Acknowledgement

This project **“Emotion Recognition with Face Detection and Face Recognition”** has been completed not only by the team members but also from joint efforts of many individuals. It has been very pleasure for us to acknowledge and contribution that were very important and supportive throughout the project.

We would like to thank to number of people who devoted their time and suggestion without which it would have been very difficult for us to complete this project.

We would like to extend our special thanks to our supervisor **Er. Saroj Bista** for providing technical support with suggestions which helped our project to achieve its goals. It would be our privilege to thank **Dr. Roshan Chitrakar** (Associate Professor, NCIT) and **Er. Madan Kadariya** (HOD of IT Engineering) for their support from every beginning of semester to completion of Project. Also we would not forgot to thank our teachers and classmates who have been knowingly and unknowingly the part of this project and lent support and views during the entire development time.

Abstract

The Project titled "**Emotion Recognition with Face Detection and Face Recognition**" is a system which recognize a person by extracting different facial features and then analyze and identify emotion of detected faces. Each detected face has its own analysis result which includes the confidence scores for seven different kinds of emotion: neutral, surprise, sadness, happiness, anger, fear, and disgust. The face is one of the easiest ways to distinguish the individual identity of each other. Face recognition is a personal identification system that uses personal characteristics of a person to identify the person's identity. Human face recognition procedure basically consists of two phases, namely face detection, where this process takes place very rapidly in humans, except under conditions where the object is located at a short distance away, the next is the introduction, which recognize a face as individuals. This stage is then replicated and developed as a special model for facial image recognition. Face recognition is one of the much-studied biometrics technology that has been developed by experts. There are two kinds of methods that are currently popular in developed face recognition pattern namely, Eigen face method and Edge histogram descriptor. Facial image recognition using Eigen face method is based on the reduction of face-dimensional space which requires Principal Component Analysis (PCA) for facial features. The main purpose of PCA on face recognition using Eigen faces was formed (face space) by finding the eigenvector corresponding to the largest eigenvalue of the face image. We prefer Edge histogram descriptor for face recognition which generated histogram using different edge types. The area of this project face detection system with face recognition is Image processing.

Face detection: identifies human faces in digital images.

Face recognition: recognizes human faces by matching different facial parameters.

Emotion recognition: Analyze and identify emotion of detected faces.

The software requirements for this project is MATLAB software.

Keywords: Face detection, Emotion recognition, Eigen face, PCA, EHD, MATLAB

Contents

List of Table

List of Figures

List of Abbreviation

PCA:	Principal Component Analysis
EHD:	Edge Histogram Descriptor
HMMD:	Hue-Max-Min-Difference
FACS:	Facial Action Coding System
HMM:	Hidden Markov Model
FERET:	Facial Recognition Technology
DWD:	Dark-White-Dark
RGB:	Red Green Blue
ANN:	Artificial Neural Network
DFD:	Data Flow Diagram
ERD:	Entity Relationship Diagram
SRS:	System Requirement Specifications
UI/UX:	User Interface / User Experience
CRC:	Class-Responsibility-Collaboration
MJPG:	Motion Joint Photographic Experts Group

1. Introduction

Face recognition is typically used in security systems. Besides that, it is also used in human computer interaction. In order to develop this project Eigen faces method is preferred for training and testing of different emotion faces. A general problem statement can be formulated as follows, given still picture or video images of a certain scene, the person detected on the scene will be identified using a pre stored database of facial images. The typical solution of this problem involves face detection, feature extraction from the face regions, face recognition and finally facial expression recognition. For the development of this project, for face recognition we prefer Edge histogram descriptor method and for emotion recognition we use Eigen faces method. Eigen faces are a set of eigenvectors used in the computer vision problem of human face as well as emotion recognition. A set of Eigen faces can be generated by performing a mathematical process called principal component analysis (PCA) on a large set of images containing different human faces. The key procedure in principal component analysis is based on Karhumen-Loeve transformation. If the elements of images are considered as random variables, then the image may be seen as a sample of a stochastic process. The main focus of this project is to find out the accuracy of the Edge histogram and Eigen faces method in face and facial expression recognition. Eigen faces approach seemed to be an adequate method to be used in facial expression recognition due to its simplicity, speed and learning capability [1]. The scheme is based on an information theory approach that decomposes face images into a small set of characteristic feature images called Eigen faces, which may be thought of as the principal components of the initial training set of face images [1].

1.1 Face Detection

Face detection is a computer technology that determines the location and size of human face in arbitrary (digital) image. The facial features are detected and any other objects like trees, buildings and bodies etc. are ignored from the digital image. To detect a face we will use the `vision.CascadeObjectDetector` that detects the location of a face in a video frame.

Following steps will be involved for detecting a facial component:

- Create a cascade detector object.
- Read a video frame and run the detector.
- Draw the returned bounding box around the detected face.

The cascade object detector uses the Viola-Jones detection algorithm and a trained classification model for detection. By default, the detector is configured to detect faces, but it can be configured for other object types. The cascade object detector uses the Viola-Jones algorithm to detect people's faces, noses, eyes, mouth etc.

1.2 Face Recognition

Edge histogram descriptor approach will be used for face recognition. Recognition is based on generating the histogram between facial landmarks, or in other words the spatial configuration of facial features. That means that the main geometrical features of the face such as the eyes, nose and mouth are first located and then faces are classified on the basis of various edge types found on the histogram of facial images.

1.3 Facial Expression Recognition

Facial Expression Recognition mainly involves two types of problems: one is the expression type and the other is the recognition method. As introduced previously, the seven types of expression were broadly accepted. While the recognition method contains expert rules, neural network, Principal Component Analysis (PCA) and hidden Markov model (HMM).

2 Problem Statement

Face recognition has become a popular area of research in computer vision, mainly due to increasing security demands and its potential, commercial and law enforcement applications. It is a very challenging problem and up to date there is no technique that provides a robust solution to all situations and different applications that face recognition may encounter. Hence, this dissertation focuses on developing a technique that provides a solution for an efficient high-speed face recognition system in different applications.

Face detection: During the detection process we may encounter with multiple unnecessary object detection in faces which may not be efficient way for capturing the facial features. So to remove such problems we need to minimize multiple unnecessary detection by varying the threshold value.

Face recognition: During the recognition process we may encounter a problem which may not effectively recognize who the person is, we store a neutral image of a person in datasets and the testing image may not be the exact position when it was stored. Blurry image may affect the proper recognition. It requires to implement neural network to predict the required face to whom it belong. Images with spectacles are the challenging issues in face recognition.

Emotion recognition: Emotion of face is the change in position of pixel values of a facial image. Change in pixel values within the specified range is acceptable. All required real valued parameters and binary parameters need to be matched to fully recognize a different types of emotions. It requires to implement neural network to identify which emotions does the detected face belongs to.

3 Objectives

The main objectives of this project are to: -

i. Detect a face of a person.

ii. Recognize a face of a person.

iii. Identify seven basic facial expressions: surprise, fear, disgust, neutral, anger, happy and sad.

4 Significance of Study

- **Face Identification:**

Face recognition systems identify people by their face images. Face recognition systems establish the presence of an authorized person rather than just checking whether a valid identification (ID) or key is being used or whether the user knows the secret personal identification numbers (Pins) or passwords.

- **Access Control:**

In many of the access control applications, such as office access or computer logon, the size of the group of people that need to be recognized is relatively small. The face pictures are also caught under natural conditions, such as frontal faces and indoor illumination. The face recognition system of this application can achieve high accuracy without much co-operation from user.

- **Image database investigations**:

Searching image databases of licensed drivers, benefit recipients, missing children, immigrants and police bookings.

- **General identity verification:**

Face recognition systems establish the presence of an authorized person rather than just checking whether a valid identification (ID) or key is being used or whether the user knows the secret personal identification numbers (Pins) or passwords.

- **Surveillance:**

Like security applications in public places, surveillance by face recognition systems has a low user satisfaction level, if not lower. Free lighting conditions, face orientations and other divisors all make the deployment of face recognition systems for large scale surveillance a challenging task.

5 Project Scope and Limitations

5.1 Scope

There are several scopes that need to be proposed for the project:

i. Face recognition of a person.
ii. Based on only a frontal view image.
iii. Using offline test image.
iv. Only single face on the image.
v. No restriction on ornamental wears.

5.2 Limitations

Several factors limit the effectiveness of facial-recognition technology:

- **Variations in lightening conditions**

Images that are captured in different environmental conditions may affect the accuracy of system. Image containing high brightness, contrast and saturation level may reduce the details of the image. System may not extract the important features from the images and may vary error in recognition process. Reconstruction of those details that are not clear in original image may be the challenging task in the image processing.

- **Image Quality**

Image quality affects how well facial-recognition algorithms work. The image quality of scanning video is quite low compared with that of a digital camera.

- **Image Size**

When a face-detection algorithm finds a face in an image or in a still from a video capture, the relative size of that face compared with the enrolled image size affects how well the face will be recognized. An already small image size, coupled with a target distant from the camera, means that the detected face is only 300 to 300 pixels on a side. Further, having to scan an image for varying face sizes is a processor-intensive activity. Most algorithms allow specification of a face-size range to help eliminate false positives on detection and speed up image processing.

- **Face Angle**

The relative angle of the target's face influences the recognition score profoundly. When a face is enrolled in the recognition system, usually multiple angles are used (profile, frontal and 45-degree are common). Anything less than a frontal view affects the algorithm's capability to generate a template for the face. The more direct the image (both enrolled and probe image) and the higher its resolution, the higher the score of any resulting matches.

- **Processing and Storage**

Although some images captured from digital camera consists more details and contains large number of pixels. So processing of such images may require inadequate processing time which may slow down the system. Storage requirement of such uncompressed image are relatively high. So that further compression techniques needs to be implemented which may be the challenging task.

