PEOPLE'S DEMOCRATIC REPUBLIC OF ALGERIA
MINISTRY OF HIGHER EDUCATION AND SCIENTIFIC RESEARCH
MOHAMED BOUDIAF UNIVERSITY - M'SILA

1985
جامعة محمد بوضياف - المسيلة
University of Mohamed Boudiaf-Msila

FACULTY OF MATHMATICS AND COMPUTER SCIENCE
DEPARTEMENT OF COMPUTER SCIENCE

In partial fulfillment of the requirements for the degree of

Master in **Computer Science**

By

Amron Abedel Arhim

Bennaama Mohamed

Title of the thesis

The conception a monitoring system based on the recognition of face, voice, and fingerprint

---

Under the supervision of

Loucif Hemza

Composition of the jury

Prof.Loucif Hemza	University of Msila	President
Prof. Attir Azeddine	University of Msila	Reporter
Prof. Therafi Abdellah	University of Msila	Examiner

June, 2022

Acknowledgments

Firstly, I am grateful to Almighty Allah for my health and well complete this project work. Secondly, I would supervisor, like **Dr. Loucif Hemza** to expressbeing that was necessary to my sincere gratitude to my for his valuable guidance, continuous support and patience. Last but not least I want to thank myself for all this hard work. Finally, I would like to thank my family for their unceasing encouragement, appreciable patience, support and attention. I also thank the honorable students who participated in this work.

Table of Contents

List of figures v

List of tables vi

Introduction 1

Chapter1: Facial Recognition 2

1.1 Introduction 2

1.2. Facial Detection 3

1.2.1. Evolution of Facial Detection 3

1.2.2. Facial Detection Approaches 4

1.2.3 Comparison between the difierent Face Detection Ap- 7

Proach : 7

1.2.4 Pre-treatment 8

1.3 Facial Recognition 10

1.3.1 Main di_culties of face recognition 10

1.3.2 Facial Recognition Approaches 11

1.3.3 Facial Recognition Applications 13

1.4 Conclusion 14

Chapter 2: Fingerprints 15

2.1 Introduction 15

2.2 Fingerprint characteristics 15

2.2.1 Classification and pattern types 16

2.2.2 Terminology 16

2.3 Enhancement techniques 19

2.4 . Material and method 19

2.4.1 Minutiae Matching 20

2.5 Conclusion 21

Chapter 3: Speech Recognition 22

3.1 Introduction 22

3.2 General Steps For Speech Recognition Systems 22

3.3 Techniques In Speech Recognition Systems 23

3.3.1 Speech Analysis 23

3.3.2 Speech Feature Extraction Techniques 24

3.3.3 Modeling Technique 24

3.3.4 Matching Techniques 25

3.4 Comparative Study Of Speech Recognition System Approaches 26
3.5 Conclusion 27
Chapter 4: Classification and Extraction of parameters 28
4.1 Introduction 28
4.2 face recognition 28
4.2.1 Principal Component Analysis 28
4.2.2 Discrete Cosine Transform 31
4.3 voice recognition 38
4.3.1 Linear Predictive Coding (LPC) 38
Conclusion 41

List of figures

Figure 1.1 Facial Detection Approaches
Figure 1.2 Face model consisting of 16 regions associated with 23 relationships
Figure 1.3 Rowley et al neuron network modal
Figure 2.1 Representation of a facial recognition system
Figure 2.2 Example of lighting variation
Figure 2.3 Example of Variation in Pose
Figure 2.4 Example of facial expressions
Figure 2.5 Facial Recognition Approaches
Figure 3.1 Example of projection following PCA
Figure 3.2 Representation of a sphere with the Euclidean distance and the City-Block distance
Figure 3.3 One dimensional cosine basis function (N=8
Figure 3.4 Two dimensional DCT basis functions (N = 8)
Figure 3.5 Decorrelation DCT Figure
Figure 3.6 Energy Compaction DCT Figure
Figure 3.7 Computation of 2-D DCT using separability property

List of tables

Table 1.1: Represents the advantages and disadvantages of each Approach
Table 2.1: Comparison of methods based on local or global characteristics

Introduction

The In today rapidly evolving world, security and surveillance systems play a crucial role in ensuring the safety of individuals and properties. Traditional methods of identification and monitoring often rely solely on one biometric modality, such as facial recognition. However, these methods may have limitations in accuracy and robustness.

To address these limitations, the conception of a monitoring system based on the recognition of face, voice, and fingerprint emerges as a comprehensive and reliable solution. This integrated system harnesses the power of multiple biometric modalities to enhance identification accuracy and provide a more robust security framework.

The first component of this system is face recognition. By utilizing advanced algorithms and techniques, the system can accurately identify individuals by analyzing unique facial features. This allows for real-time monitoring and identification, even in challenging environmental conditions.

The second component is voice recognition. Voice biometrics offer a distinct advantage in identifying individuals based on their unique voice patterns. The system captures voice samples and compares them against pre-registered voice models, enabling accurate identification and verification.

The third component is fingerprint recognition, a widely recognized and trusted biometric modality. By capturing and analyzing fingerprints can establish a reliable and secure method for identifying individuals.

By integrating these three biometric modalities - face, voice, and fingerprint recognition - into a single monitoring system, a higher level of security and accuracy can be achieved. This multi-modal approach minimizes the chances.

Chapter1: Facial Recognition

1.1 Introduction

Facial recognition has become a must in modern times as it ensures the protection of individuals and institutions. Compared to other password based and other easily hacked systems, facial recognition based security and protection systems are reliable because facial recognition systems are based on a major part of the human face.

Face recognition is the step after face recognition in this section. It is used to solve the main difficulties of face recognition and methods to improve the face recognition system. It is divided into three general methods: global method, local method and hybrid method. Finally, we mentioned some applications that work with these systems.

1.2. Facial Detection

The first step in face recognition is face recognition. In this chapter, we introduce the development of face recognition technology and the methods used, as well as the preprocessing stage. Finally, we see some databases that provide us with facial images and some applications that use the technology.

1.2.1. Evolution of Facial Detection

Early efforts in face detection have dated back as early as the beginning of the 1970s, where simple heuristic and anthropometric techniques were used [1] .

Because face detection techniques require a priori information of the face, they can be effectively organized into two broad categories distinguished by their different approach to utilizing face knowledge. The techniques in the first category make explicit use of face knowledge and follow the classical detection methodology in which low level features are derived prior to knowledge-based analysis [2].

1.2.2. Facial Detection Approaches

There are several techniques for face detection. They can be subdivided into four different categories [3].

