EKC 499 FINAL YEAR PROJECT
REPORT 3

Student   : Dharishaan A/L Vengadesan

Matric No   : 120495

H/P   : 0195586256

Supervisor   : Assoc. Prof. Dr Zainal Ahmad

Research Title   : Pattern Recognition using clustering method

Check and approved (Supervisor):

…………………………

Date:

ACKNOWLEDGEMENT

First and foremost, I would like to convey my sincere gratitude to my supervisor, Associate Professor Dr Zainal for his precious encouragement, guidance and generous support throughout this work.

I would also extend my gratitude towards my Dean of School of Chemical Engineering Prof Dr Azalina Harun @ Kamarudin for everything that she has done in helping me out. Her moral support and kindness has helped me during my darkest moments.

Apart from that, I would also like to thank all SCE staffs for their kindness cooperation and helping hands. Indeed their willingness in sharing ideas, knowledge and skills are deeply appreciated. I would like to express my deepest gratitude to my beloved family for their continuous support throughout my life and especially when you really need them for their continuous love and support.

Once again, I would like to thank all the people, including those whom I might have missed out and my friends who have helped me to the accomplishment of this project. Thank you very much.

Dharishaan Vengadesan

May 2018

LIST OF TABLE

LIST OF FIGURES

LIST OF ABBREVIATIONS

(PCA)   Principal Component Analysis

(UV)   Ultraviolet

(MATLAB)   MATrix LABoratory

ABSTRAK

Pengenalan muka telah menjadi satu isu penting dalam banyak aplikasi seperti sistem keselamatan, pengesahan kad kredit dan mengenal pasti penjenayah. Pengenalan muka adalah lebih selamat dalam sistem keselamatan kerana imej wajah telah digunakan sebagai ID. Ia juga membantu untuk mengelakkan sebarang peniruan pengenalan. Pengenalan muka membantu untuk mengenali imej muka terutama dalam mengenalpasti penjenayah tertentu. Mengenal pasti dan membandingkan wajah dalam imej adalah tugas yang sangat kompleks, ini telah melibatkan ramai penyelidik pada tahun-tahun terkini. Objektif projek ini adalah untuk merekabentuk dan membangunkan pengenalan muka menggunakan perisian MATLAB disamping memahami kaedah eigenfaces yang digunakan untuk mengiktiraf wajah imej. Ruang muka ditakrifkan oleh "eigenface", yang mana merupakan eigenvektor bagi sesuatu set muka, mereka tidak semestinya sesuai dengan ciri-ciri yang terpencil seperti mata, telinga dan hidung. Pendekatan eigenfaces seolah-olah menjadi satu kaedah yang mencukupi untuk digunakan dalam menghadapi pengiktirafan kepada kelajuan kesederhanaan, dan keupayaan pembelajaran. Keputusan kajian dilampirkan untuk menunjukkan daya maju cadangan kaedah pengenalan muka ini.

ABSTRACT

Face recognition has become an important issue in many applications such as security systems, credit card verification and criminal identification. Face recognition is more secure in security system because facial image had been used as the ID. It also helps to avoid any duplicated identification. Face recognition helps to recognize the facial image especially to indentifying certain criminals. Identifying and comparing faces in images is a very complex task, this is probably why it has attracted so many researchers in the latest years. Common method used in face recognition like eigenface method will be discussed. The objectives of this project are to design and develop a face recognition using MATLAB software beside to comprehend eigenfaces method of recognizing faces images. The face space is defined by the "eigenface", which are the eigenvectors of the set of faces, they do not necessarily correspond to isolated features such as eyes, ears and noses. Eigenfaces approach seems to be an adequate method to be used in face recognition due to its simplicity, speed and learning capability. Experimental results are given to demonstrate the viability of the proposed face recognition method. On the whole, when the system is modelled and tested, it was unreliable at times. The accuracy of the system was not high and has to be further improved.