6 Literature Review

6.1 Eigen Faces

Eigen face is one of the most thoroughly investigated approaches to face recognition. It is also known as Karhunen-Loève expansion, Eigen picture, eigenvector, and principal component. References [2, 3] used principal component analysis to efficiently represent pictures of faces. They argued that any face images could be approximately reconstructed by a small collection of weights for each face and a standard face picture (Eigen picture). The weights describing each face are obtained by projecting the face image onto the Eigen picture. Reference [4] used Eigen faces, which was motivated by the technique of Kirby and Sirovich, for face detection and identification. In mathematical terms, Eigen faces are the principal components of the distribution of faces, or the eigenvectors of the covariance matrix of the set of face images. The eigenvectors are ordered to represent different amounts of the variation, respectively, among the faces. Each face can be represented exactly by a linear combination of the Eigen faces. It can also be approximated using only the "best" eigenvectors with the largest eigenvalues. The best M Eigen faces construct an M dimensional space, i.e., the "face space". The authors reported 96 percent, 85 percent, and 64 percent correct classifications averaged over lighting, orientation, and size variations, respectively. Their database contained 2,500 images of 16 individuals. As the images include a large quantity of background area, the above results are influenced by background. The authors explained the robust performance of the system under different lighting conditions by significant correlation between images with changes in illumination. However, [5] showed that the correlation between images of the whole faces is not efficient for satisfactory recognition performance. Illumination normalization [4] is usually necessary for the Eigen faces approach. Reference [6] proposed a new method to compute the covariance matrix using three images each was taken in different lighting conditions to account for arbitrary illumination effects, if the object is Lumberton. Reference [7] extended their early work on Eigen face to eigenfeatures corresponding to face components, such as eyes, nose, and mouth. They used a modular Eigen space which was composed of the above eigenfeatures (i.e., Eigen eyes, Eigen nose, and Eigen mouth). This method would be less sensitive to appearance changes than the standard Eigen face method. The system achieved a recognition rate of 95 percent on the FERET database of 7,562 images of approximately 3,000 individuals. In summary, Eigen face appears as a fast, simple, and practical method. However, in general, it does not provide invariance over changes in scale and lighting conditions. Recently, in [8] experiments with ear and face recognition, using the standard principal component analysis approach , showed that the recognition performance is essentially identical using ear images or face images and combining the two for multimodal recognition results in a statistically significant performance improvement. For example, the difference in the rank-one recognition rate for the day variation experiment using the 197-image training sets is 90.9% for the multimodal biometric versus 71.6% for the ear and 70.5% for the face. [7] Extended their early work on Eigen face to eigenfeatures corresponding to face components, such as eyes, nose, and mouth. They used a modular Eigen space which was composed of the above eigenfeatures (i.e., Eigen eyes, Eigen nose, and Eigen mouth). This method would be less sensitive to appearance changes than the standard Eigen face method. The system achieved a recognition rate of 95 percent on the FERET database of 7,562 images of approximately 3,000 individuals. In summary, Eigen face appears as a fast, simple, and practical method. However, in general, it does not provide invariance over changes in scale and lighting conditions. Recently, in [8] experiments with ear and face recognition, Using the standard principal component analysis approach, showed that the recognition performance is essentially identical using ear images or face images and combining the two for multimodal recognition results in a statistically significant performance improvement. For example, the difference in the rank-one recognition rate for the day variation experiment using the 197-image training sets is 90.9% for the multimodal biometric versus 71.6% for the ear and 70.5% for the face. There is substantial related work in multimodal biometrics. For example [9] used face and fingerprint in multimodal biometric identification, and [10] used face and voice. However, use of the face and ear in combination seems more relevant to surveillance applications.

6.2 Geometrical Feature Matching

Geometrical feature matching techniques are based on the computation of a set of geometrical features from the picture of a face. The fact that face recognition is possible even at coarse resolution as low as 8x6 pixels [11] when the single facial features are hardly revealed in detail, implies that the overall geometrical configuration of the face features is sufficient for recognition. The overall configuration can be described by a vector representing the position and size of the main facial features, such as eyes and eyebrows, nose, mouth, and the shape of face outline. One of the pioneering works on automated face recognition by using geometrical features was done by [12] in 1973. Their system achieved a peak performance of 75% recognition rate on a database of 20 people using two images per person, one as the model and the other as the test image. References [13, 14] showed that a face recognition program provided with features extracted manually could perform recognition apparently with satisfactory results. Reference [15] automatically extracted a set of geometrical features from the picture of a face, such as nose width and length, mouth position, and chin shape. There were 35 features extracted form a 35 dimensional vector. The recognition was then performed with a Bayes classifier. They reported a recognition rate of 90% on a database of 47 people. Reference [16] introduced a mixture-distance technique which achieved 95% recognition rate on a query database of 685 individuals. Each face was represented by 30 manually extracted distances. Reference [17] used Gabor wavelet decomposition to detect feature points for each face image which greatly reduced the storage requirement for the database. Typically, 35-45 feature points per face were generated. The matching process utilized the information presented in a topological graphic representation of the feature points. After compensating for different centroid location, two cost values, the topological cost, and similarity cost, were evaluated. The recognition accuracy in terms of the best match to the right person was 86% and 94% of the correct person's faces was in the top three candidate matches. This project work adopts the following methodologies for the application of knowledge, skills, tools and techniques to a broad range of activities to meet the requirements of our project **'Emotion Recognition with Face Detection and Face Recognition.'**

7 Methodology

7.1 Face Detection

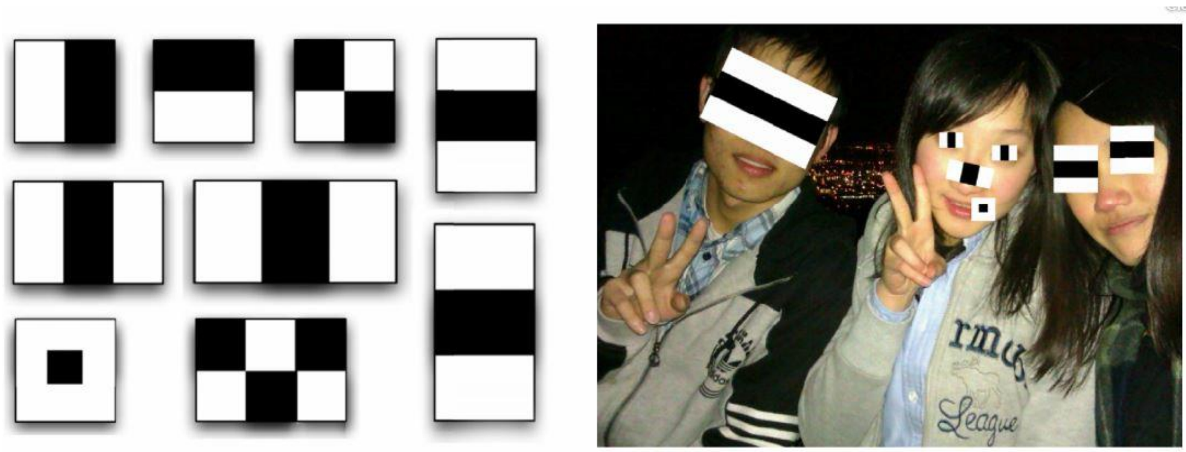

Figure 1: Haar like Features

Haar like features uses three different types of rectangular features i.e. Edge, line and combinations of four rectangles in order to detect a single object. A Haar object detector defines thousands of these rectangular features in different region and combines these features in order to define an object and detect it efficiently.

7.1.1 Eye Detection

Eyes are detected based on the hypothesis that they are darker than other part of the face, finding eye analogue segments searching small patches in the input image that are roughly as large as an eye and are darker than their neighborhoods.

7.1.2 Nose Detection

- **Similarity of both sides:** The left and right sides of nose are similar in a front-view face, this property of similarity can be measured using Euclidean distance between both sides.

- **Dark-White-Dark (DWD) property:** The lower part of nose region is characterized by two dark nostrils and a light sub region due to the reflection of light on the nose.

7.1.3 Mouth Detection

Detection and extraction features from the mouth that is composed of weak classifiers, based on a decision stump, which uses Haar features to encode mouth details. The basic concept of the proposed algorithm is to extract and then verify the desired components, including skins, lips, by applying some kind of color segmentation in addition to some geometry. Based on the color distribution of the lips and skins. The lip detection method based on RGB chromaticity diagram is used to separate lip color from other colors (face skin color) by a color segmentation method.

7.2 Facial Expression Recognition

The proposed system contains four stages: face detection, pre-processing, principle component analysis (PCA) and classification as shown in Figure.

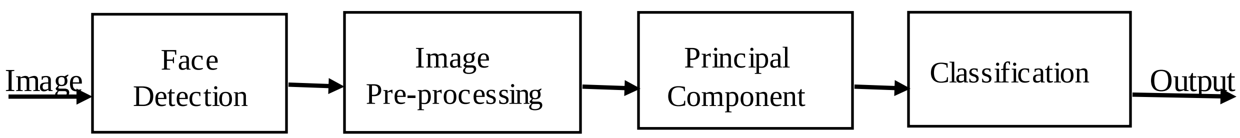

Figure 2: Facial Expression Recognition System

The first stage is face detection method. In this method the database of images are almost identical environment of distance, background, etc. the collection of all the images includes different poses of several neutral, anger, happiness, etc. expressions. For creating any type of database some images used for training and some for testing, both of which include number of expressions. The proposed technique is depend on coding and decoding method. First the information is extracted, encoded and then matched with the database of model. Next is the pre-processing module, in this the image gets normalized and it also remove the noise from the image. In Eigen face library the database image set divides into two sets- training dataset and testing dataset. The Eigen faces are calculated from the training set. These training set images are matched with the best Eigen faces, which have the largest Eigen values. For calculating those Eigen values the principle component analysis algorithm (PCA) is used. At the last stage of architecture the neural network trained the function in various field of application. The Artificial

Neural Network (ANN) can be used for the database in which the face descriptors are used as an input to train the network. For all positive result the network shows 1 in output and for all negative result 0 is present in output. If the new database is obtained for training then first the neural network match all the new result to the pre-built dataset and match the maximum threshold values and provide the output. Then it is conformed that the new facial expression is belong to the recognized person with the maximum output.