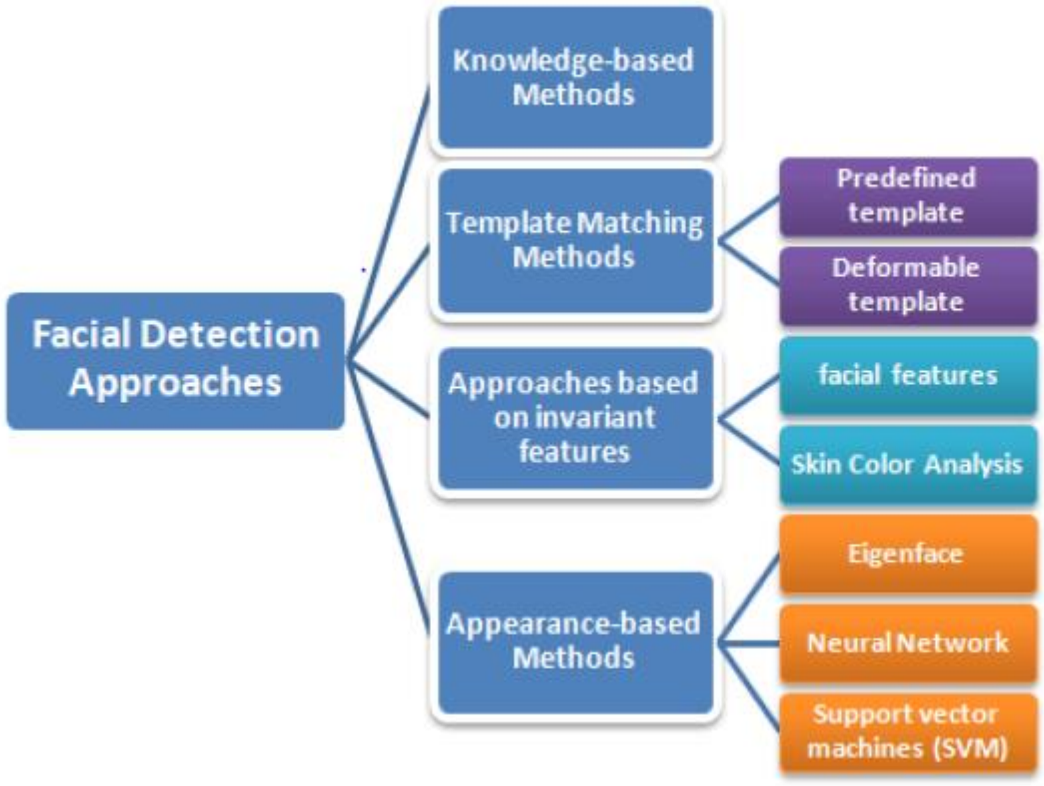

Figure 1.1: Facial Detection Approaches [4]

1.2.2.1 Knowledge-based Methods :

These methods are based on the definition of strict rules based on the relationships between facial features. They are interested in the characteristic parts of the face such as the nose, mouth and eyes. These methods are designed primarily for localization of face [4] . Kotropoulous and Pitas [5] use a rule-based method. The facial features are localized using the projection method proposed by Kanade [6] to detect facial contours .

1.2.2.2 Template Matching Methods :

Definition of templates matching is either "manual" or configured using function.

the idea in this approach is to calculate the relationship between the filtered faces and the template. However, these methods still face problems of strength associated with differences in light, scale, etc . Sinha [7] uses a set of invariants describing the model of the face. To determine the invariants to changes in brightness to characterize different parts of the face, this algorithm calculates the ratio of luminance between the regions of the face and retains the directions of these reports.

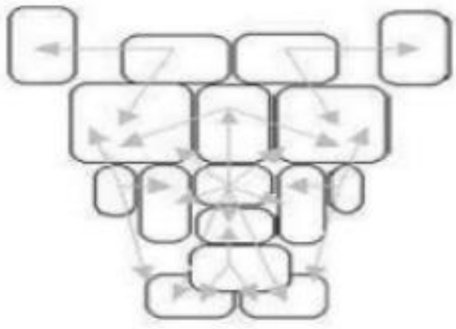

Figure 1.2: Face model consisting of 16 regions (rectangles) associated with 23 relationships(arrows) [5].

This predefined model corresponding to 23 relationships. These predefined relationships are classified into 11 key relationships (arrows) and 12 confirmations relationships (gray). Each arrow represents a relationship between two regions. A relationship is verified if the relationship between the two regions exceeds a threshold. The face is localized if the number of essential relationships and confirmation also exceeds a threshold.

Yuille et al [9] used a deformable Template to model facial features, they created a Template elastic adaptive to facial features such as eyes, mouth...etc. The Template parameterize this technique allows to describe the facial features.

1.2.2.3 Approaches based on invariant features :

These methods are mainly used to locate faces. new algorithm Designed to create existing structural features, even if they are constructs, Display or change the lighting state. You can then use these immutable traits to Identify faces. The algorithms of this method can be divided into two categories which one:

- **Based on facial features**
- **Based on skin color analysis**

1.2.2.4 Approaches based on invariant features:

- **facial features :** In general these algorithms use first a hypothesis about the position of the upper face then the search algorithm runs up and down the face in order to find the axis of the eyes characterized by a sudden increase in the density of contours (measured by the black/white ratio along the horizontal planes)[5].

- **Skin Color Analysis :** Detection methods based on skin color analysis are efficient and fast methods. They reduce the research space of the facial area in image. In addition, the color of the skin is a robust information about the rotations, the changes of scale and partial occultations. Several color spaces can be used to detect, in the image, the pixels that have the color of the skin. The effectiveness of the detection depends essentially on the chosen color space [5].

1.2.2.5 Appearance-based methods :

These approaches generally apply machine learning techniques.Thus, the models are learned from a set of images representative of the variability of the facial appearance [5].

- **Eigenface :** The first to develop this method was Turk et pentland [10] in 1991. Eigenface is one of the most known methods of facial detection. The principle of this method is to display an image in space and then calculate the distance between the original image and its projection, encoding an image in space that serves to analyze the information found in the class.the image is not correctly represented in the space it does not contain a face. This method gives rather encouraging results, but the calculation time is very important [11]

- **Neural Network :** In Rowley et al [12], the authors propose a facial detection system based on classification by neural networks. Their technique is divided into two stages: location of faces using a neural network and verification of results obtained. The authors constructed a network of neurons that, from a pre-processed image 20x20 pixels, indicates whether it is a face or not.

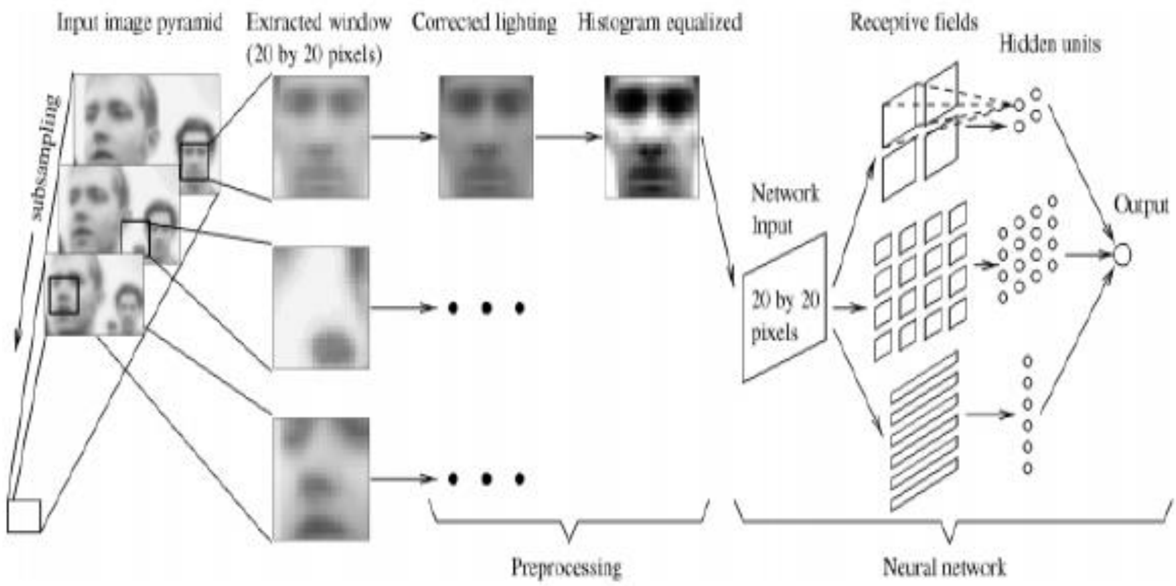

Figure 1.3: Rowley et al neuron network modal [12]

- **Support vector machines (SVM) :** One of the first statistical methods based on information theory for face detection, SVM is considered as a new model to classify training polynomial function, neural network or radial basis function (RBF).

