. Images are taken in daytime in room light condition with uniform background using Samsung A7 mobile 8MP camera. Images of 5 persons with 5 different expressions including happy, sad, disgust, surprise, and neutral with different pose (frontal, side, and top) and with spectacles are collected. Images are pre-processed by median filtering, and Imadjust. Face database of persons with different expressions and pose and Expression database of expressions of all persons and different pose are separately created. Now, feature extraction is performed on every class (every person's faces) of face data base and also on every class (every expression) of expression data base. These extracted features are stored separately in training phase. Test image is projected onto these feature space i.e. eigen space and distance between projected test image and projected train image is calculated by using cosine angle distance and Euclidean distance measure. Minimum distance decides the class of the face and expression. These steps are demonstrated below briefly.

4.1 Algorithm

The implementation basically involves two stages. In the first stage we perform training of the data base and in the second stage we perform testing of images.

4.1.1 Training stage:

In this stage data base is created for face and expression recognition mechanism.in the face data base we have separate class for each person with their images. And in expression data base we have separate class for each expression with images of each person. The images captured were of different pose and with spectacles also. Next, feature extraction is done using PCA and extracted features are stored separately.

Step1: Image pre-processing

- Read image.
- Convert RGB image to grayscale image.
- Median filter is applied to remove salt and pepper noise if present any.
- Illumination compensation is done by Imadjust [12].

Step 2: Face detection of images using Viola Jones algorithm.

- Viola Jones algorithm is applied on the images which are pre-processed in the previous stage. In this stage we detect the face part from the given image ignoring background and other area because the required information for face and expression recognition lies only in the face part.
- The results obtained in this stage are detected faces of the persons with different expressions which are resized to 200*200 size for further processing.

Step 3: Feature extraction using Principal Component Analysis (PCA).

- PCA is applied separately on face data base and expression data base. For face recognition PCA extracts features separately from different person's faces of face data base. And then PCA extracts features from different expression images of expression data base. And these extracted features are stored separately in matlab.dat file.

- PCA applied on every class of data bases involves the following steps
    1. Calculate the mean face of all the images.
    2. Subtract the mean from all the images.
    3. Convert two dimensional images into one dimensional column vectors.
    4. Calculate covariance matrix A of these column vectors.
    5. Compute Eigen vectors of these covariance matrix using
       [A2, eigVec, eigValue]=pca(A).
       A2 matrix contains coefficients of principal components.
       We need to extract eigen vectors which are called eigen faces or eigen space.
    6. Now these eigen space of each class of face data base and expression data base are stored separately.

4.1.1 Testing stage:

In the testing stage we have tested 58 unique images of five persons with five expressions including happy, sad, surprise, disgust and neutral. These images also include images of different pose variation and with spectacles. The steps involved in testing are demonstrated below:

Step 1: read the test image and perform the preprocessing steps including RGB to gray conversion, median filter and imadjust for image enhancement.
Step 2: Detect the face part of the image using Viola Jones algorithm.

Step 3: now reconstruct this test image separately using the top best Eigen faces extracted from different classes of data base. This reconstructed image is called projected test image.

Step 4: Even the train images are projected onto their respective eigen space and the distance between the projected train images and projected test image is calculated using distance measures namely Euclidean distance and cosine angle distance.

Step 5: minimum distance is chosen among the distances calculated above. Minimum distance for the images of face data base gives the recognised person's face. And Minimum distance for the images of expression data base gives the recognised expression.

Step 6: recognised person's name and recognised expression are displayed in a message box using Matlab.