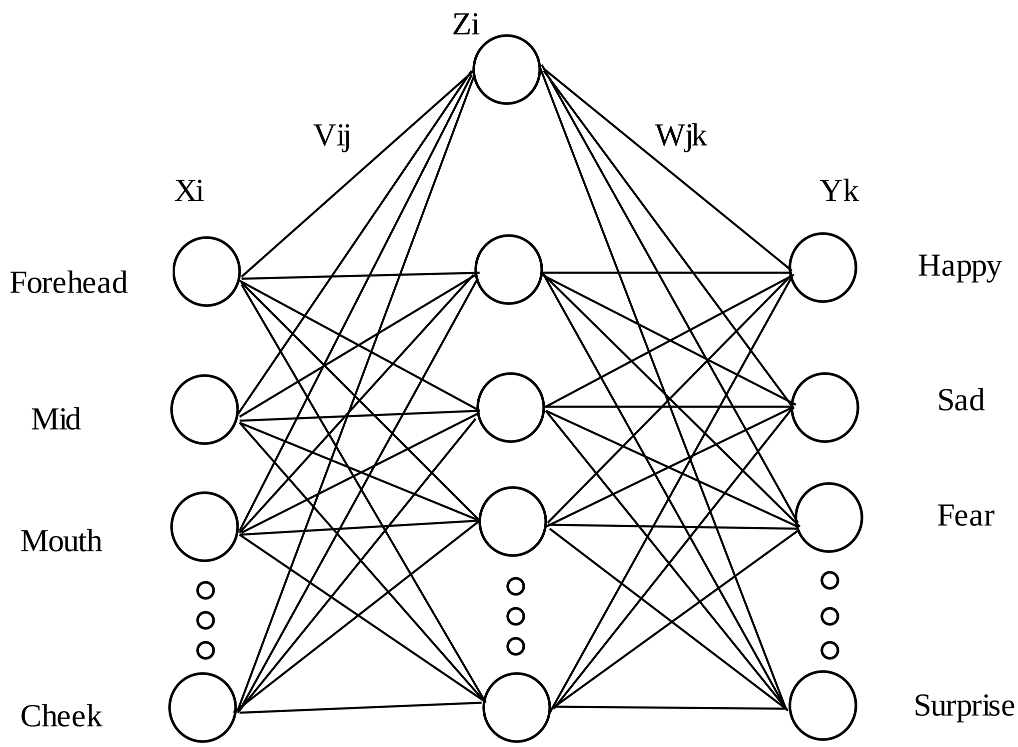

Figure 3: Architecture of Feed-Forward Back Propagation Neural Network

The figure gives an example of feed-forward backpropagation neural network, in which the input layer composed of neurons. These neurons provide the data of forehead, mid forehead, and mouth to the next layer of neuron. The next layer is called a hidden layer which calculate the values and provided to the output layer, where the system provides the different expression as an output.

Layers used in generating Neural Network:

Input layer size = 15;

Hidden layer size = 100;

Output layer size = 7;

7.3 Feature Extraction: Emotion Recognition

Principal Components Analysis PCA is preferred for feature extraction which is powerful method in image formation, Data Patterns, similarities and differences between them can be identified efficiently. The common features that lies in human face like, nose, mouth, eyes, lips etc. in which principal component analysis is used to extract these features from the human face which are the principal components. The main advantages of PCA is dimension will be reduced by avoiding redundant information without much loss. Two types of parameters can be extracted from the facial images.

- **Real valued parameters.**
- **Binary parameters.**

The real valued parameters will have a definite value depending upon the distance measured. This definite value can be measured in number of pixels. The binary measures give either a present (= 1) or an absent (= 0) value. In all, eight real valued measures and seven binary measures need to be obtained.

A number of parameters, both real-valued and binary, needs to be extracted and analyzed to decide their effectiveness in identifying a certain facial expression. The features which do not provide any effective information of the facial expression portrayed in the image needs to be eliminated and will not be considered in the final study. The real valued and binary feature selection is inspired by the Facial Action Coding System (FACS). The following real valued and binary parameters will be finally used in this system.

Real valued parameters used:

1. Eyebrow raise distance, 2. Upper eyelid to eyebrow distance, 3. Inter eyebrow distance, 4. Upper eyelid to lower eyelid distance, 5. Top lip thickness, 6. Lower lip thickness, 7. Mouth width, 8. Mouth opening.

Binary parameters used:

1. Upper teeth visible, 2. Lower teeth visible, 3. Forehead lines, 4. Eyebrow lines, 5. Nose lines, 6. Chin lines, 7. Nasolabial lines.

7.4 Features Extraction: Face Recognition:

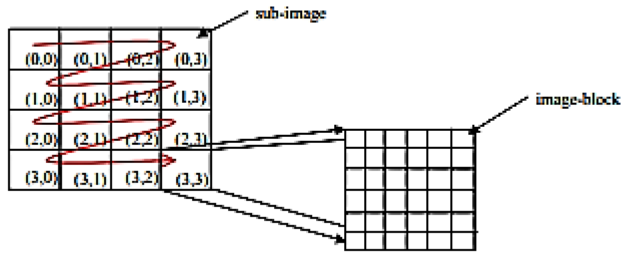

Figure 4: Sub-Image Division

We use Edge Histogram Descriptor (EHD) for the facial features extraction. Histogram is the most commonly used characteristic to represent the global feature composition of an image. Edge in the image is considered an important feature to represent the content of the image. To localize edge distribution to a certain area of the image, we divide the image space into 4x4 sub-images. Then, for each sub-image, we generate an edge histogram to represent edge distribution in the sub-image. To define different edge types, the sub-image is further divided into small square blocks called image-blocks. The size and the number of image-blocks in each sub-image.

- **HMMD (Hue-Max-Min-Difference):**

Captured image from webcam is in RGB form. RGB color space is converted into HMMD (Hue-Max-Min-Difference) and therefore is quantized. The component names, "Max", "Min" and "Diff" are according to the following transform equations between RGB and HMMD:

Max = max(R, G, B);

Min = min(R, G, B);

Diff = Max – Min;

Even though the four components are identified in the name of the HMMD color space, one more component, Sum can be defined.

Sum = (Max+Min) /2;

Therefore, a total of five components are identified in this color space. However, a set of three components, {H, Max, Min} or {H, Diff, Sum}, is sufficient to form the HMMD color space and specify a color point.

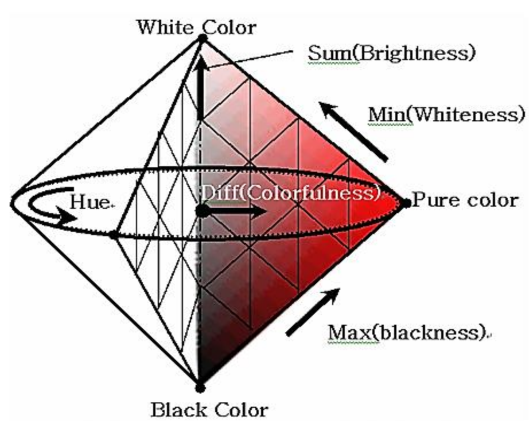

Figure 5: HMMD Color Space

- **Edge Type:**

Five edge types are defined in the edge histogram descriptor. They are four directional edges and a non-directional edge. Four directional edges include vertical, horizontal, 45 degree, and 135 degree diagonal edges. These directional edges are extracted from the image-blocks. If the image-block contains an arbitrary edge without any directionality, then it is classified as a non-directional edge.

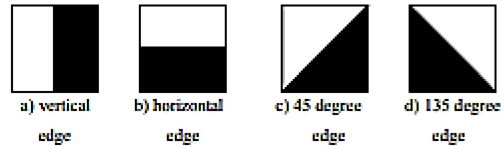

Figure 6: Edge Type

7.5 Backpropagation as a Classifier: Emotion Recognition

Backpropagation is a practical realization of the gradient descent algorithm in multilayered neural networks. Applying the Backpropagation algorithm to a neural network is a two-way process: we first propagate the input values through the networks and calculates errors, and then we back propagate the errors through the network backwards to adjust the connection weights to minimize the error. The algorithm calculates the gradient of the loss function with respect to the weights between the hidden layer and output layer nodes, and then it proceeds to calculate the gradient of the loss function with respect to the weights between the input layer and hidden layer nodes. After calculating the gradients, it subtracts them from the corresponding weight vectors to get the new weights for the connections. This process is repeated until the network produces the desired outputs.

7.6 Matching Similarities: Face Recognition

After the features has been extracted i.e. histogram with different edges has been stored as a facial features in the datasets. As for the recognition part captured image which has been divided into 4x4 sub-image, histogram is generated for each sub-image so that it can be easily identified how the edges are distributed in the sub-image. Shape of the edges is required to be identified, which is done by dividing the sub-image into small square blocks called image blocks. For identification of size of image block, following formula is used:

$$x = \sqrt{\frac{image\_width \times image\_height}{desired\_num\_block}}$$

EHD assumes that size of image blocks must be multiple of 2.

Different edge types vertical, horizontal, diagonal, anti-diagonal are extracted from those image blocks as a histogram. These histogram of different edges extracted from the captured image as a training images are stored in the datasets as a features for image recognition, which are then compared with the test images for face recognition. Histogram of different edges of trained images are compared with histogram of test images and similarities are matched. As a result face is recognized.