1.2.3 Comparison between the difierent Face Detection Ap-Proach :

Approach	Advantages	Disadvantages
Knowledge-based Methods	-reduce the calculation time necessary by use of sub-sampled images	-causes many false detection -difficult to translate the human knowledge in well-defined rules
Template Matching Methods	-Simple at the level of detection process. -Gives fairly good results encourage.	-Update to each change in focus
Approaches based on invariant features	-Resists the small ones lighting changes and the face position -reduced skin color the search box	-The use of the method based on the color of the skin requires processes to complete the detection of face.
Appearance-based methods	-Eigenface gives good results. -Neural networks robust to noise.	-Lots of calculation time -difficult to build-learning phase difficult to conduct

Table 1.1: Represents the advantages and disadvantages of each Approach.

1.2.4 Pre-treatment

The preprocessing stage precedes the detection stage. It enables image faces to be processed in a way that can be used during the registration phase. It is also called the normalization stage because it returns all the images extracted from the original image. It usually consists of centering faces in the image and eliminating non-informative areas.

1.2.4.1 Geometric Normalization

is required because the size of the face in the captured image may vary depending on the distance between the capture module and the person. So the face has to be extracted from the image and a geometric transformation applied to get a fixed size. The standard approach is to define the position of the center of the eye as a constant position within the output image.

1.2.4.2 Photometric Normalization

Photometric Normalization: Attempts to remove or reduce the effect of lighting on the image.

1.2.5 Facial Detection Applications

Face detection technology can be useful and necessary in a wide range of applications, including biometric identification, video conferencing, indexing of image and video databases, and intelligent human–computer interfaces. In this section we give a brief presentation of some of these applications.

1.2.5.1 Face Recognition Systems

As mentioned earlier, face detection is most widely used as a preprocessor in face recognition systems. Face recognition is one of many possible approaches to biometric identification; thus many biometric systems are based on face recognition in combination with other biometric features such as voice or fingerprints. In BioID [14], a model-based face detection method based on edge information [15] is used as a preprocessor in a biometric systems which combines face, voice, and lip movement recognition. Other biometric systems using face detection include template-based methods in the CSIRO PC-Check system [16] and eigenface methods [17] [18] in FaceID from Visage Technologies.

1.2.5.2 Images and Videos

The idea behind CBIR systems using face detection is that faces represent an important cue for image content; thus digital video libraries consisting of terabytes of video and audio information have also percieved the importance of face detection technology. One such example is the Informedia project [20]which provides search and retrieval of TV news and documentary broadcasts. Name-It [21]also processes news videos, but is focused on associating names with faces. Both systems use the face detection system of Rowley et al [12].

1.2.5.3 Video Conferencing Systems

In video conferencing systems, there is a need to automatically control the camera in such a way that the current speaker always has the focus. One simple approach to this is to guide the camera based on sound or simple cues such as motion and skin color. A more complex approach is taken by Wang et al [22],who propose an automatic video conferencing system consisting of multiple cameras, where decisions involving which camera to use are based on an estimate of the head's gazing angle. Human faces are detected using features derived from motion, contour geometry, color, and facial analysis. The gazing angle is computed based on the hairline (the border between an individuals hair and skin)

1.3 Facial Recognition

Face recognition is a technique of identifying persons based on their faces, Since the sixties this technique has grown and has been widely applied to identify individuals and suspects in the largest security companies and government [23].

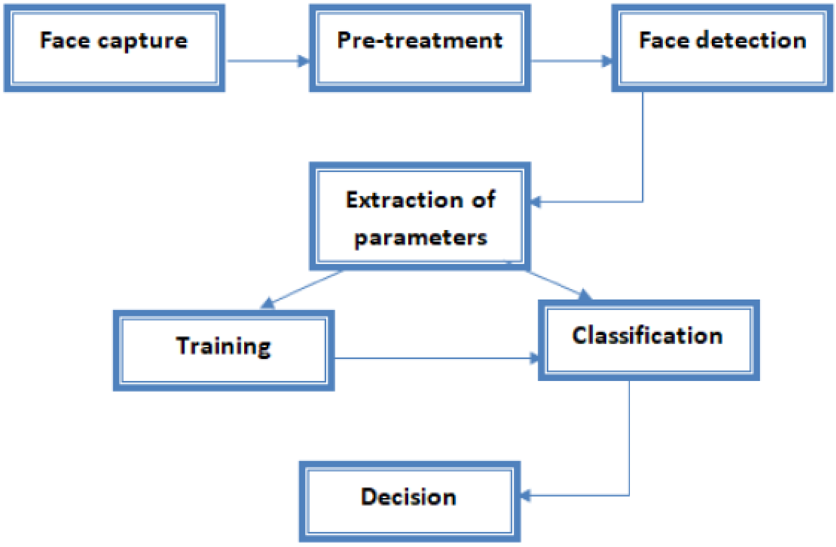

Figure 2.1: Representation of a facial recognition system

1.3.1 Main di_culties of face recognition

Facial recognition for the human brain is a high-level visual task. Although humans can easily detect and identify faces in the landscape, building an automatic system that performs these tasks is a major challenge.

1.3.1.1 Change in illumination

The appearance of a face in an image varies greatly depending on the illumination of the scene during shooting .

Figure 2.2: Example of lighting variation

1.3.1.2 Variation in Pose

The face recognition rate decreases considerably when variations in pose are present in the images. This difficulty has been demonstrated by evaluation tests developed on the FERET and FRVT bases [25] [26].

Figure 2.3: Example of Variation in Pose

1.3.1.3 Facial expressions

Another factor that affects the appearance of the face is facial expression The deformation of the face due to facial expressions is localized mainly on the lower part of the face.

Figure 2.4: Example of facial expressions

1.3.1.4 Presence or absence of structural components

The presence of structural components such as beard, moustache, or glasses can dramatically alter facial features such as shape, color, or face size. In addition, these components can hide the basic facial features causing a failure of the recognition system.

1.3.1.5 Partial occultations

The face can be partially masked by objects in the scene, or by the port accessories such as glasses, wrap... In the context of biometrics, the proposed systems must be non-intrusive, that is, they must not rely on active cooperation on the subject. Therefore, it is important to recognize partially hidden faces. Gross et al [27] studied the impact of wearing sunglasses, and the masking of the lower part of the face on facial recognition. They used the AR [28] database. Their experimental results suggest that under these conditions, the performance of the recognition algorithms remains low.

1.3.2 Facial Recognition Approaches

There are many facial recognition approaches and these have been classified into 3 categories according to Tan et al [29] we can devise or classify approaches to facial recognition into three large families that are:

1.3.2.1 Global Methods (Holistic Methods)

And is called the holistic approach, The principle of this approach according to O'Toole et al [30] is to represent the image of the face by a single vector of large dimension n x m, in concatenating the gray levels of all the pixels of the face.The advantage of this representation is that it implicitly preserves the texture and shape information needed for face recognition. In addition, it allows a better capture of the overall aspect of the face than local representations [30] However, its major disadvantage lies in the very large size of the image space it requires, which makes classification very difficult.

1.3.2.1.1 Principal Component Analysis (PCA)

Also called "Eigenfaces" a very popular method in the field of recognition proposed by Turk and pentland [10]. A mathematical method used to simplify and reduce the dimensions of a data set and to represent facial images that can be reconstructed from a standard face and a set of points.It is a matter of finding all the main components of face in an image set of faces.

1.3.2.1.2 Latent Dirichlet Allocation (LDA)

It's also known as "Fisherfaces" Belhumeur et al []are the first ones who introduced this algorithm in 1997, it performs a class separation and to be able to use it it is necessary to organize a learning base of images in several classes, one class per person and several images per class.

This method is used effectively in many classification and dimension reduction. However, if the data are too large, it is not possible to apply this method directly to the images without first decreasing the size of the data. In this case, instead of using the image pixel value directly, a PCR is first applied to the data and it is the image representation in the face space that is used [32].