4.2 Flowchart of the Mechanism

4.2.1 Flowchart of the training stage

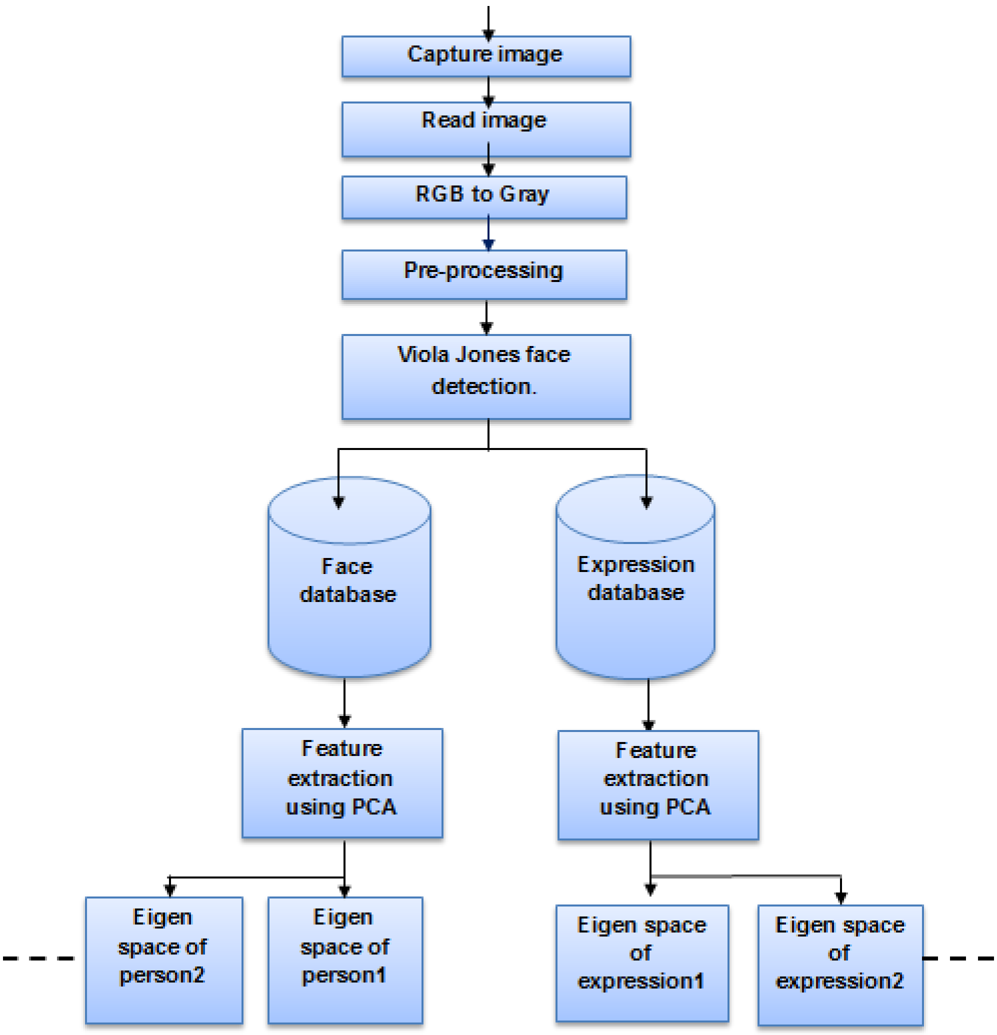


Figure 4.1 Flowchart for training stage

.