7.7 Software Development Life Cycle

The proposed framework for developing this project is Incremental model and Extreme Programming. The incremental model combines linear sequential model with the iterative prototype model. New functionalities will be added as each increment is developed. The phases of the linear sequential model are: Analysis, Design, Coding and Testing. The software repeatedly passes through these phase in iteration and an increment is delivered with progressive changes.

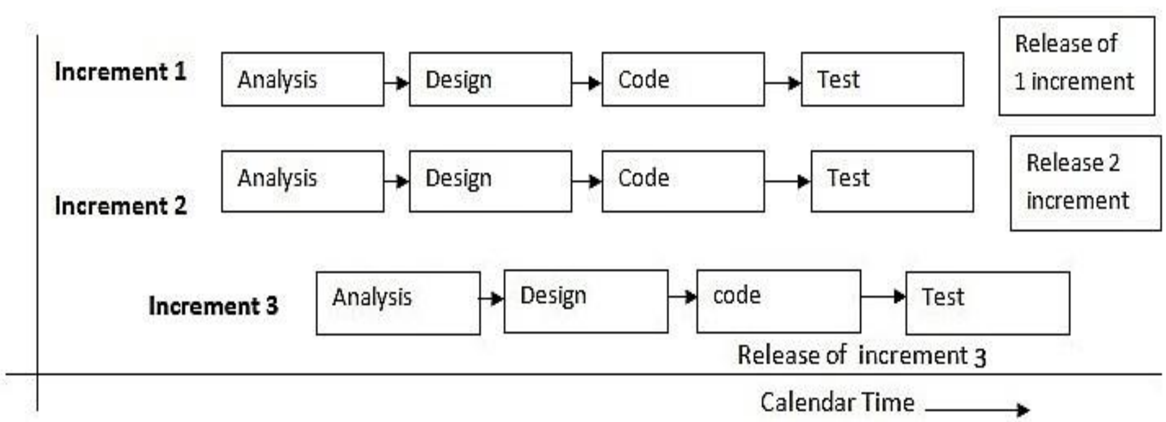

Figure 7: Incremental Model

7.7.1 Analysis Phase

In this phase, analysis will be performed in order to find out the requirements of the system. The outcome of this phase would be a SRS which is an acronym for "System Requirement Specifications".

7.7.2 Design Phase

In this phase, the SRS would be translated into the system's design. Context Diagram, DFD, ER – Diagram, Use Case Diagram and Class Diagram will be developed.

7.7.3 Coding Phase

In this phase, coding will be done according to the design and a working system will be developed by the end of this process.

7.7.4 Testing Phase

In this phase, the system will be tested. With each testing, a list of changes to the system developed, is suggested and the changes will be applied to the software and the software would be delivered as a successive increment until a satisfying system is achieved.

The tools used for documentation, designing and developing UI/UX, testing is listed below in table 1.

7.8 Extreme Programming

Figure 8: Extreme Programming

The Extreme Programming software development process is an agile method focused on providing the highest value for the customer in the fastest way possible.

7.8.1 Design

The first step in extreme programming software developing process is design. The three major guiding techniques adopted during the design phase are considerations of simplicity, use of CRC cards, and spike solutions.

7.8.2 Iteration Meetings

Extreme Programming progresses through iterations. The iterations in Extreme Programming methods ensure development of objects and classes that, when combined, provide the complete product.

7.8.3 Coding

Extreme Programming software development process gives priority to the actual coding over all other tasks.

Standards related to coding include:

- Adherence to metaphor or standards on names, class names, and methods.
- Using uniform styles and formats to ensure compatibility among the work of different team members.
- Pair programming or developing code by two programmers working together on a single machine to produce higher quality code at the same or less cost.

7.8.4 Acceptance Tests

Acceptance testing is of key importance in Extreme Programming. The customer provides the developer with functional or acceptance tests to validate the features. The developers subject the developed code to this test and continue to modify the code until the code passes the test. The test runs several times for confirmation.

7.8.5 Feedback

Extreme Programming adopts a system of continual feedback to make changes at the development change itself rather than at the end of the project. During the project development process, developers provide customers with acceptance test scores and demonstrations and take feedback. Such feedback might entail change in requirements, and the code is redone at this stage itself. A feedback system also exists between developers and project managers through daily meetings.

7.8.6 Completion

At the end of the iteration, the programmers deliver a working system without bugs to the customer. The developed code passing the specified acceptance test and the customer having no additional requirements from the user story developed at the start of the iteration ends the specific iteration.

7.9 Tools and Techniques Used

Tools	**Usages**
Photoshop	Image Editing
Integrated Webcam	Capture image
MATLAB	Coding, Database, UI

Table 1: Tools Used

7.10 System Design

There are a number of diagrams during system design which are listed below:

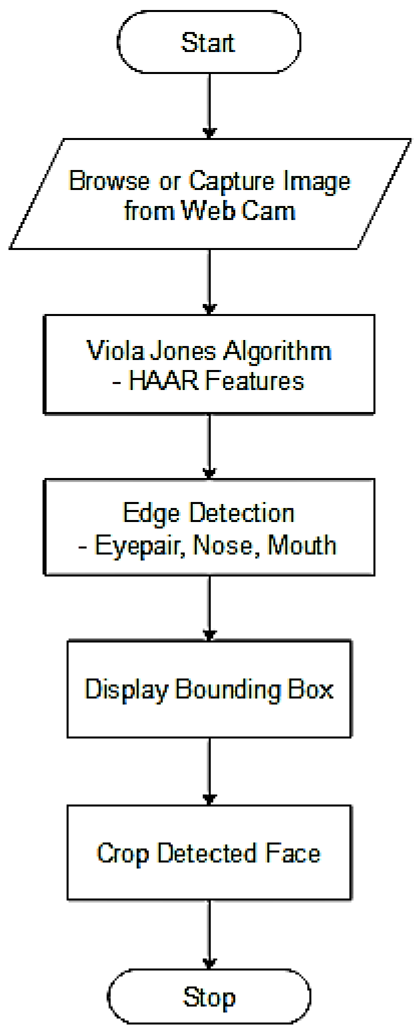

Figure 9: Flowchart of Face Detection

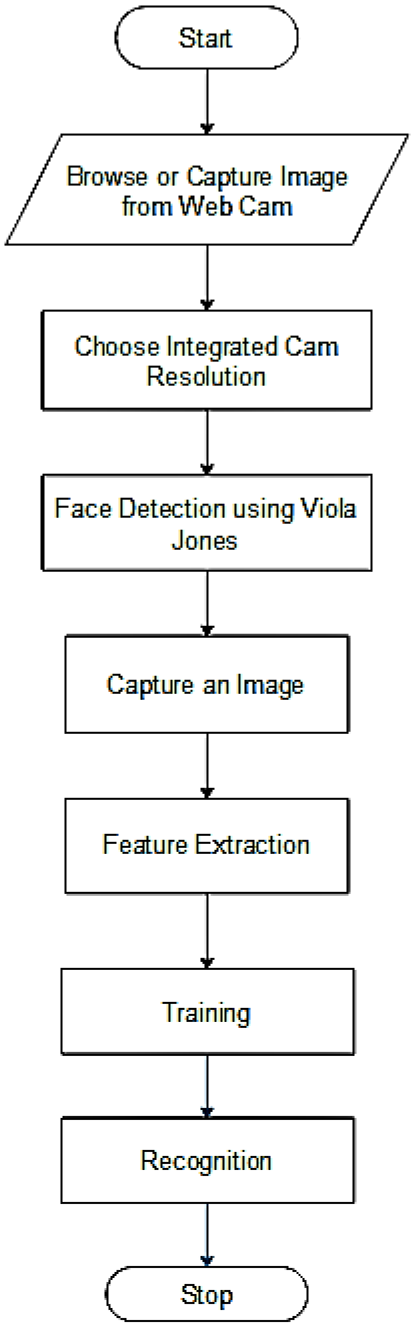

Figure 10: Flowchart of Face Recognition

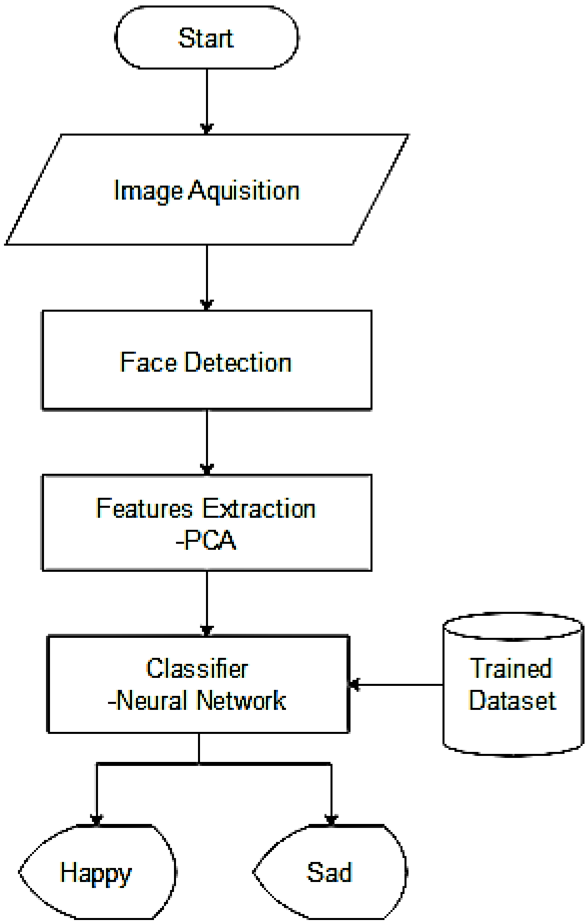

Figure 11: Flowchart of Emotion Recognition

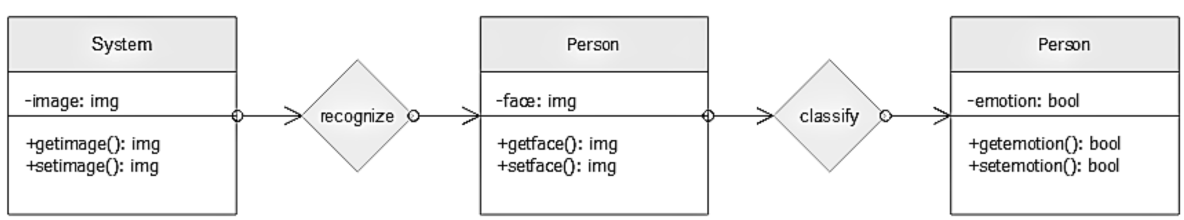

Figure 12: Class Diagram of Face and Emotion Recognition

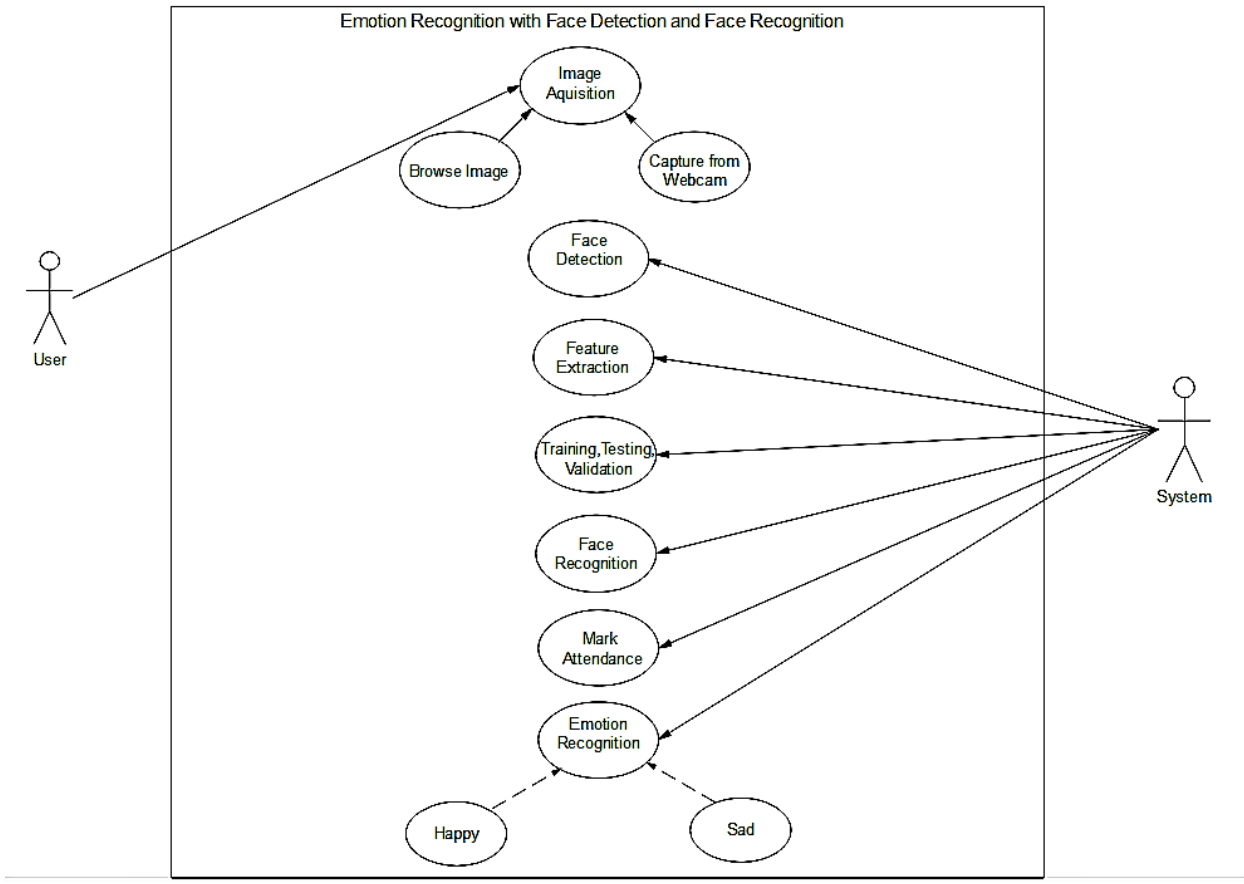

Figure 13: Use case Diagram of Face and Emotion Recognition

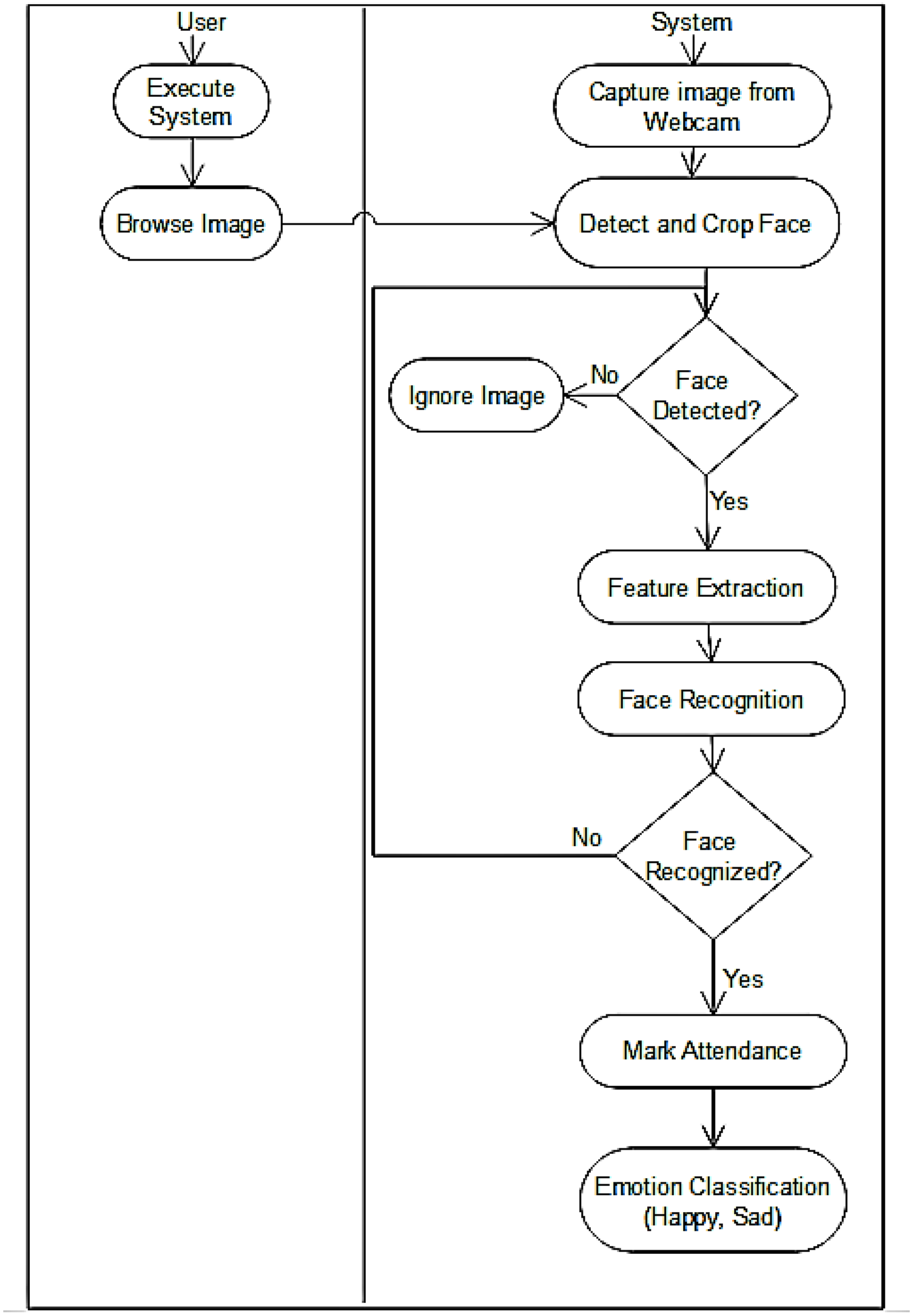

Figure 14: Activity Diagram of Face and Emotion Recognition

8 Proposed Outcome

Face Recognition process requires facial features. Every time proposed system tries to recognize a person it requires extracted facial features. For the purpose of features extraction and database creation it requires to capture the image of a person.

8.1 Opening Integrated Webcam

The proposed system requires image acquisition toolbox which is freely available in Matlab environment. It requires to access the integrated webcam of a machine with an appropriate resolution listed. MJPG 1280x720 or YUY2 1280x720 resolution is chosen for high quality image and that helps to extract a lossless features from facial images. Alternative way is to browse an image from local drive where no opening of webcam is required for features extraction.

Figure 15: Opening Integrated Webcam

8.2 Face Detection

Viola Jones algorithm has been implemented for detection of facial edges. It uses Haar like features for accurate detection of facial components. Matlab special features Bounding box is used for detection of face boundary. Only single person is accepted for face detection. Multiple faces are ignored and cannot be used for further process of recognition although which provides the advantage for the proposed system for not

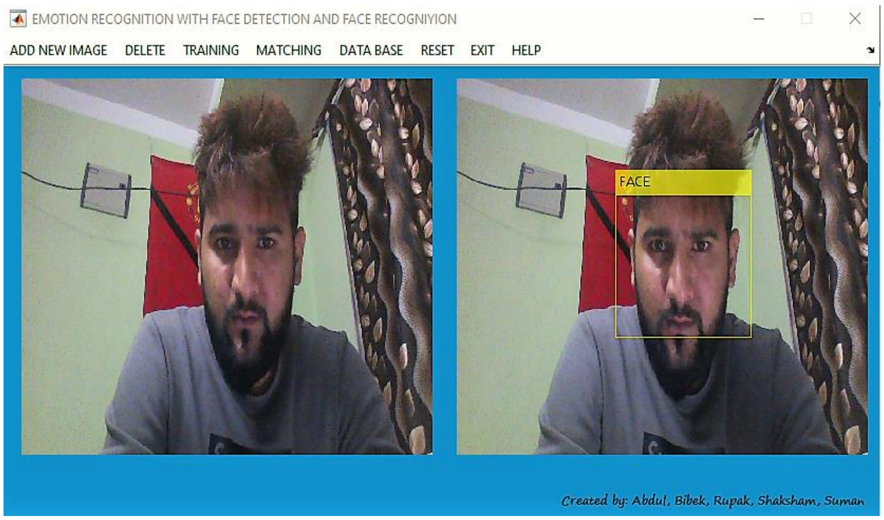

Figure 16: Face Detection

allowing more than one person to mark the attendance at a time. As for successful detection yellow rectangle box is shown in above image. From the above captured image, image in left side is moving image of integrated webcam and image in right side is successfully face detected image.