1.3.2.2 Local Methods

Local methods use local facial features for recognition face. They are relatively mature compared to holistic methods. In these methods, the face is represented by a set of characteristic vectors of small dimensions, rather than by a single vector of large size [33]. The local approach can be divided into two categories:

- **Methods based on local features:** extractions and location of features Points.

- **Methods based on local appearances:** scores of facial images in the regions features.

1.3.2.3 Hybrid methods

Hybrid methods are approaches that combine holistic methods with local methods to improve facial recognition performance. In fact, local and global characteristics each have quite different characteristics. It is hoped that their integration can be used to improve classification.

Factors for change	local feature	Global feature
illumination	very sensitive	sensitive
expression	not sensitive	sensitive
Pose	sensitive	very sensitive
noise	very sensitive	very sensitive

Table 2.1: Comparison of methods based on local or global characteristics.

1.3.3 Facial Recognition Applications

Facial recognition technology has become a necessary technology in some applications. In fact, the use of this technology is not only in the computer science industry but can be provided in many applications useful to all types of industries, such as educational systems and law.

1.3.3.1 Security

Face recognition technology can be used as another form of authentication when attempting to access personal accounts. Whether that be personal computers, banking accounts, or even email accounts having an extra layer of security can be invaluable when a person's livelihood is on the line. Automated Teller Machines (ATMs) already have cameras installed to help prevent tampering. This camera can instead be used as a form of identification, only allowing the user to access their account if their face is in the system. Securing this data behind a biometric security system, such as face recognition, can be vital. Facial recognition is already being used as a form of access control. Jeffrey S. Coffin developed a security method that uses face recognition to identify individuals in order to determine their clearance access (Coffin, 1999). In the media,biometric security measures are generally boiled down to retinal scans, and finger print scans.Facial recognition is a very powerful biometric security system that can improve any access system [37].

1.3.3.2 Attendance Systems

The educational systems have some processes that can be improved by employing facial recognition technology. Taking attendance is an important part of class time to ensure that all students are present; however, the larger the class the more difficult and longer it takes to check.The problem is compounded in classes that fill entire auditoriums. These systems, though not used often, already exist. Abhishek Jha developed an automated attendance system that can invaluable in the education industry [38]. This system can help cut down the amount of time needed in class by allowing the professor to jump straight into the material rather than waste time calling everyone's name one at a time. Many universities have already forgone the use of attendance because sometimes it just isn't feasible with the number of students they receive.By automating this process, the university can track which students are attending lectures.

1.3.3.3 Law Enforcement

Law Enforcement is another industry that can benefit from facial recognition software.One of the jobs required of policemen is tracking down criminals. Face recognition software can automate this process by automatically tracking these criminals without needing any manpower besides making the arrest. Cameras are set up all around the country, from traffic cameras, street cameras, to private security cameras. The law enforcement officials can feed their database of criminals into a facial recognition algorithm and use these cameras to track criminals all around the world. Writing speeding tickets is another job police officers bog their time down with.Sensors can be set up to check whether the car is going too fast. Then cameras can recognize the face and automatically send a speeding ticket. Setting up cameras and speeding sensors may not be very cost effective now; however, as technology progresses the cost of these things continue to lower. [37].

1.4 Conclusion

In this chapter, we have demonstrated the mechanisms of face detection and the two most popular facial recognition algorithms, highlighting their most important characteristics. So we define the PCA algorithm and detail the steps of the algorithm as well as the method for calculating Euclidean distances. , and then we define the DCT algorithm and present its mathematical method in a one- and two-dimensional manner, and we also highlight the most important characteristics of DCT and compare it with DFT and KLT.

Chapter 2: Fingerprints

2.1 Introduction

By seven months of age, the fetus's fingerprints are fully developed. The properties of fingerprints do not change throughout life, except in the event of injury, disease, or decomposition after death. However, after a minor injury to the fingertip, the pattern will grow back as the fingertip heals. [41, 49]

This chapter begins with some important historical events related to fingerprints, in particular fingerprints as a means of identification. What follows is a brief introduction to how fingerprints are viewed in contemporary society. Fingerprint features and enhancement techniques are also discussed to provide readers with a better platform before reading the following chapters.

2.2 Fingerprint characteristics

At some point in your life, you may have seen your own fingerprints and Notice the ridgeline above. In fingerprint literature, the terms ridge and Valleys are used to describe the higher and lower parts of a ridgeline. That The reason we have ridges and valleys on our fingers is due to the frictional ability of the skin [48].

The formation of ridges and valleys is a combination of genetic and environmental factors. psychological factors. DNA dictates the formation of fetal skin, But the exact formation of a fingerprint is a series of random events. That The exact position and exact position of the fetus in the womb at a given time The composition and density of the surrounding amniotic fluid is decisive, as is each individual Glitch formation. [25]

This is also why fingerprints are on different fingers of the same person Completely different and why identical twins have different fingerprints, see section 3.2

2.2.1 Classification and pattern types

Throughout history, fingerprints can and have been classified in different ways, See section 3.1 on page 13. The Henry classification system is the basis of modern classification. Day AFIS classification method until the 1990s. In recent years, the Henry classification system has been superseded by the first-class classification method used in most forensic departments. These new classification methods use distances between core and triangular points, minutiae locations, and pattern types (the latter using the Henry classification system). [twenty two] Fingerprints can be divided into three main pattern types, arcs, rings and threads, as shown in **Figure 3.1**. Looping is the most common fingerprint pattern [27].

These main pattern types can come in different variations. For example, you can find simple or tented (narrow) arches, right or left loops, and spirals or concentric circles as threads. In addition, different pattern types can be combined into a fingerprint, such as a fingerprint. Double or looped bows [5].

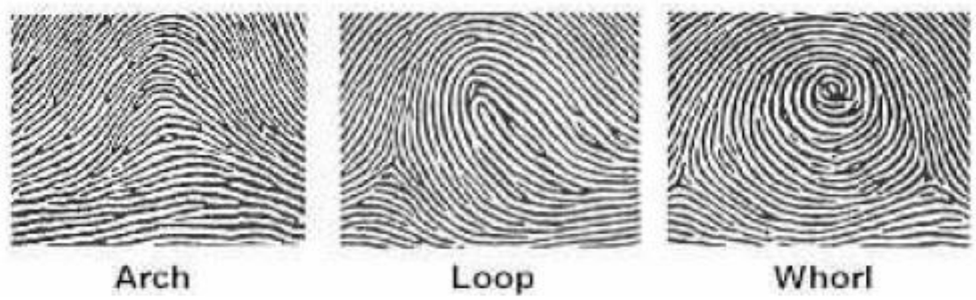

Figure 2.1. The three major pattern types: arches, loops, and whorls. These major pattern types can be divided further into di®erent subgroups: right or left loops, plain or tented (narrow) arches, and spiral or concentric circles as whorls. There are also combinations of these di®erent pattern types. [22]

2.2.2 Terminology

To understand the basics of fingerprinting, the same approach used by [41] is presented here. Fingerprints can be viewed at different levels; global level, local level, and extremely fine-grained level [41].

On a global level, you can find singularities called core and delta points, see Figure 3.2 These singularities are important for fingerprint classification, but not sufficient for accurate matching [41].