4.2.2 Flowchart of the testing stage

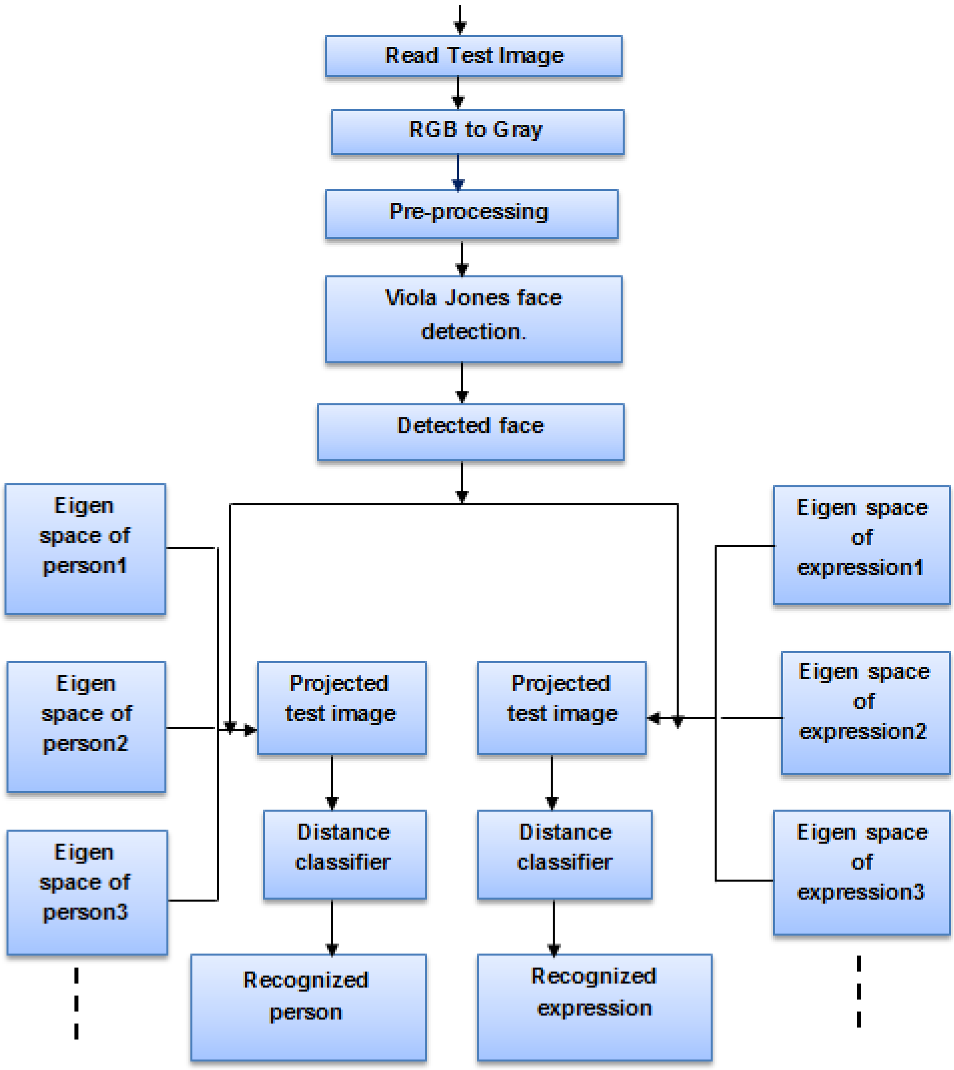


Figure 4.2 Flowchart for testing stage

Chapter 5

Simulation Results

In the previous chapter we have discussed the algorithm implementation of face and expression recognition. Now in this chapter we are going to show the simulation results of each stage of the algorithm implementation. The different stages of algorithm implementation are image pre-processing, face detection, feature extraction followed by classification. At the end the recognised face and expression are displayed.

5.1 RGB to Gray

The images captured are converted from RGB to Gary after reading through Matlab software.

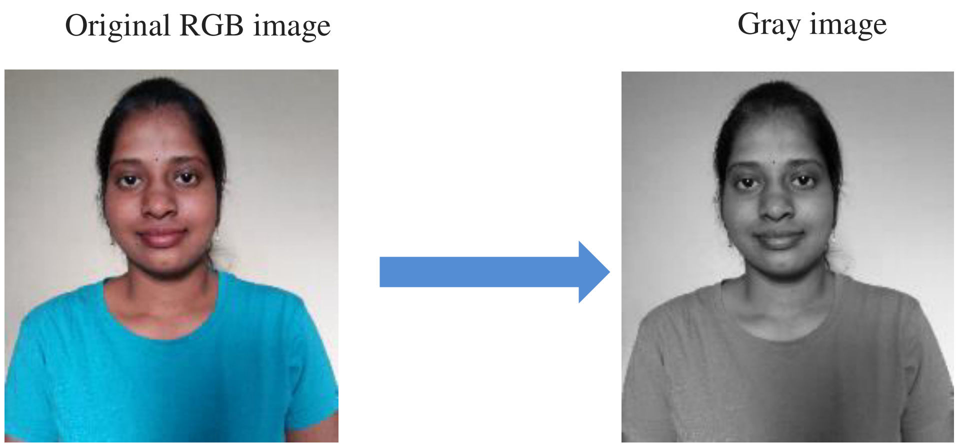


Figure 5.1 RGB to gray conversion.

5.2 Face detection by Viola Jones Algorithm

Face part is detected from the given image by Viola Jones algorithm. The results are shown below.