8.3 Cropping of an Image

After successful detection of face the proposed system crops the detected face area for further feature extraction process. From the above captured image, image in the left side is cropped image sized 300x300 pixels which is further stored in database for training purpose as well as for features extraction using Edge Histogram Descriptor (EHD). Image in right side is original image captured from webcam with appropriate face detection.

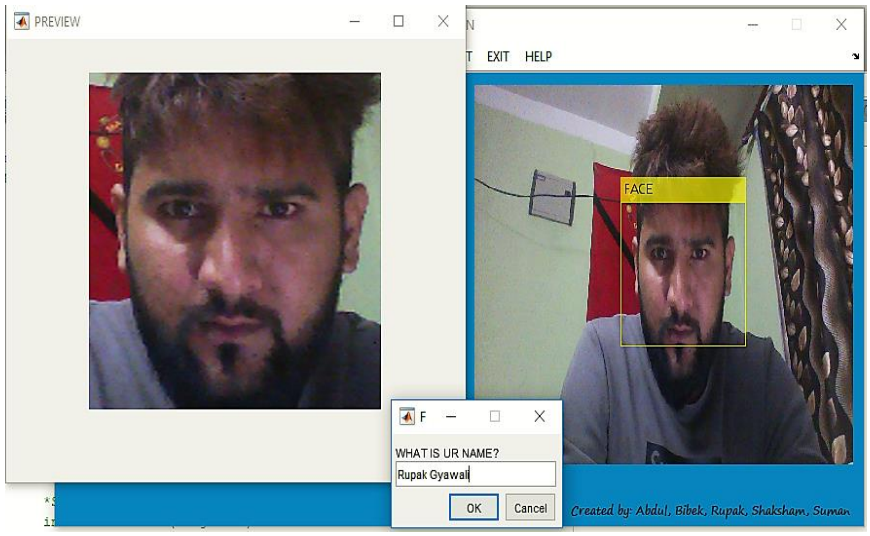

Figure 17: Cropping an Image

8.4 Features Extraction: Face Recognition

After capturing an image next step is to extract the facial features using Edge Histogram Descriptor. EHD divides original image into 4x4 sub-images and for each sub- images we generate an edge histogram to represent edge distribution. To define different edge types, the sub-image is further divided into small square blocks called image-blocks. Four directional edges namely vertical, horizontal, diagonal and anti-diagonal are extracted from these image-blocks. Now we can extract edge information from each block. These edge information are the main features that are extracted and stored in database.

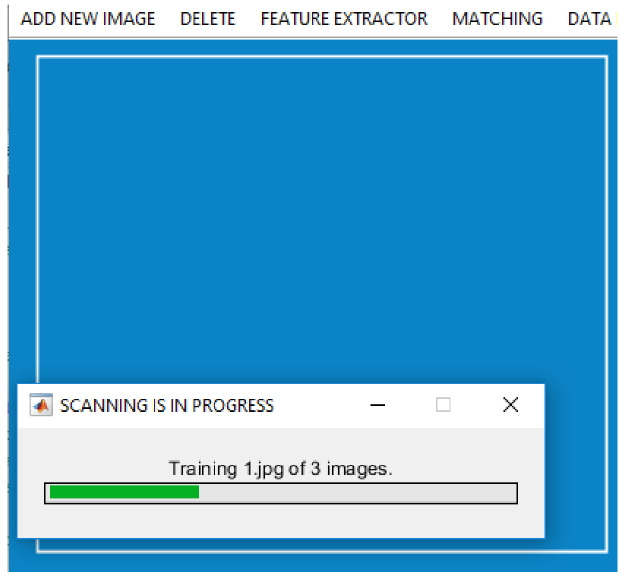

Figure 18: Feature Extraction

8.5 Database Creation

Detected face is cropped and is stored as in the form of image with extension *.jpg inside the database folder. Every captured image with accurately detected face contains a special facial features that varies between different people. Proposed system generates a database of features based on the images in the database folder. And the captured image is prompted with a dialogue box saying, enter name of a person. Image with name is tagged with the captured image and stored in database folder. Database is created to match the test image with pre-stored database image for recognition process. Figure above shows images stored in database.

Figure 19: Datasets

8.6 Testing of an Image

After the successful training, next step is to test for an unrecognized image. So the error rate can be computed and know how accurate the proposed system is. Extracted features of test image is compared with trained image and the recognition process is completed. And the result is either recognized or not recognized. In real world implementation once the image authorized person is stored in the database, for a daily attendance system test image refers to the image captured by webcam who is in front for marking an attendance. Hence test image is compared with trained images.

8.7 Face Recognition

Features are extracted using Edge Histogram Descriptor (EHD). Distances are calculated from different edge types and are stored in datasets as trained images. Again distances are calculated for test image and for recognition extracted edge distances are compared, if similarities found considered as recognized face else unrecognized face. From the figure, image in the left side is test image which is used for face recognition process whereas image on the right is pre stored dataset image.

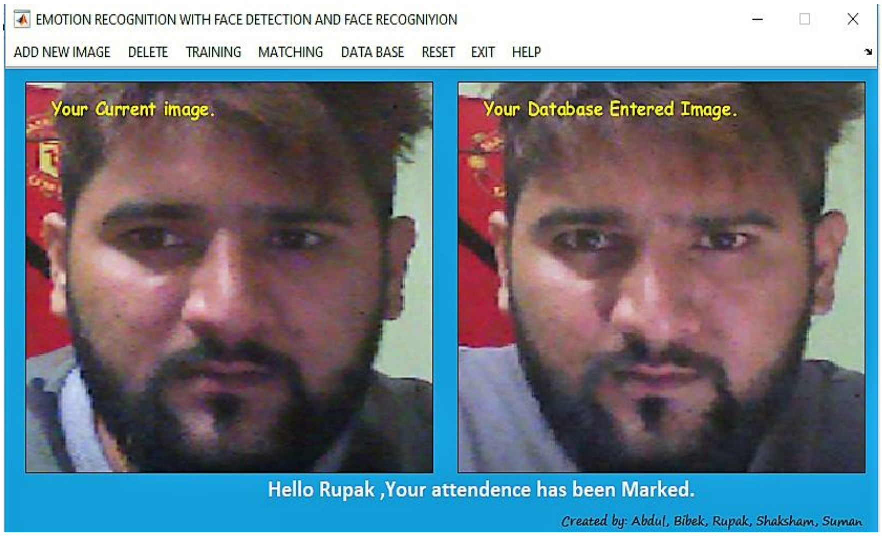

Figure 20: Face Recognition

8.8 Attendance Record

After the successful recognition of face the next step of proposed system is to mark an attendance. Name, Date, Time and the attendance status are recorded in text file.

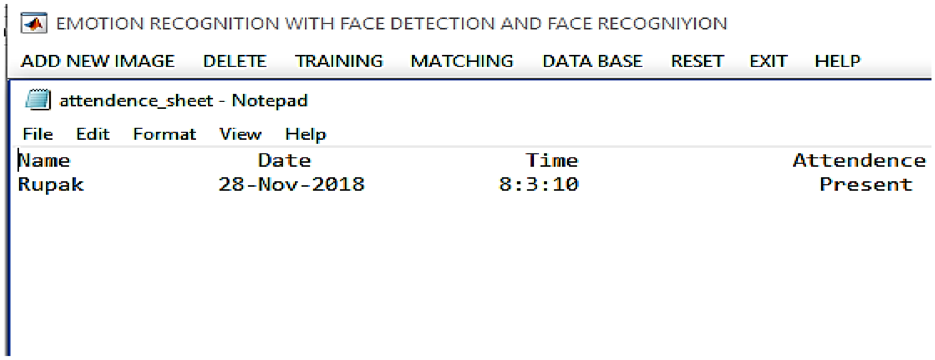

Figure 21: Attendance Record

8.9 Emotion Recognition

Facial Action Coding System (FACS) has provided seven different facial expression including Happy, Sad, Angry, Disgust, Fear, Neutral, Surprised. PCA has successfully extracted facial features and features of different emotions are stored in datasets and these features are trained using Neural Network. Following are the 7 different outcomes:

8.9.1 Emotion Recognition: Happiness

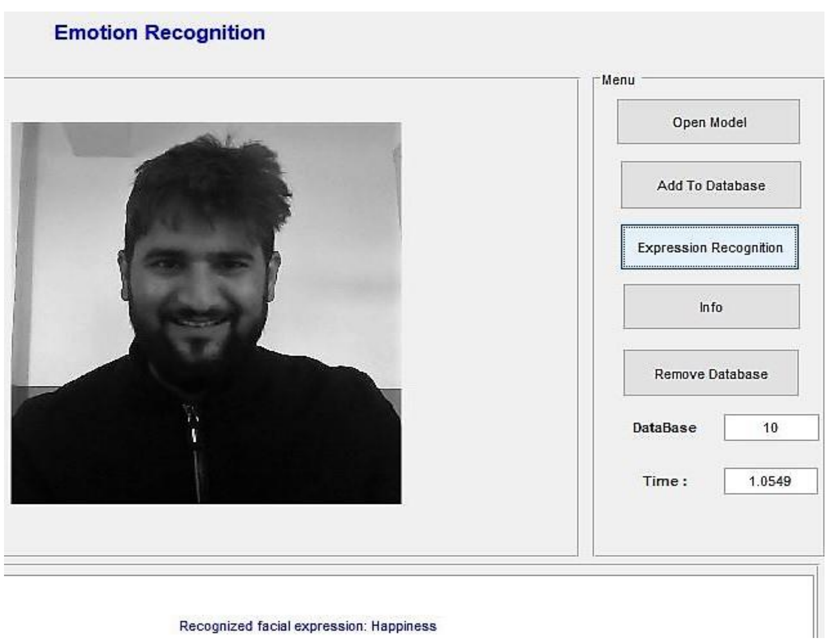

Figure 22: Emotion Recognition 'Happiness'

8.9.2 Emotion Recognition: Sadness

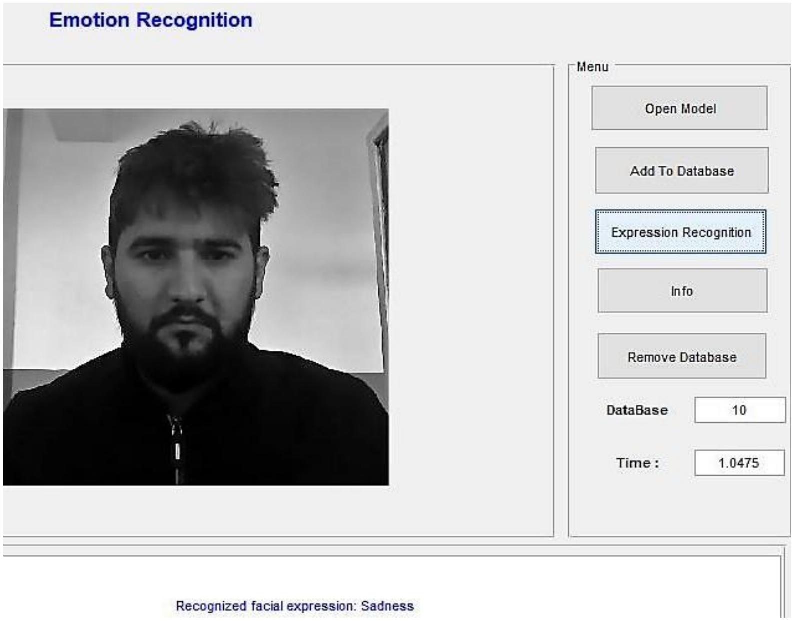

Figure 23: Emotion Recognition 'Sadness'

8.9.3 Emotion Recognition: Anger

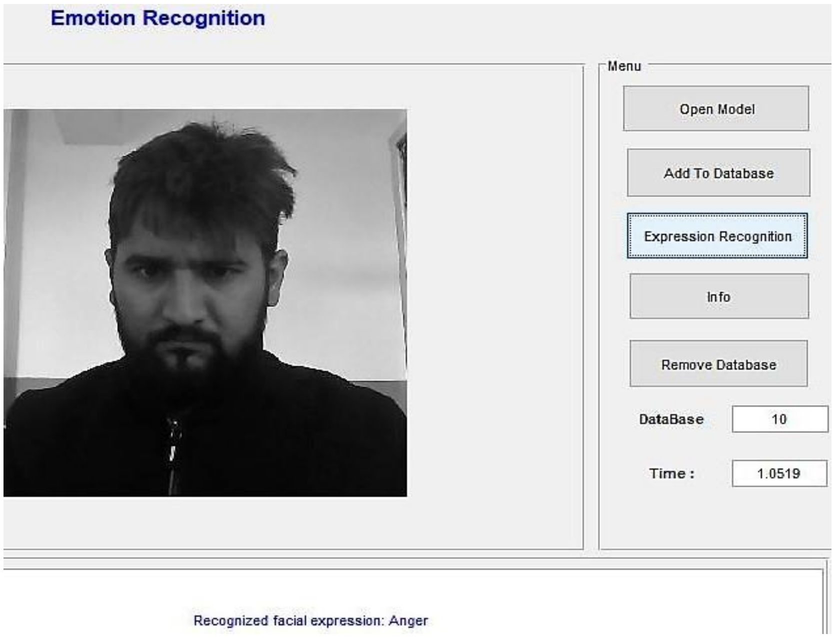

Figure 24: Emotion Recognition 'Anger'

8.9.4 Emotion Recognition: Neutral

Figure 25: Emotion Recognition 'Neutral'

8.9.5 Emotion Recognition: Surprise

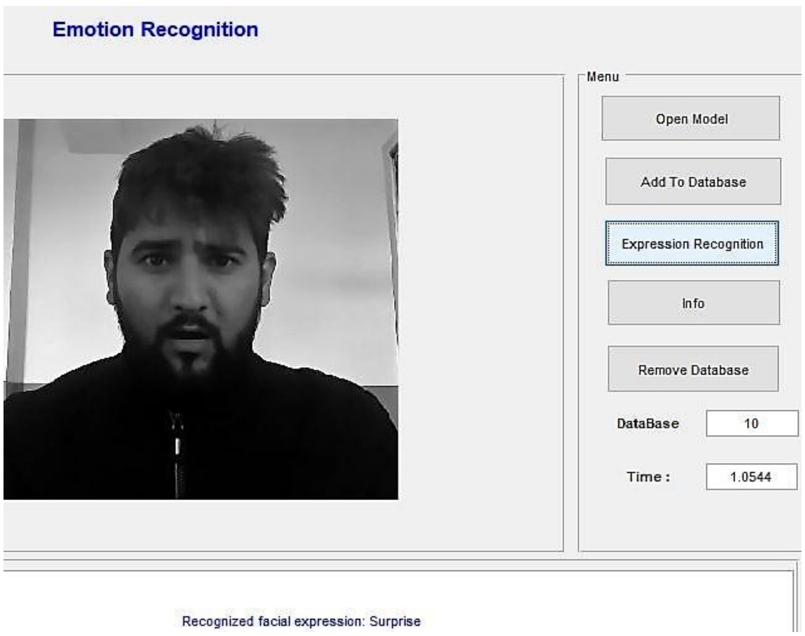

Figure 26: Emotion Recognition 'Surprise'

8.9.6 Emotion Recognition: Fear

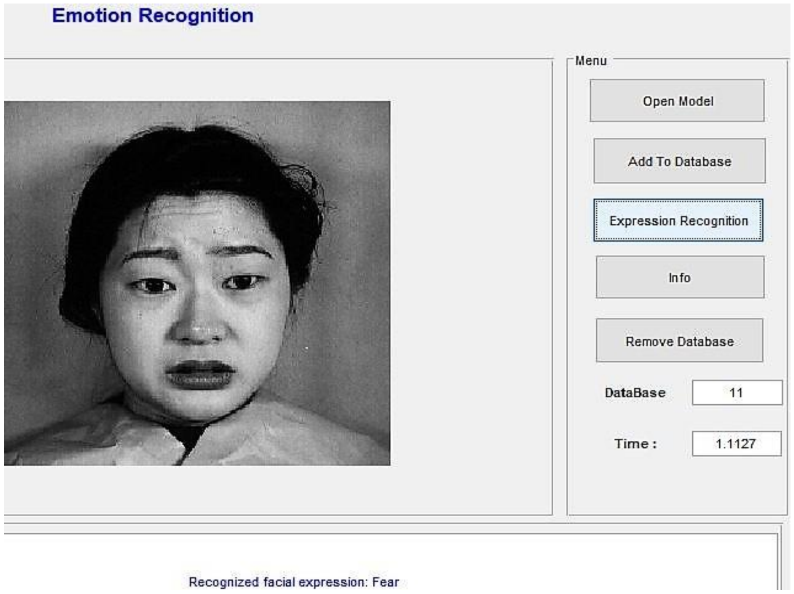

Figure 27: Emotion Recognition 'Fear'

8.9.7 Emotion Recognition: Disgust

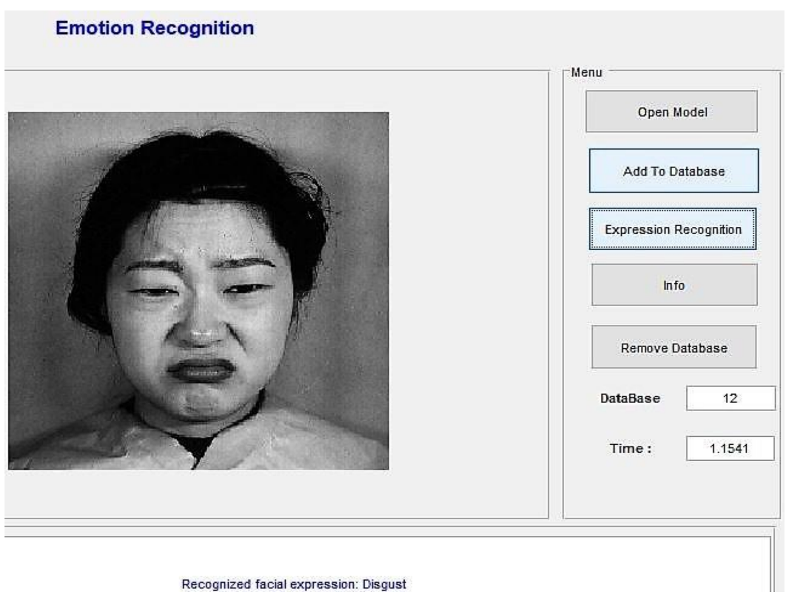

Figure 28: Emotion Recognition 'Disgust'