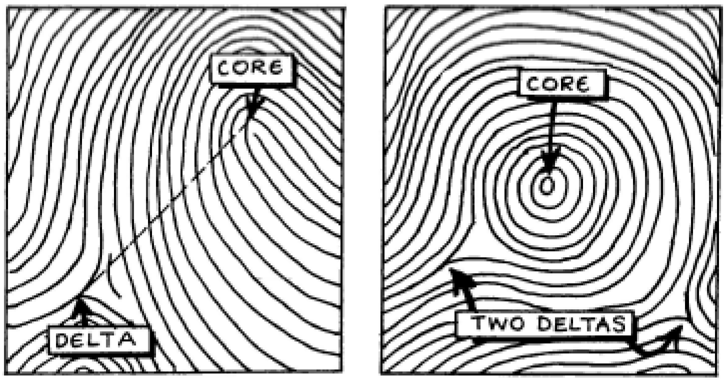

Figure 2.2. Core and delta points marked on sketches of the two fingerprint patterns loop and whorl. Loops have one delta, whorls have two. Minutiae details are not shown. The number of intervening ridges from delta to core in the leftmost pattern (loop) is 12. A ridge tracing from left to right delta on the rightmost pattern (whorl) determines an inner tracing, meaning that when following a ridge emanating from the left delta, the ridge passes inside the other delta. [7]
At the local level, you find the *minutiae details* (sometimes called minutiae points). One way to classify the minutiae details are in terms of ridge termination, bifurcation, independent ridge, dot or island, lake, spur, and crossover [7]. These are depicted in **Figure 3.3**. The two most prominent minutiae details, are ridge termination (ending) and ridge bifurcation [41].

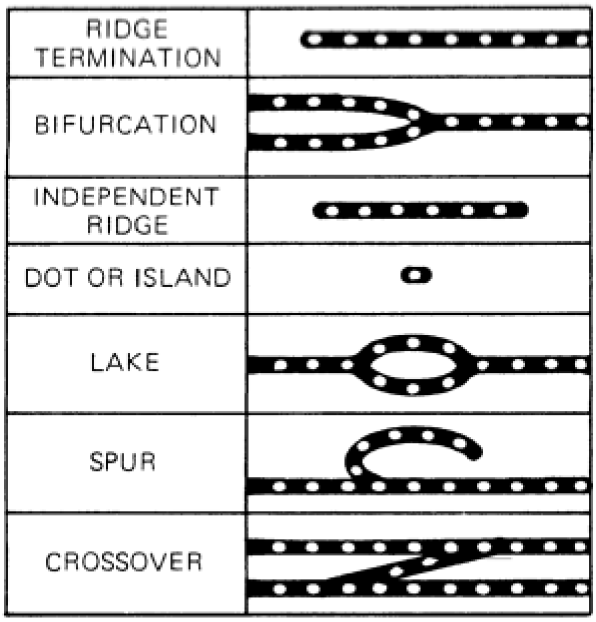

Figure 2.3. Minutiae details, also known as ridge characteristics, ridge details, or Galton's details. Most of the identi¯cations of ¯ngerprints during this century, were made by matching corresponding minutiae details between two prints. [7]
. The position and shape of the pores can be used to help identify a person.
To be able to use this information, a high-resolution image of the fingerprint .

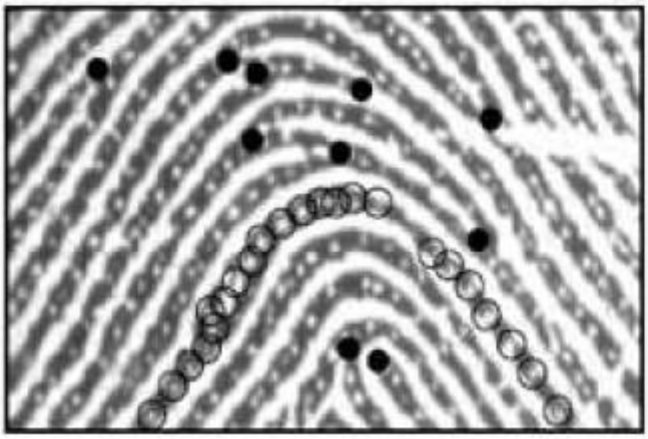

Figure 2.4. Part of a fingerprint image with sweat pores and minutiae details visible. The black lines in the image correspond to the ridges in the fingerprint, and the white lines in the image correspond to the valleys in the fingerprint. The white dots on the ridges correspond to the sweat pores in the fingerprint and are marked with empty circles
on a single ridge line. Minutiae details are marked with black-filled circles. [41]
Figure 2.5 shows a cross-section of a papillary line. The sweat glands supply the papillary skin with moisture and when touching a surface with a finger, the sweat from these pores is transferred to the pattern of the fingerprint, see figure 3.4. The outer skin layer is called epidermis, and the inner skin layer is called dermis.

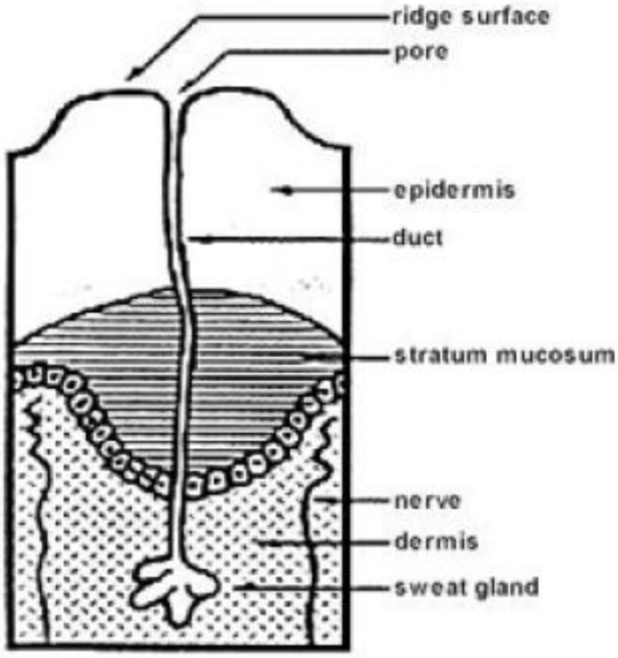

Figure 2.5. Cross-section of a papillary line of a fingerprint. [48]

2.3 Enhancement techniques

A latent fingerprint results from the reproduction of friction ridges found on fingers. To be able to identify the owner of the fingerprint, the fingerprint must in most cases first be enhanced in order for it to be visible. Enhancing a fingerprint will also be used in the experiments described .

A print consists of a combination of different chemicals that originate from natural secretions, blood, and contaminants. Some contaminants found in fingerprints result from contact with different materials in the environment. [56]

Latent fingerprints can be found on all types of surfaces. In general, surfaces can be characterized as porous, nonporous, or semiporous. Understanding these characteristics helps in deciding the processing technique of the latent fingerprint.[57]

2.4 . Material and method

Fingerprint identification system is mainly divided into three modules; fingerprint image preprocessing , minutiae extraction and minutiae matching [9]. The input fingerprint image is processed for skeleton image by the fingerprint image preprocessing stage and subsequently processed by minutiae extraction stage for extracting minutiae using crossing number concept. After minutiae extraction stage, if input fingerprint image is processed for enrollment then the skeleton image is saved as template fingerprint image in database, otherwise skeleton image is given to matching stage. In matching stage system compares skeleton image with template fingerprint images from database and make decision whether input fingerprint match or not.

Minutiae matching and detection is employed in this research which is the most common method of extraction using the Crossing Number (CN) concept. This method involves the use of the skeleton image where the ridge flow pattern is eight-connected. The minutiae are extracted by scanning the local neighbourhood of each ridge pixel in the image using a 3×3 window. The CN value is then computed, which is defined as half the sum of the differences between pairs of adjacent pixels in the eight-neighbourhood [3]. The CN for a ridge pixel P is given by

$$CN = \frac{1}{2}\sum_{i=1}^{8} | Pi - 1 | , P_9 = P_1$$

Where Pi is the pixel value in the neighbourhood of P. For a pixel P, its eight neighbouring pixels are scanned in an anticlockwise direction as display in figure 1: After Cross Number for a ridge pixel has been computed, the pixel can then be classified according to the property of its CN value. Using the properties of the CN as shown in Table 1, the ridge pixel can then be classified as a ridge ending, bifurcation or non-minutiae point. For example, a ridge pixel with a CN of one corresponds to a ridge ending, and a CN of three corresponds to a bifurcation. Table 1: The table below shows the Properties of Crossing Number. [4].