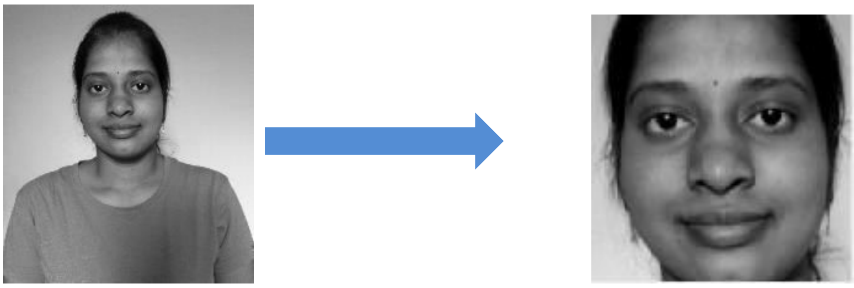


Figure 5.2 Face detection by Viola Jones algorithm

5.3 Image Enhancement

The images are filtered by median filter to remove salt and pepper noise. After filtering illumination compensation is performed by Imadjust [12].

Figure 5.3 Effects of image filtering and illumination compensation

5.4 Feature extraction using PCA

After the face detection and image enhancement we need to extract features from each class of face data base and expression data base. A class here indicates faces of a person in face data base and images of an expression in expression data base.

5.4.1 Feature extraction from different classes of expressions

Each class of expression consists of faces with single expression of five persons.

5.4.1.1 Neutral faces

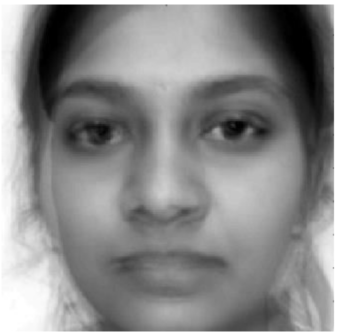


Figure 5.4 Neutral mean face **Figure 5.5 Eigen space of neutral faces**

5.4.1.2 Sad faces

Figure 5.6 Sad mean face **Figure 5.7 Eigen space of sad faces**

5.4.1.3 Happy faces

Figure 5.8 Happy mean face **Figure 5.9 Eigen space of happy faces**

5.4.1.4 Surprise faces

Figure 5.10 Surprise mean fa

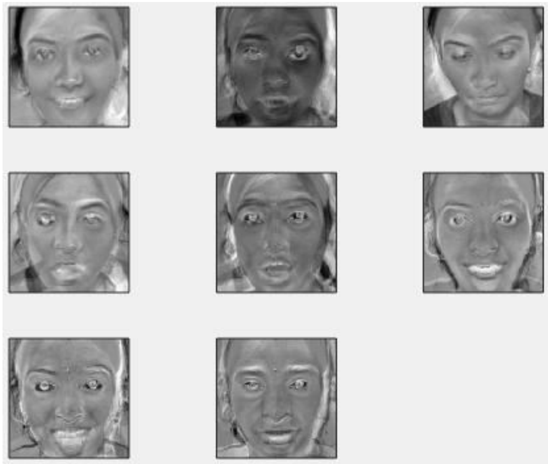


Figure 5.11 Eigen space of surprise faces

5.4.1.5 Disgust faces

Figure 5.12 disgust mean face

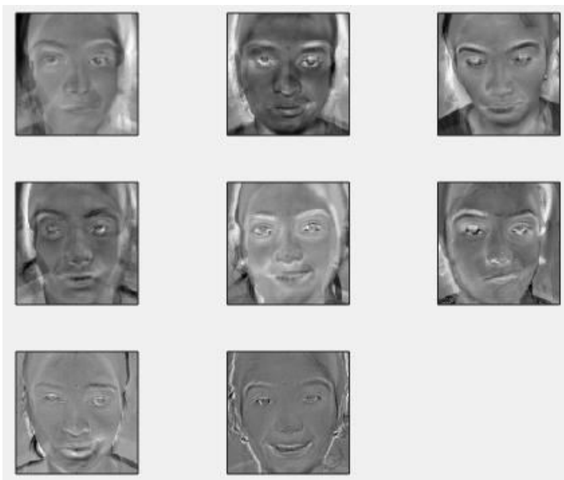


Figure 5.13 Eigen space of disgust

5.4.2 Feature extraction from different classes of persons

Each class of person consists of faces of single person with five different expressions.