9 Testing and Results

Expression	Anger	Fear	Surprise	Sadness	Happiness	Disgust	Neutral
Anger	46	0	0	4	0	0	0
Fear	0	50	0	0	0	0	0
Surprise	0	0	50	0	0	0	0
Sadness	0	0	0	95	0	5	0
Happiness	0	0	0	0	100	0	0
Disgust	0	0	0	3	0	47	0
Neutral	0	0	0	0	0	0	50

Table 2: Test Images Result

Facial Expression	Total Number of Test	Successful Test	Failure Test
Anger	50	46	4
Fear	50	50	0
Surprise	50	50	0
Sadness	100	95	5
Happiness	100	100	0
Disgust	50	47	3
Neutral	50	50	0

Table 3: Experiment Result

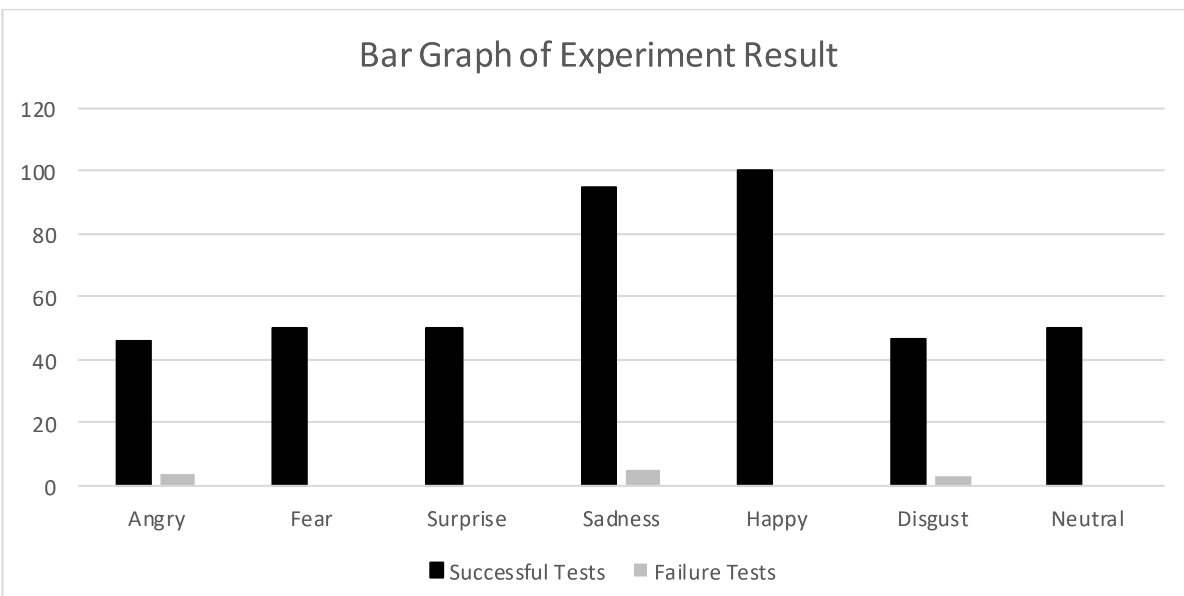

Figure 29: Bar Graph of Experiment Result

10 Conclusion

We propose an emotion recognition based on Principal Component Analysis (PCA) and face recognition based on Edge Histogram Descriptor (EHD). The proposed system will utilize the Eigen face method that is information reduction for the images for emotion recognition and computes edges information for face recognition. Viola jones algorithm is preferred for face detection where Haar features can be used for detection of facial features and Adaboost removes the redundancy occurred in facial features. If facial features (eye pair, nose, mouth) etc. are detected then face is also detected otherwise face is not detected. For face recognition extracted features i.e. edges information are computed using Edge Histogram Descriptor and are stored in datasets. Test image is compared with those extracted edges and hence the face is recognized. After successful recognition an attendance of the person can be recorded. For an emotion recognition training images are stored in datasets where common features are removed and unique features are extracted. Average features of all training images are taken and are compared to every single input image. The result is recognized emotion of the image and hence emotion is recognized. Distance of the real valued parameters are calculated and binary parameters are checked if it is present or not. For these computation Artificial Neural Network is preferred in which Sigmoid can be used as an activation function. According to these parameters result is obtained as different facial emotions.

11 Further Works

Face and emotion is recognized after an image is captured. In future face and facial expression can be recognized in live integrated cam which is less time consuming. After the integrated cam is opened pre stored images can be directly recognized in less time and hence emotion is recognized in live cam as per the person changes his/her facial expression.

12 Project Task

This project will be design as per requirements and constraints involve. This project is scheduled to be completed in about 160 days. Given table clears the schedule of the task and labor division.

12.1 Schedule and Work Breakdown

Task	People
Requirement analysis and specification	Abdul, Bibek, Rupak
Undertake analysis of system	Shaksham, Suman
Design system	Abdul, Rupak
Coding	Bibek , Rupak, Suman
Testing and debugging	Abdul, Rupak, Shaksham
Test system model	Suman, Shaksham
Overall system test	Abdul, Bibek, Rupak
Documentation	Bibek, Rupak, Suman, Shaksham

Table 4: Work Breakdown Schedule

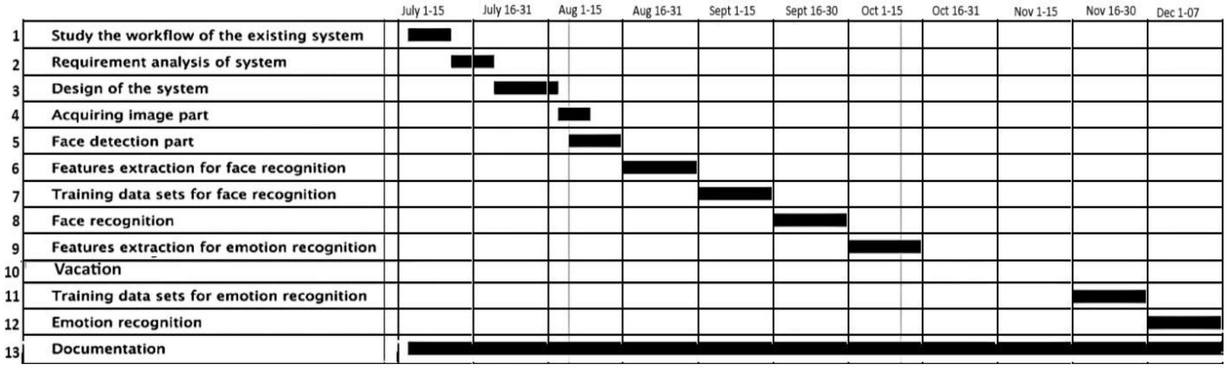

Figure 30: Gantt Chart

13 Bibliography

[1] M. Turk and A. Pentland, “Eigenfaces for recognition,” *J. Cognitive Neuroscience*, vol. 3, pp. 71-86, 1991.

[2] L. Sirovich and M. Kirby, “Low-Dimensional procedure for the characterisation of human faces,” *J. Optical Soc. of Am.*, vol. 4, pp. 519-524, 1987.

[3] M. Kirby and L. Sirovich, “Application of the Karhunen- Loève procedure for the characterisation of human faces,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 12, pp. 831-835, Dec. 1990.

[4] M. Turk and A. Pentland, “Eigenfaces for recognition,” *J. Cognitive Neuroscience*, vol. 3, pp. 71-86, 1991.

[5] M.A. Grudin, “A compact multi-level model for the recognition of facial images,” Ph.D. thesis, Liverpool John Moores Univ., 1997.

[6] L. Zhao and Y.H. Yang, “Theoretical analysis of illumination in pcabased vision systems,” *Pattern Recognition*, vol. 32, pp. 547-564, 1999.

[7] A. Pentland, B. Moghaddam, and T. Starner, “View-Based and modular eigenspaces for face recognition,” *Proc. IEEE CS Conf. Computer Vision and Pattern Recognition*, pp. 84-91, 1994.

[8] K. Chang, K.W. Bowyer and S. Sarkar, “Comarison and combination of ear and face images in appearance-based biometrics,” *IEEE Trans. On Pattern analysis and machine intelligence,* vol. 25, no. 9, September 2003.

[9] L. Hong and A. Jain, “Integrating faces and fingerprints for personal identification,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 20, no. 12, pp. 1295-1307, Dec. 1998.

[10] P. Verlinde, G. Matre, and E. Mayoraz, “Decision fusion using a multilinear classifier,” Proc. *Int’l Conf. Multisource-Multisensor Information Fusion,* vol. 1, pp. 47-53, July 1998.

[11] S. Tamura, H. Kawa, and H. Mitsumoto, “Male/Female identification from 8_6 very low resolution face images byneural network,” *Pattern Recognition*, vol. 29, pp. 331-335, 1996.

[12] T. Kanade, “Picture processing by computer complex and recognition of human faces,” technical report, Dept. Information Science, Kyoto Univ., 1973.

[13] A.J. Goldstein, L.D. Harmon, and A.B. Lesk, “Identification of human faces,” *Proc. IEEE,* vol. 59, pp. 748, 1971.

[14] Y. Kaya and K. Kobayashi, “A basic study on human face recognition,” *Frontiers of Pattern Recognition*, S. Watanabe, ed., pp. 265, 1972.

[15] R. Bruneli and T. Poggio, “Face recognition: features versus templates,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 15, pp. 1042-1052, 1993.

[16] I.J. Cox, J. Ghosn, and P.N. Yianios, “Feature-Based face recognition using mixture-distance,” *Computer Vision and Pattern Recognition*, 1996.

[17] B.S. Manjunath, R. Chellappa, and C. von der Malsburg, “A Feature based approach to face recognition,” *Proc. IEEE CS Conf. Computer Vision and Pattern Recognition*, pp. 373-378, 1992.