CN	Property
0	Isolated Point
1	Ridge Ending Point
2	Continuing Ridge Point
3	Bifurcation Point
4	Crossing Point

Table 2.1 International Journal of Computer Applications (0975 – 8887) Volume 179 – No.21, Febru ary 2018

2.4.1 Minutiae Matching

Let T and I be the representation of the template and input fingerprint, respectively. Each minutia is considered as a triplet m={x,y, θ} that indicates the x, y minutia location coordinates and the minutia angle θ

T={m1,m2,….mm}
m={x1,y1, θ1}
i=1,2,….m
I={m'1,m'2,….m'n}
m'j={x'j,y'j, θ'j}
j=1,2,….n

where m and n denote the number of minutiae in T and I, respectively. A minutia m′j in I and a minutia mi in T are considered "matching", if the spatial distance (sd) between them is smaller than a given tolerance r0 and the direction difference (dd) between them is smaller than an angular tolerance θ0 [8].

P4	P3	P2
P5	P	P1
P6	P7	P8

Table 2.2: this figure displays 3 by 3 windows for minutiae searching

Figure 2.6: Original Image

Figure 2.7: Minutiae Extracted using Cross Number concept

2.5 Conclusion

The reliability of any automatic fingerprint system strongly relies on the precision obtained in the minutia extraction process. A number of factors are detrimental to the correct location of minutia. Among them, poor image quality is the most serious one. In this project, we have combined many methods to build a minutia extractor and a minutia matcher. The following concepts have been used- segmentation using Morphological operations,minutia marking by specially considering the triple branch counting, minutia unification by decomposing a branch into three terminations and matching in the unified coordinate system after a step transformation in order to increase the precision of the minutia localization process and elimination of spurious minutia with higher accuracy. The proposed alignment-based elastic matching algorithm is capable of finding the correspondences between minutiae without resorting to exhaustive research.

Chapter 3: Speech Recognition

3.1 Introduction

Designing a machine that converse with human, particularly responding properly to spoken language has intrigued engineers and scientists for centuries. Speech Recognition System(SRS) is also known as Automatic Speech Recognition (ASR) or computer speech recognition which is the process of converting a speech signal to a sequence of words by means of an algorithm implemented as a computer program. It has the potential of being an important mode of interaction between humans and computers [1]. Today speech technology enabled applications are commercially available for a limited but interesting range of tasks. Veryuseful and valuable services are provided by these technology enabled machines, by responding correctly and reliably to human voices. In order to bring us closer to the "Holy Grail" of machines that recognize and understand fluently spoken speech, many important scientific and Technological advances have been took place, but still we are far from having a machine that mimics human behavior. Speech recognition technology has become a topic of great interest to general population, through many block buster movies of 1960's and 1970's [2].

3.2 General Steps For Speech Recognition Systems

In speech recognition system an unknown speech signal is transformed into sequence of feature vectors by different speech processing techniques. It converts feature vector to phoneme lattice by applying an algorithms [4]. A recognition module transforms the phoneme lattice into a word lattice by lexicon and then grammar is applied to word lattice to recognize the specific words or text. Figure 1 shows the information for general steps in speech recognition system (SRS).

The speech recognition process is divided into several steps.

- ***Step 1:***In this step speech signal is divided intoequally spaced blocks to get signal characteristics such as, total energy, zero crossing strength across various frequency ranges etc. By using these characteristics feature vectors combine each block with the phoneme to produce a string of phonemes.
- ***Step 2:***In this step spectrum analysis is appliedon each block by using linear predictive coding technique, fast Fourier transform(FFT) and bank of frequency filters.
- ***Step 3:***In this step decision process is performedon each block. Each phoneme has distinguished features which narrow the field.
- ***Step 4:*** This step is used to enhance theperformance of decision process to get high degree of success using different algorithms. For each word of vocabulary an algorithm is constructed and then string of phonemes is compared against each algorithm.

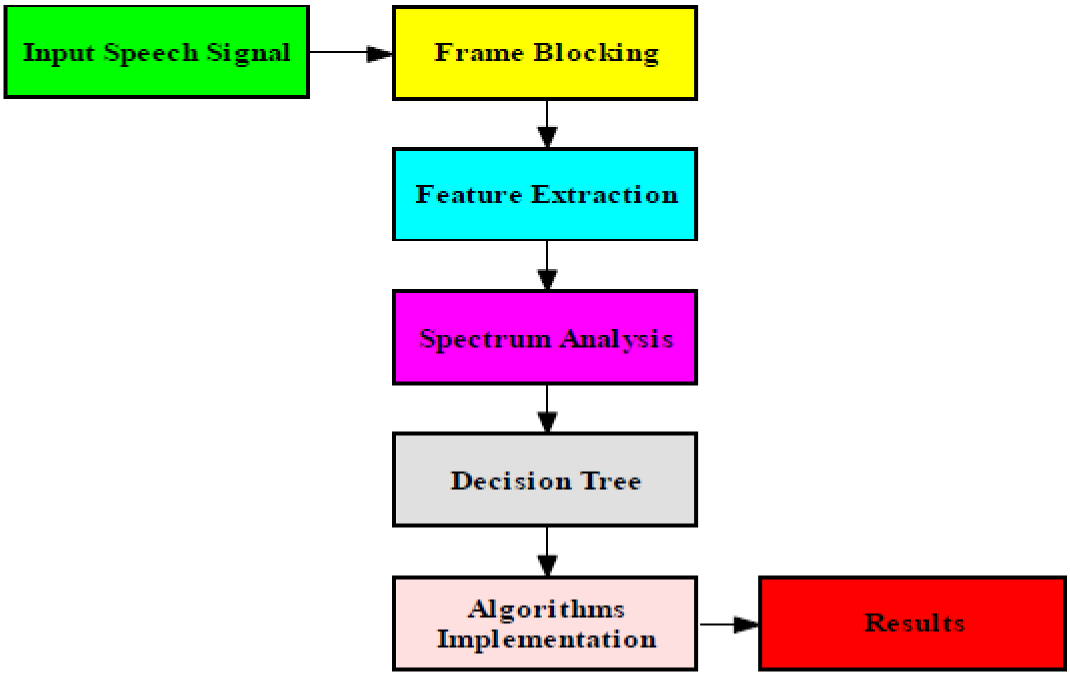

Figure 3.1: General steps for Speech Recognition System

3.3 Techniques In Speech Recognition Systems

The goal of speech recognition is for a machine to be able to "hear," understand," and "act upon" spoken information. The earliest speech recognition systems were first attempted in the early 1950s at Bell Laboratories. Davis, Biddulph and Balashek developed an isolated digit recognition system for a single speaker. The speaker recognition system may be viewed as working in a four stages.

3.3.1 Speech Analysis

In speech analysis technique Speech data contains different types of information that shows a speaker identity. This includes speaker specific information due to vocal tract, excitation source and behavior feature. The physical structure and dimension of vocal tract as well as excitation source are unique for each speaker. The speech analysis deals with stages with suitable frame size for segmenting speech signal for further analysis and extracting [5]. The speech analysis is done with following three techniques.

- **Segmentation Analysis:**In this case, speech is analyzed using the frame size and shift in the range of 10-30 ms toextract speaker information. Studies have been made in using segmented analysis to extract vocal tract information of speaker recognition.
- **Sub-segmental Analysis:**Speech analyzed using the frame size and shift in range 3-5 ms is known as Sub segmentalanalysis. This technique is used mainly to analyze and extract the characteristic of the excitation state. [6]. The excitation source information is relatively fast varying compared to vocal tract

information, so small frame size and shift are required to best capture the speaker-specific information [7].