5.4.2.1 Faces of Person1

Figure 5.14 Mean face of person1

Figure 5.15 Eigen space of person1

5.4.2.2 Faces of Person 2

Figure 5.16 Mean face of person 2

Figure 5.17 Eigen space of person 2

5.4.2.3 Faces of person 3

Figure 5.18 Mean face of person3

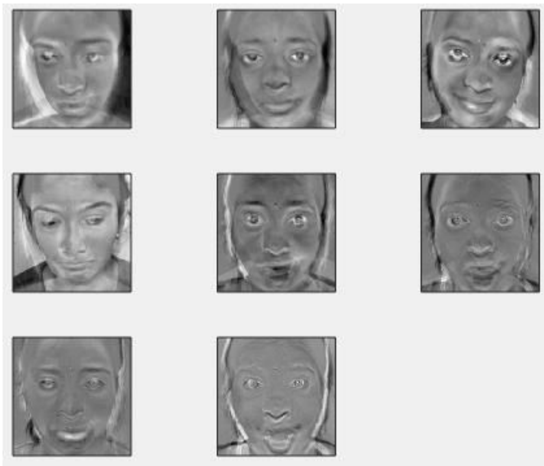


Figure 5.19 Eigen space of person3

4.2.4 Faces of Person 4

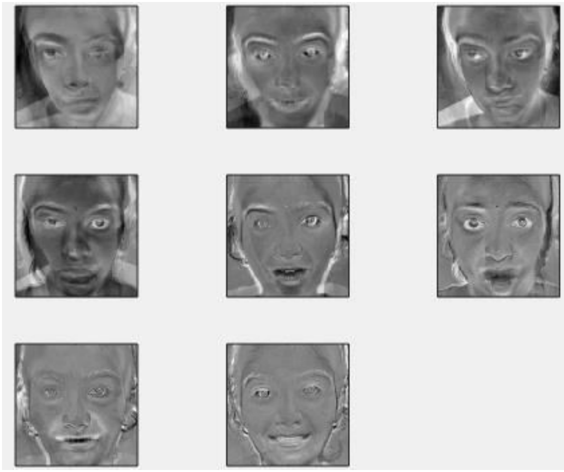


Figure 5.20 Mean face of person4

Figure 5.21 Eigen space of person 4

5.4.2.4 Faces of Person 5

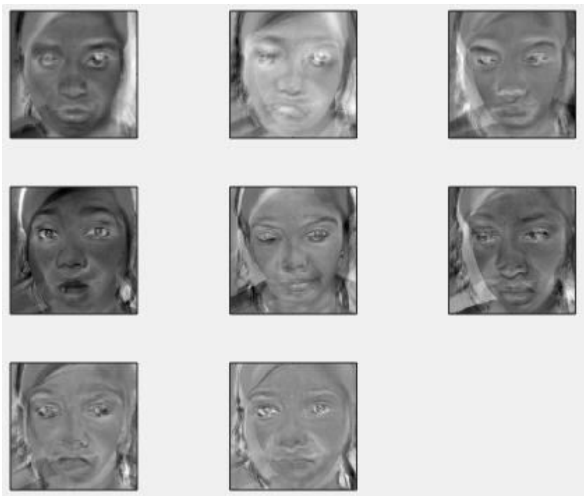


Figure 5.22 Mean face of person 5

Figure 5.23 Eigen space of person 5

5.5 Testing stage

Testing is performed on 58 unique faces including five expressions of five persons. The testing data base is shown below. These images include faces of different poses and images with spectacles as well.