CHAPTER ONE

INTRODUCTION

1.1 Project background

Face recognition is typically used in security systems. Besides that, it is also used in human computer interaction. In order to develop this project eigenfaces method is used for training and testing faces. It has received significant attention up to the point that some face recognition conferences have emerged. A general statement of the problem can be formulated as follows, given still or video images of a scene, one or more persons in the scene can be identified using a stored database of faces. The solution of the problem involves face detection, feature extraction from the face regions and recognition. To develop this project we used the eigenfaces method.

Eigenfaces are a set of eigenvectors used in the computer vision problem of human face recognition. A set of eigenfaces can be generated by performing a mathematical process called principal component analysis (PCA) on a large set of images depicting different human faces. The key procedure in PCA is based on Karhumen-Loeve transformation. If the image elements are considered to be random variables, the image may be seen as a sample of a stochastic process. The focus of the research is to find the accuracy of eigenfaces method in face recognition.

We have focused our research toward developing a sort of unsupervised pattern recognition scheme that does not depend on excessive geometry and computations like deformable templates. Eigenfaces approach seemed to be an adequate method to be used in face recognition due to its simplicity, speed and learning capability [2]. The scheme is based on an information theory approach that decomposes face images into a small set of characteristic feature images called eigenfaces, which may be thought of as the principal components of the initial training set of face images[2].

1.2 Problem statements

In security system, many type of password are used to access the private and confidential data. Such password can be as insert characters (key in pin) and touch smart card using RFID technology. Passwords and PINs are hard to remember and can be stolen or guessed; cards, tokens, keys and the like can be misplaced, forgotten, purloined or duplicated; magnetic cards can become corrupted and unreadable. By developed face recognition it more secure because facial image had been used as the ID. It also helps to avoid any duplicated identification.

Other problem is to indentifying certain criminals especially in identification technique used by the police. Face recognition helps to recognize the facial image in more efficient and accurate in order to match with the identity stored in the database.

1.3 Objectives

(1)To comprehend eigenfaces method of recognizing faces images and tests its accuracy.

(2) To design and develop a face recognition using MATLAB

(3) To set up test platform for determining the accuracy of this technique.

(4) To design graphic user interface (GUI) using MATLAB to generated the program.

1.4 Scopes of project

(1) Focus on research about face recognition in eigenfaces method by using MATLAB

(2) Based on software programming and GUI. Simulation by using MATLAB programming to implement face recognition system.

(3) Three main steps to recognition the faces:

    i) Construct face database of known face image

    ii) Input is the unknown face image

    iii) Output is recognizing result. Indentifying the unknown image by comparing with the image in the database

```
┌─────────────────────────────────────────┐
│                                         │
│              Load Image                 │
│                                         │
└─────────────────────────────────────────┘
                    │
                    ▼
┌─────────────────────────────────────────┐
│           Construct Image               │
│                                         │
│             (Eigenface)                 │
└─────────────────────────────────────────┘
                    │
                    ▼
┌─────────────────────────────────────────┐
│                                         │
│          Classified New Faces           │
│                                         │
└─────────────────────────────────────────┘
                    │
                    ▼
┌─────────────────────────────────────────┐
│                                         │
│           Reconstruct Images            │
│                                         │
└─────────────────────────────────────────┘
                    │
                    ▼
┌─────────────────────────────────────────┐
│                                         │
│                Images                   │
│                                         │
└─────────────────────────────────────────┘
```
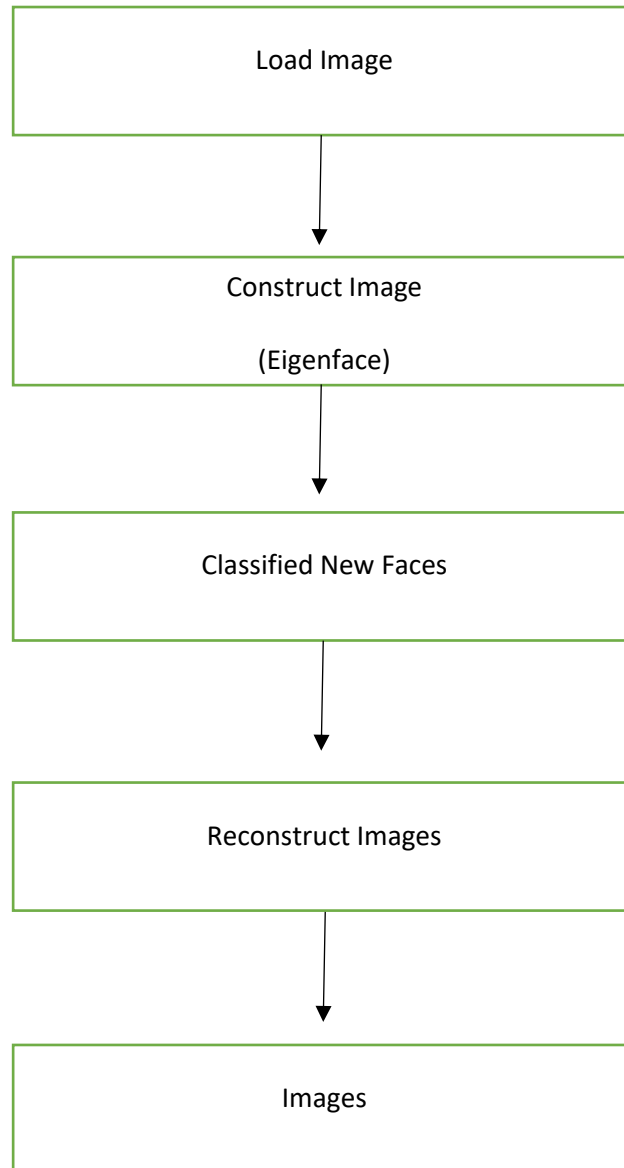
Figure 1.1: Flow of Eigenface

1.5 Report Structure

This report is organized in the following manner: Chapter 2 deals with the literature review concepts of face recognition, three main phases during a face recognition process and research about eigenface method such as eigenface recognition procedure. Chapter 3 is based on the methodology of this project. This chapter will describe about the flow chart of the eigenface. Chapter 4 will show the result and discuss about the face recognition software that was developed to demonstrate the eigenfaces approach. Finally, Chapter 5 gives the conclusion drawn from the research and possible directions for future work.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

Face recognition presents a challenging problem in the field of image analysis and computer vision, and as such has received a great deal of attention over the last few years because of its many applications in various domains. Face recognition techniques can be broadly divided into three categories based on the face data acquisition methodology: methods that operate on intensity images; those that deal with video sequences; and those that require other sensory data such as 3D information or infra-red imagery [6].