- **Supra-segmental Analysis:** In this case, speech is analyzed by using the frame size and shift of 100-300 ms to extractspeaker information mainly due to behavioral tract and here speech is analyzed using the frame size. This technique is used mainly to analyze and characteristic due to behavior character of the speaker. These include word duration, intonation, speaker rate, accent etc.

3.3.2 Speech Feature Extraction Techniques

Feature Extraction is the most important part of speech recognition since it plays an important role to separate one speech from other. The utterance can be extracted from a wide range of feature extraction techniques proposed and successfully exploited for speech recognition task. But extracted feature should meet some criteria while dealing with the speech signal such as:

- Easy to measure extracted speech features
- It should not be susceptible to mimicry
- It should show little fluctuation from one speaking environment to another
- It should be stable over time
- It should occur frequently and naturally in speech

3.3.3 Modeling Technique

Speech recognition is a vast and emerging field nowadays. Every speech signal has different characteristics depend on utterances, to achieve this task of recognition different approaches are used. Each technique has its own benefits depends upon the scenario. Some approaches include template based approach, neural network based approach, statistical based approach, Hidden Markov Model (HMM) based speech recognition etc. Most famous of all is the HMM approach because It is easy, simple and reliable, it can be automatically trained and feasible to use. Some approaches are listed below.

- **Acoustic Phonetic Recognition :**Acoustic phonetic recognition performs the function at phoneme level. It exist distinctive, finite phonemes which are characterized by a set of acoustic properties that occur in a speech signal [12]. English language includes forty different phonemes and doesn't depend on the vocabulary. It is the earliest approach of SRS to recognize speech by providing labels to the speech. This approach includes highly variable phonetic units, the variability in these unit are straight forward which are easily learned by machine [13].the information about Acoustic phonetic recognition block diagram. This approach is divided into three steps
  - Feature Extraction
  - Segmentation and Labeling
  - Word-level recognition
- **Pattern Recognition Approach:**By using mathematical framework this approach formulates speech pattern representation from formal training algorithms by set of labeled training samples via formal training algorithms [14]. This approach involves two major steps.
  - Pattern training
  - Pattern compression

- **Template Based Approach** In this approach by using a collection of speech pattern, dictionary of words are created which is stored as a reference, after that matching of unknown speech is done with the reference template [15]. Then selection of the best matching template is done. It is also a modified form of pattern recognition.It is the oldest and least effective method but dominating approach in 1950's to 1960's. It is still successful for small dictionaries and specifically for isolated word recognition. It is the primary technology for verification of speaker.shows the template based approach block diagram.
- **Vector Quantization Approach:**For efficient reduction of data of automatic SRS this approach is used, because in automatic SRS transmission rate is not a major issue. The efficiency of this approach lies on using best code book which yields the lowest distance measure [16].
- **Dynamic Time Wrapping:**This approach finds similarities between two different sequences which vary with time and speech [17].
- **Statistical Based Approach** Uncertainty or incompleteness occurs in speech recognition due to many reasons like mixing of sound, speaker variability, confusable words etc.
- **Hidden Markov Model :**It is widely use stochastic approach today because in this approach speech is generating from number of states for each HMM model. In HMM we mixture multi vibrate Gaussian distribution, probabilistic mean, variance and mixture weight for speech [19].
- **Artificial Neural Network Approach** This approach is designed for complicated tasks but it is not as efficient as HMM in the case of large vocabularies. Phoneme recognition is the general approach of neural networks.
-

3.3.4 Matching Techniques

Speech-recognition engines match a detected word to a known word using one of the following techniques (Svendsen et al., 1989)[21].

- **Whole-word matching** The engine compares the incoming digital-audio signal against a prerecorded template of the word. This technique takes much less processing than sub-word matching, but it requires that the user (or someone) prerecord every word that will be recognized - sometimes several hundred thousand words. Whole-word templates also require large amounts of storage (between 50 and 512 bytes per word) and are practical only if the recognition vocabulary is known when the application is developed [21].
- **Sub-word matching** The engine looks for sub words usually phonemes and then performs further pattern recognition on those. This technique takes more processing than whole word matching, but it requires much less storage (between 5 and 20 bytes per word). In addition, the pronunciation of theword can be guessed from English text without requiring the user to speak the word beforehand. Discuss that research in the area of automatic speech recognition had been pursued for the last three decades [21].

3.4 Comparative Study Of Speech Recognition System Approaches

The comparative study of speech recognition system approaches have given in the table 3 with their advantages and disadvantages:

Sr. No.	SRS Techniques	Advantages	Disadvantages
1	Acoustic Phonetic Recognition	1. It reduces Processing Time forconnected words	1. Not widely used in commercial application due to large time execution of each isolated word
2	Dynamic Time wrapping	1. Continuity is less important because it can match sequence with missing information. 2. Reliable time alignment between reference and test pattern.	1. It matches between two given sequence with certain restrictions. 2. It requires maximum time for complex computational work. 3. Limited number of templates.
3	Pattern Recognition Approach	1. It recognize pattern quickly, easily and automatically because word to word matching will occur.	1. It is useful for word to word matching. 2. Template is the main problem. 3. Slow process. 4. It doesn't recognize speech if new variation of pattern occur.
4	Vector Quantization Approach	1. It is useful for efficient data reduction.	1. It is text depended because need codebook for matching.
5	Template Base Approach	1.It is better for discrete words 2. Less error occur due to segmentation and classification of small variable units.	1. Expensive due to large vocabulary size in each word has reference templates for it. 2. Template matching and preparation requires more time. 3. It difficult to recognize similar templates.
6	Artificial Neural Network Approach	1. It can solve complex computational task effectively within less time. 2. It has ability to automatically train the data and taught the system changing from initial training model without error. 3. It can handle noisy, low quality data efficiently and require minimum training data vocabulary.	1. It gives inefficient result for large vocabulary 2. It is expensive because for training it requires much iteration over large amount of training data. 3. Full nature of neural network is not fully understood still. 4. It requires more training time 5. More error variation occurs due to complex architecture of neural networks.

Table 3: Comparative Study of Speech Recognition Modeling Technique

3.5 Conclusion

Speech Recognition System is growing day by day and has unlimited applications. Thestudy has shown the overview of the speech recognition process, its basic model, and applications. In this studytotal seven different approaches which are widely used for SRS have been discussed and after comparative study of these approaches it is concluded that Hidden markov method(HMM) is best suitable approach for a SRS because it efficient, robust, and reduces time and complexity.

Chapter 4: Classification and Extraction of parameters

4.1 Introduction

In face, voice and fingerprint recognition, we need to extract the features of each one of them and compare them with the features in the databases using custom algorithms. In this chapter, we'll show the algorithms we'll use to extract and classify properties, how they work mathematically, and how to calculate Euclidean distances.

4.2 face recognition

4.2.1 Principal Component Analysis

The idea of using principal components to represent human faces was developed by Sirovich and Kirby [44] in 1987, and was used by Turk and Pentland [10] in 1991 for facial detection and recognition .The eigenface approach is considered by many to be the first practical facial recognition technology, and has served as the basis for many facial recognition technology products. Since its initial development and dissemination, there have been many extensions of the original method and many new methods.