Figure 5.24 Testing Data Base

5.5.1 Test Image 1

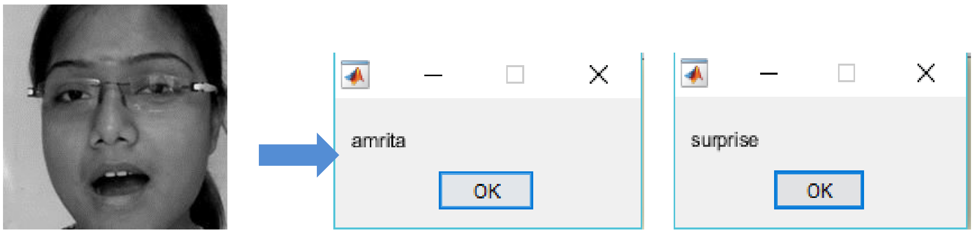


Figure 5.25 recognized face and expression of test image1

5.5.2 Test image 2

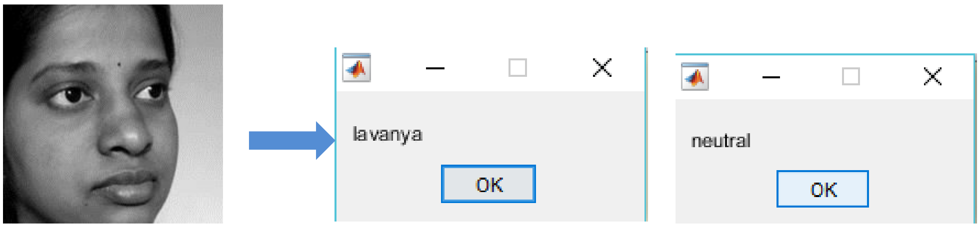


Figure 5.26 recognized face and expression of test image2

5.5.3 Test image 3

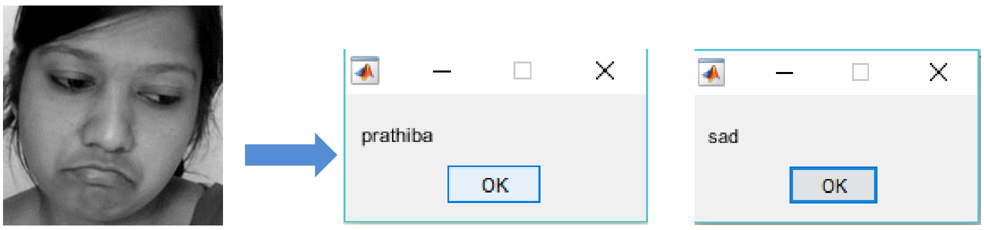


Figure 5.27 recognized face and expression of test image3

5.5.4 Test image 4

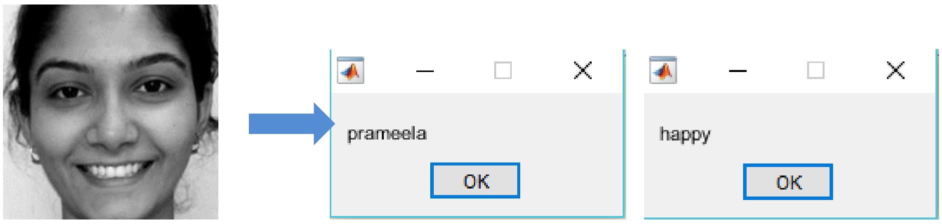


Figure 5.28 recognized face and expression of test image4

5.5.5 Test image 5

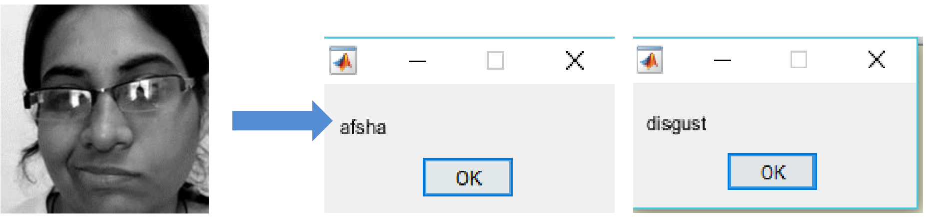


Figure 5.29 recognized face and expression of test image5

Chapter 6

Conclusion and Future Work

Conclusion- Face and Expression recognition algorithm implemented here is based on Viola Jones face detection, PCA based feature extraction followed by Distance classifier. Testing is performed on 58 unique images including frontal faces, images with different views (poses) while the training is done in two ways.

Training database with images of frontal, side, top views and with spectacles.