In the language of information theory, the objective is to extract the relevant information in a face image, encode it as efficiently as possible, and compare one face encoding with a database of models encoded in the same way [1].
In mathematical terms, the objective is to find the principal components of the distribution of faces, or the eigenvectors of the covariance matrix of the set of face images. These eigenvectors can be thought of as a set of features which together characterize the variation between face images. Each image location contributes more or less to each eigenvector, so that we can display the eigenvector as a sort of ghostly face called an eigenface [1].
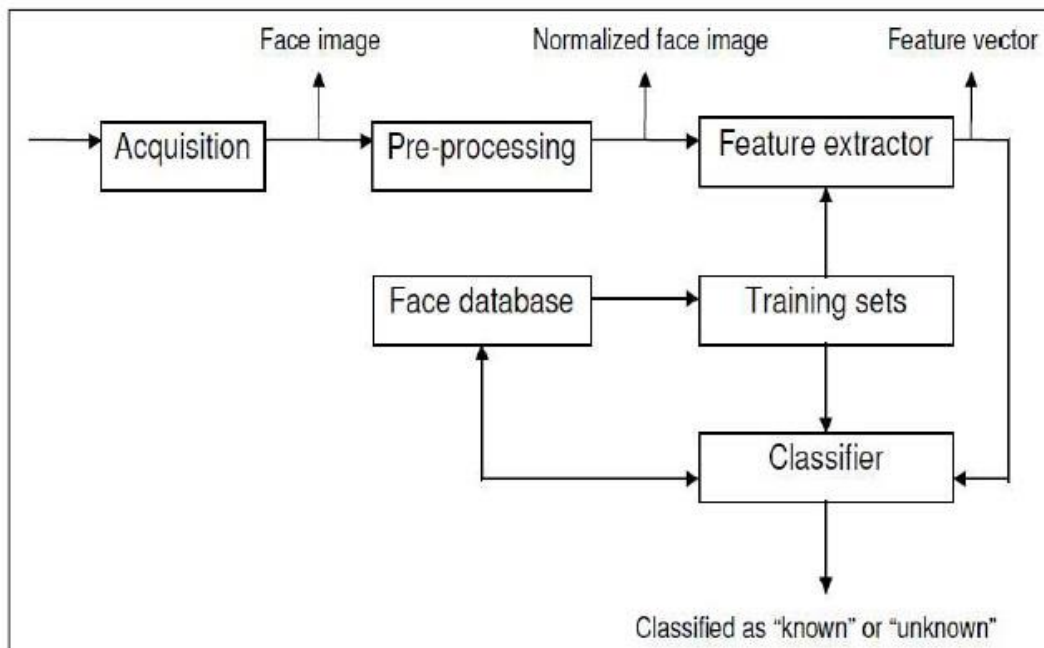
Figure 2.1: Outline of a typical face recognition system [6].

There are six main functional blocks, whose responsibilities are given below [6]:

a) The acquisition module

This is the entry point of the face recognition process. It is the module where the face image under consideration is presented to the system. Another words, the user is asked to present a face image to the face recognition system in this module. An acquisition module can request a face image from several different environments: The face image can be an image file that is located on a magnetic disk, it can be captured by a frame grabber or it can be scanned from paper with the help of a scanner.

b) The pre-processing module.

In this module, by means of early vision techniques, face images are

normalized and if desired, they are enhanced to improve the recognition performance of the system.

c) The feature extraction module

After performing some pre-processing (if necessary), the normalized face image is presented to the feature extraction module inorder to find the key features that are going to be used for classification. Another words, this module is responsible for composing a feature vector that is well enough to represent the face image.

d) The classification module.

In this module, with the help of a pattern classifier, extracted features of the face image is compared with the ones stored in a face library (or face database). After doing this comparison, face image is classified as either known or unknown.

e) Training set.

Training sets are used during the "learning phase" of the face recognition process. The feature extraction and the classification modules adjust their parameters in order to achieve optimum recognition performance by making use of training sets. Face library or face database. After being classified as "unknown", face images can be added to a library (or to a database) with their feature vectors for later comparisons. The classification module makes direct use of the face library.