	Face recognition		Expression recognition	
Test images	Euclidean distance	Cosine angle distance	Euclidean distance	Cosine angle distance
1)frontal faces	26/30 **86.66%**	18/30 **60%**	24/30 **80%**	24/30 **80%**
2)faces with different pose orientation	10/11 **90.90%**	9/11 **81.81%**	11/11 **100%**	8/11 **72.72%**
3)faces with spectacles	14/17 **82.35%**	13/17 **76.47%**	14/17 **82.35%**	15/17 **88.23%**
Recognition accuracy (%)	50/58 **86.2%**	40/58 **68.96%**	49/58 **84.48%**	47/58 **81.03%**

Table 1: Recognition accuracy 1

Training database with only frontal faces.

	Face recognition		Expression recognition	
Test images	Euclidean distance	Cosine angle distance	Euclidean distance	Cosine angle distance
1)frontal faces	26/30 **86.66%**	18/30 **60%**	21/30 **70%**	25/30 **83.33%**
2)faces with different angles	4/11 **36.36%**	5/11 **45.45%**	7/11 **63.63%**	2/11 **18.18%**
3)faces with spectacles	10/17 **58.82%**	12/17 **70.59%**	9/17 **52.94%**	6/17 **35.30%**
Recognition accuracy (%)	40/58 **68.96%**	35/58 **60.34%**	37/58 **63.79%**	33/58 **56.89%**

Table 2: Recognition accuracy 2

Highest recognition accuracy of **86.2%** for face recognition and **84.48%** for expression recognition is obtained with Euclidean distance classifier while the training database has images of frontal, side, top views, and with spectacles. While the training database has only frontal faces the highest recognition accuracy of **68.96%** for face recognition and **63.79%** for expression recognition is obtained with Euclidean distance classifier.

Future Scope –

(1) The recognition accuracy can be improved by making the algorithm robust to illumination and pose variation with limited number of training faces.

(2) We can look forward for the hard ware implementation of this work in real time. One of the real time applications is “patient monitoring system”. The data base of patients in the hospital is maintained and the recognition of pain expression along with that particular patient’s face recognition makes it possible to attend and treat the patient immediately.

Bibliography

[1] S.Agrawal and P.Khatri, "Facial Expression Detection Techniques: Based on Viola and Jones Algorithm and Principal Component Analysis," *2015 Fifth International Conference on Advanced Computing & Communication Technologies*, Haryana, 2015, pp. 108-112.

[2] Meijing Li, Xiuming Yu, Keun Ho Ryu, Sanghyuk Lee, NiponTheera-Umpon, **“**Face recognition technology development with Gabor, PCA and SVM methodology under illumination normalisation condition**”** Cluster Comput (2017). doi:10.1007/s10586-017-0806-7

[3] D.Reney and N. Tripathi, "An Efficient Method to Face and Emotion Detection," *2015 Fifth International Conference on Communication Systems and NetworkTechnologies*, Gwalior, 2015,pp.493-497.doi: 10.1109/CSNT.2015.155

[4] G. Donato, M. S. Bartlett, J. C. Hager, P. Ekman, and T. J. Sejnowski, “Classifying facial actions,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 21, no. 10, pp. 974–989, Oct. 1999.

[5] L. Zhang, Student Member, IEEE, and D. Tjondronegoro,”Facial Expression Recognition Using Facial Movement Features’’ IEEE Transaction on Affective Computing, VOL. 2, NO. 4, October-December 2011.

[6] Kuldeep Singh Sodhi and Madan Lal."Comparative analysis of PCA-based face recognition system using different distance classifiers." Int. J. of Appl. or Innovation in Eng. &Manag 2 (2013): 341-348.

[7] Bartlett, M. S., Movellan, J. R., &Sejnowski, T. J. (2002). Face Recognition by Independent Component Analysis. *IEEE Transactions on Neural Networks / a Publication of the IEEE Neural Networks Council*, *13*(6), 1450–1464.

[8] Yu-Buhee and Sukhanlee" ETRI Journal, Volume 33,No.4, August 2011.