2.2 Outline of the Proposed Face Recognition System

The proposed face recognition system passes through three main phases during a face recognition process. Three major functional units are involved in these phases and they are depicted in Figure 2.2. The characteristics of these phases in conjunction with the three functional units are given below:
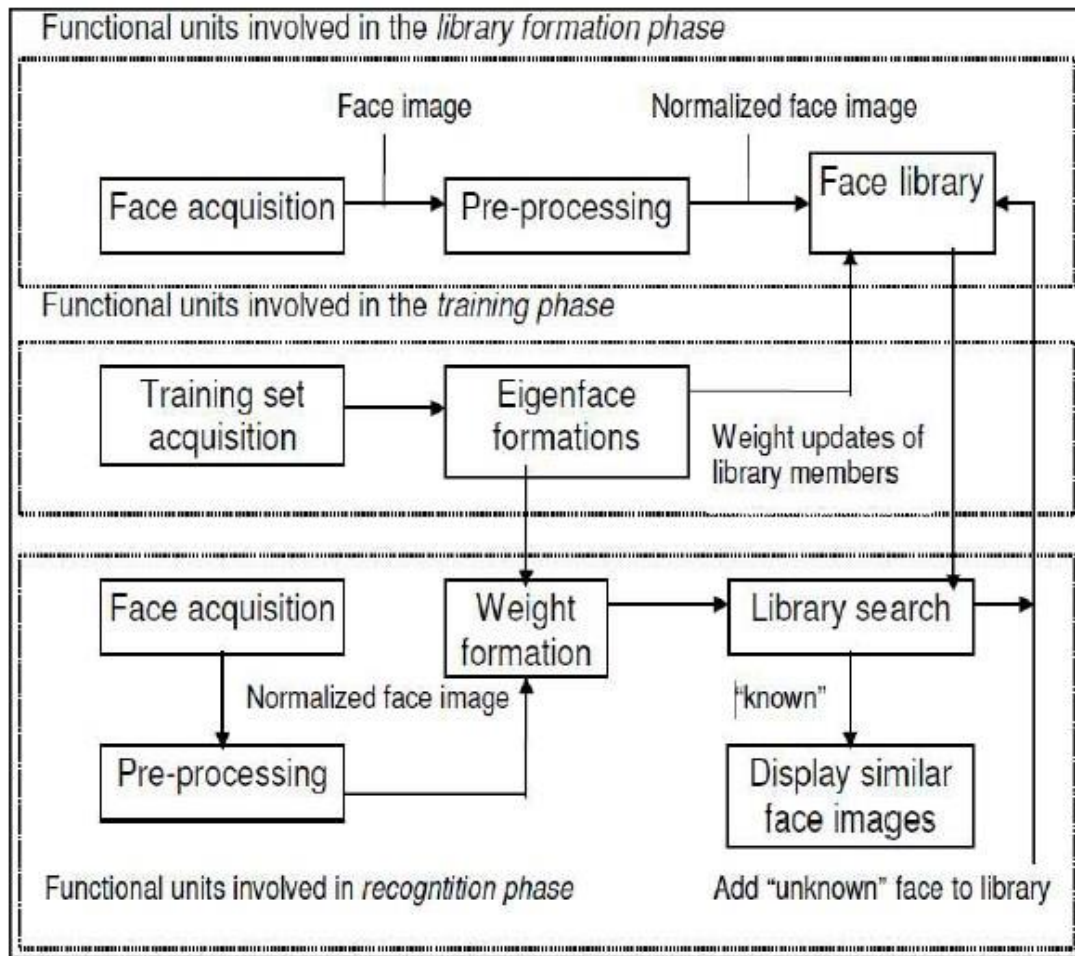
Figure 2.2: Functional Block Diagram of the Proposed Face Recognition System [2]

2.2.1 Face Library Formation Phase

In this phase, the acquisition and the preprocessing of the face images that are going to be added to the face library are performed. Face images are stored in a face library in the system. We call this face database a "face library" because at the moment, it does not have the properties of a relational database. Every action such as training set or eigenface formation is performed on this face library. Face library is initially empty. In order to start the face recognition process, this initially empty face library has to be filled with face images.

The proposed face recognition system operates on 128 x 128 x 8, HIPS formatted image files. At the moment, scanner or camera support is unavailable. In order to perform image size conversions and enhancements on face images, there exists the "pre-processing" module. This module automatically converts every face image to 128 x 128 (if necessary) and based on user request, it can modify the dynamic range of face images (histogram equalization) in order to improve face recognition performance. Also, we have considered implementing a "background removal" algorithm in the pre-processing module, but due to time limitations, we have left this for future work.