[9] Chaur-Chin Chen and Hsueh-Ting Chu, "Similarity measurement between images," *29th Annual International Computer Software and Applications Conference (COMPSAC'05)*, 2005, pp. 41-42 Vol. 1. doi: 10.1109/COMPSAC.2005.140

[10] Yow and Cipolla "Feature Based Human Face Detection" Image and Vision Computing vol15,No.9pp 713-735,1997

[11] Pushpaja V. Saudagare, D.S. Chaudhari, Facial Expression Recognition using Neural Network –An Overview, International Journal of Soft Computing and Engineering (IJSCE) ISSN: 2231-2307, Volume-2, Issue-1, March 2012

[12] Teik-Toe TEOH Siu-Yeung CHO "Human Emotional States Modeling by hidden Markov Model" 2011 Seventh International Conference on Natural Computation

[13] Hemalatha, G., & Sumathi, C. P. (2014). A study of techniques for facial detection and expression classification. *International Journal of Computer Science and Engineering Survey*, *5*(2), 27.

[14] Sandeep K. Gupta, ShubhLakshmi Agrawal, Yogesh K. Meena, Neeta Nain "A Hybrid Method of Feature Extraction for Facial Expression Recognition" 2011 Seventh International Conference on Signal Image Technology & Internet-Based Systems.

[15] Sanjay Kr.singh, D.S.Chauhan,Mayank Vatsa, Richa Singh "A robust Skin Color Based Face Detection Algorithm" Tamkang Journal of Science and Engineering, Vol 6, No4,pp227-234(2003).

[16] Jiaming Li, Geoff Poulton, Ying Guo,Rong-Yu Qiao "Face Recognition Based on Multiple Region Features" Proc.VIIth Digital Image Computing:Techniques and Applications,Sunc,TalbotH,OurselinS. and Adriaansen T.(Eds), 10-12 Dec 2003,Sydney.

[17] Paul Viola, Michael J Jones “Robust Real-Time Face Detection” International Journal of Computer Vision 57, pp. 137-154, Netherlands, 2004.

[18] https://en.wikipedia.org/wiki/Feature_extraction

[19] A. Aghagolzadeh S.Dabbaghchian and M.S. Moin. Feature extraction

using discrete cosine transform for face recognition. IEEE, 2007.

[20] https://en.wikipedia.org/wiki/Median_filter

[21] https://en.wikipedia.org/wiki/Gabor_filter

[22] https://cvtuts.wordpress.com/2014/04/27/gabor-filters-a-practical-overview/

[23]http://www.cs.otago.ac.nz/cosc453/student_tutorials/principal_components.pdf

[24] https://en.wikipedia.org/wiki/Eigenface

[25] http://www.serc.iisc.ernet.in/~venky/SE263/slides/EigenFaces.pdf

[26]http://ijarcet.org/wp-content/uploads/IJARCET-VOL-1-ISSUE-9-135-139.pdf

[27] https://en.wikipedia.org/wiki/MATLAB

[28] http://in.mathworks.com/matlabcentral/mlcdownloads/downloads/submissio-ns/16201/versions/3/previews/toolbox_image/html/content.html

[29] https://en.wikipedia.org/wiki/Homomorphic_filtering

[30] https://en.wikipedia.org/wiki/Multiplicative_noise

[31] https://in.mathworks.com/help/images/ref/imgaussfilt.html

[32] https://en.wikipedia.org/wiki/Wiener_filter

[33] http://edoras.sdsu.edu/doc/matlab/toolbox/images/deblurr6.html

[34] https://en.wikipedia.org/wiki/Histogram_equalization

[35] http://www.cambridgeincolour.com/tutorials/gamma-correction.htm

[36] https://www.cgsd.com/papers/gamma_intro.html

[37] http://in.mathworks.com/help/images/ref/imadjust.html

[38] https://en.wikipedia.org/wiki/Image_analysis

[39] http://in.mathworks.com/help/matlab/ref/rgb2gray.html

[40] https://en.wikipedia.org/wiki/Grayscale

[41] https://en.wikipedia.org/wiki/Digital_image_processing

[42] Digital Image Processing by Rafael C. Gonzalez and Richard Eugene Woods, 3rd Edition by Pearson

[43] Image Processing: The Fundamentals by Maria Petrou, 2nd Edition by Wiley

[44] http://www.imageprocessingbasics.com/rgb-to-grayscale/