After acquisition and pre-processing, face image under consideration is added to the face library. Each face is represented by two entries in the face library: One entry corresponds to the face image itself (for the sake of speed, no data compression is performed on the face image that is stored in the face library) and the other corresponds to the weight vector associated for that face image. Weight vectors of the face library members are empty until a training set is chosen and eigenfaces are formed. [2].

2.2.2 Training Phase

After adding face images to the initially empty face library, the system is ready to perform training set and eigenface formations. Those face images that are going to be in the training set are chosen from the entire face library.

Because that the face library entries are normalized, no further pre-processing is necessary at this step. After choosing the training set, eigenfaces are formed and

stored for later use. Eigenfaces are calculated from the training set, keeping only the M images that correspond to the highest eigenvalues. These M eigenfaces define the M dimensional"face space". As new faces are experienced, the eigenfaces can be updated or recalculated. The corresponding distribution in the M-dimensional weight space is calculated for each face library member, by projecting its face image onto the "face space" spanned by the eigenfaces. Now the corresponding weight vector of each face library member has been updated which were initially empty. The system is now ready for the recognition process.

Once a training set has been chosen, it is not possible to add new members to the face library with the conventional method that is presented in "phase 1" because, the system does not know whether this item already exists in the face library or not. A library search must be performed [2].

2.2.3 Recognition and Learning Phase

After choosing a training set and constructing the weight vectors of face library members, now the system is ready to perform the recognition process. User initiates the recognition process by choosing a face image. Based on the user request and the acquired image size, pre-processing steps are applied to normalize this acquired image to face library specifications (if necessary). Once the image is normalized, its weight vector is constructed with the help of the eigenfaces that were already stored during the training phase.

After obtaining the weight vector, it is compared with the weight vector of every face library member within a user defined "threshold". If there exists at least one face library member that is similar to the acquired image within that threshold then, the face image is classified as "known". Otherwise, a miss has occurred and the face image is classified as "unknown". After being classified as unknown, this new face image can be added to the face library with its corresponding weight vector for later use (learning to recognize) [2].

2.3 Why choose Face Recognition?

Face recognition appears to offer several advantages over other biometric methods, a few of which are outlined here. Almost all these technologies require some voluntary action by the user, for examples, the user needs to place his hand on a hand-rest for fingerprinting or hand geometry detection and has to stand in a fixed position in front of a camera for iris or retina identification [6].
However, face recognition can be done passively without any explicit action or participation on the part of the user since face images can be acquired from a distance by a camera. This is particularly beneficial for security and surveillance purposes [6].

2.3.1 Disadvantage for other biometrics:

(i) Techniques that rely on hands and fingers can be rendered useless if the epidermis tissue is damaged in some way (i.e., bruised or cracked).
(ii) Iris and retina identification require expensive equipment and are too sensitive to any body motion.
(iii) Voice recognition is susceptible to background noises in public places and auditory fluctuations on a phone line or tape recording.
(iv) Signatures can be modified or forged. However, facial images can be easily obtained with an inexpensive fixed camera. Good face recognition algorithms and appropriate reprocessing of the images can compensate for noise and slight variations in orientation, scale and illumination.

2.3.2 Fundamental Issues in Face Recognition

Due to the dynamic nature of face images, a face recognition system encounters various problems during the recognition process. It is possible to classify a face recognition system as either "robust" or "weak" based on its recognition performances under these circumstances [2].

Robust face recognition requires the ability to recognize identity despite many variations in appearance that the face can have in a scene. The face is a 3D object which is illuminated from a variety of light sources and surrounded by arbitrary background data (including other faces). Therefore, the appearance a face has when projected onto a 2D image can vary tremendously. If we wish to develop a system capable of performing non-contrived recognition, we need to find and recognize faces despite these variations. In fact, 3D pose, illumination and foreground background segmentation have been pertinent issues in the field of computer vision as a whole.

Additionally, our detection and recognition scheme must also be capable of tolerating variations in the faces themselves. The human face is not a unique rigid object. There are billions of different faces and each of them can assume a variety of deformations.

2.4 Face recognition system structure

Face Recognition is a term that includes several sub-problems. There are different classifications of these problems in the bibliography. Some of them will be explained on this section. Finally, a general or unified classification will be proposed.

2.4.1 A generic face recognition system

The input of a face recognition system is always an image or video stream. The output is an identification or verification of the subject or subjects that appear in the image or video. Some approaches define a face recognition system as a three step process. From this point of view, the

Face Detection and Feature Extraction phases could run simultaneously.

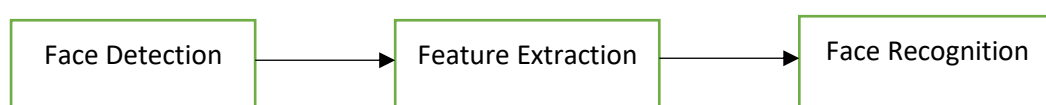Face Detection → Feature Extraction → Face Recognition

Figure 2.3

Face detection is defined as the process of extracting faces from scenes. So, the system positively identifies a certain image region as a face. This procedure has many applications like face tracking, pose estimation or compression. The next step -feature extraction- involves obtaining relevant facial features from the data. These features could be certain face regions, variations, angles or measures, which can be human relevant (e.g. eyes spacing) or not. This phase has other applications like facial feature tracking or emotion recognition. Finally, the system does recognize the face. In an identification task, the system would report an identity from a database. This phase involves a comparison method, a classification algorithm and an accuracy measure. This phase uses methods common to many other areas which also do some classification process -sound engineering, and data mining. These phases can be merged, or new ones could be added. Therefore, we could find many different engineering approaches to a face recognition problem. Face detection and recognition could be performed in tandem, or proceed to an expression analysis before normalizing the face.

2.5 MATLAB

The name MATLAB stands for MATrix LABoratory. MATLAB was written originally to provide easy access to matrix software developed by the LINPACK (linear system package) and EISPACK (Eigen system package) projects. MATLAB is a high-performance language for technical computing. It integrates computation, visualization, and programming environment. MATLAB has many advantages compared to conventional computer languages (e.g., C, FORTRAN) for solving technical problems. MATLAB is an interactive system whose basic data element is an array that does not require dimensioning. Specific applications are collected in packages referred to as toolbox. There are tool boxes for signal processing, symbolic computation, control theory, simulation, optimization, and several other fields of applied science and engineering.

2.5.1 MATLAB's Power of Computational Mathematics

MATLAB is used in every facet of computational mathematics. Following are some commonly used mathematical calculations where it is used most commonly:

* • Dealing with Matrices and Arrays

* • 2-D and 3-D Plotting and graphics

* • Linear Algebra

* • Algebraic Equations

* • Non-linear Functions

* • Statistics

* • Data Analysis

* • Calculus and Differential Equations Numerical Calculations

* • Integration

* • Transforms

- • Curve Fitting

- • Various other special functions

2.5.2 FEATURES OF MATLAB

Following are the basic features of MATLAB

It is a high-level language for numerical computation, visualization and application development.

- • It also provides an interactive environment for iterative exploration, design and problem solving.

- • It provides vast library of mathematical functions for linear algebra, statistics, Fourier analysis, filtering, optimization, numerical integration and solving ordinary differential equations.

- • It provides built-in graphics for visualizing data and tools for creating custom plots.

- • MATLAB's programming interface gives development tools for improving code quality, maintainability, and maximizing performance.

- • It provides tools for building applications with custom graphical interfaces.

- • It provides functions for integrating MATLAB based algorithms with external applications and languages such as C, Java, .NET and Microsoft Excel.

2.5.3 USES OF MATLAB