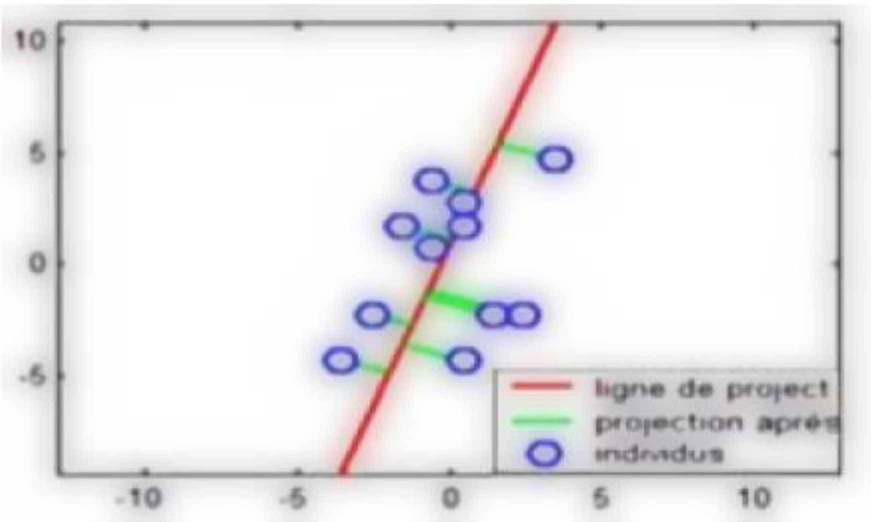

Figure 3.1: Example of projection following PCA .

4.2.1.1 PCA Algorithm

The steps of the eigenfaces calculation are :

- **Prepare the data**

  A 2-D facial image can be represented as 1-D vector by concatenating each row (or column) into a long thin vector. Let's suppose we have M vectors of size N (= rows of image × columns of image) representing a set of sampled images. Then the training set becomes: $\Gamma 1$, 2, $\Gamma 3$..... ΓM.

- **Subtract the mean**

  The average matrix Ψ has to be calculated, then subtracted from the original faces (Γi) and the result stored in the variable Φi:

$$\Psi = \frac{1}{M}\sum_{n=1}^{M}\Gamma_n \quad (3.1)$$

$$\Phi_i = \Gamma_i - \Psi \quad (3.2)$$

- **Calculate the co-variance matrix**

  In the next step the covariance matrix A is calculated according to:

$$A = \Phi^T\Phi \quad (3.3)$$

- **Calculate the eigenvectors and eigenvalues**

  In this step, the eigenvectors (eigenvectors) Xi and the corresponding eigenvalues λi should be calculated.

- **Calculate eigenfaces**

$$[\Phi]X_i = f_i \quad (3.4)$$

where x_i are eigenvectors and f_i are eigenfaces

- **Classifying the faces**

  The new image is transformed into its eigenface components. The resulting weights form the weight vector Ω_{new}^{T} :

$$\omega\kappa = \Omega k T \Gamma_{new} - \Psi \qquad \text{and} \qquad k = 1.2.3....M \tag{3.5}$$

$$\Omega_{new}^{T} = [\omega_1 \omega_2 \omega_3 .... \omega_M] \tag{3.6}$$

  The Euclidean distance [13-19] between two weight vectors $d(\Omega_i, \Omega_j)$ provides a measure of similarity between the corresponding images i & j. If the Euclidean distance between Γ_{new} and other faces exceeds some threshold value θ, one can assume that Γ_{new} is not a face at all, $d(\Omega_i, \Omega_j)$ also allows one to construct "clusters" of faces such that similar faces are assigned to one cluster [45]

4.2.1.2 Distance Measures

When comparing two characteristic vectors from the Biometric Feature Extraction Module, either a similarity (resemblance) or a distance (divergence) measurement can be performed. The first category of distances consists of Euclidean distances and are defined from the distance of Minkowski of order p in an Euclidean space RN (N determining the dimension of the Euclidean space) [46].

Consider two vectors $X = (x_1, x_2, ..., x_n)$ et $Y = (y_1, y_2, ..., y_n)$, the distance of Minkowski of order p noted L_p is defined by:

$$L_p = \left(\sum_{i=1}^{n} |x_i - y_i|^p \right)^{1/p} \tag{3.7}$$

- **Euclidean distances:**

- Distance City Block (L1) Pour p = 1, on obtient la distance City-Block (ou distance de Manhattan) :

(3.8)

$$L_1(x, y) = \sum_{i=1}^{n} |x_i - y_i|$$

- Euclidean distances (L2) For p = 2, the Euclidean distance is obtained:

(3.9)

$$L_2(x, y) = \sqrt{\sum_{i=1}^{n} |x_i - y_i|^2}$$

The objects can then appear in very different ways depending on the distance measurement chosen.

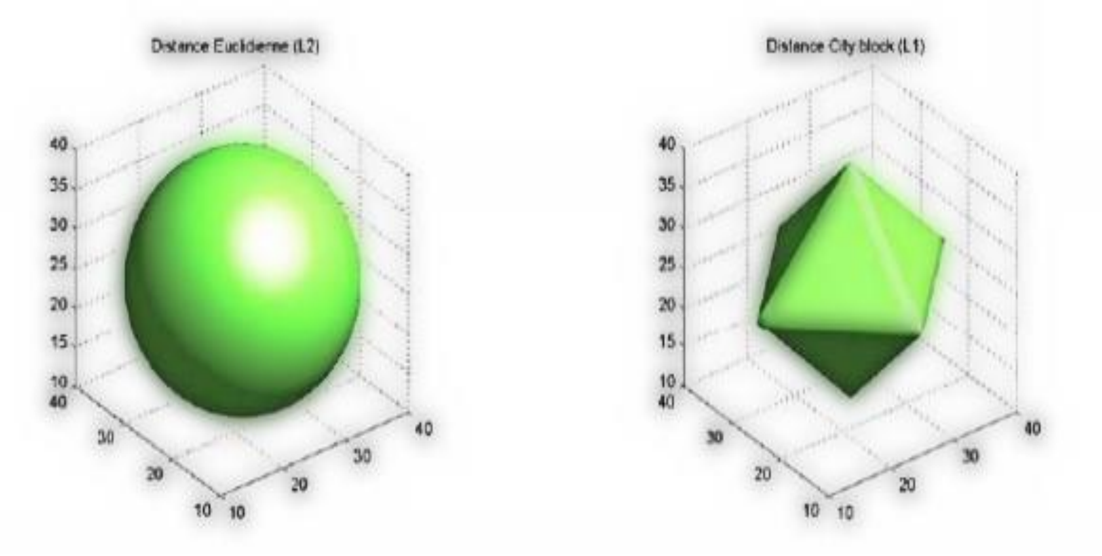

Figure 3.2: Representation of a sphere with the Euclidean distance and the City-Block distance [46]

4.2.2 Discrete Cosine Transform

The use of DCT in face recognition field is one of the most recent methods. It uses the discrete transformation into a cosine to eliminate the redundancies in an image and extract from them the most significant elements (i.e. coefficients) in order to use them for recognition [47].

4.2.2.1 Discrete Cosine Transform Encoding

- **The One-Dimensional DCT**

The general equation for a 1D (N data items) DCT is defined as follows:

$$F(u) = \alpha(u) \sum_{i=0}^{N-1} A(x) * \cos\left(\frac{u(2x+1)\Pi}{2N}\right) \quad (3.10)$$

for u = 0,1,2,....,N-1 . Similarly, the inverse transformation is defined as

$$A(x) = \sum_{u=0}^{N-1} \alpha(u) F(u) \cos\left(\frac{u(2x+1)\Pi}{2N}\right) \quad (3.11)$$

for x = 0,1,2,....,N-1 . In both equations Equation 3.10 and Equation 3.11 $\alpha(u)$ is defined as

$$\alpha(u) = \begin{cases} \sqrt{\frac{1}{N}} & \text{for } u = 0 \\ \sqrt{\frac{2}{N}} & \text{for } u \neq 0 \end{cases} \quad (3.12)$$

It is clear from Equation 3.10 that for u = 0

$$F(u=0) = \sqrt{\frac{1}{N}} \sum_{x=0}^{N-1} A(x) \quad (3.13)$$

the first transform coefficient is the average value of the sample sequence. In literature, this value is referred to as the DC Coefficient. All other transform coefficients are called the AC Coefficients 4.

The plot of $\sum_{x=0}^{N-1} \cos\left(\frac{u(2x+1)\Pi}{2N}\right)$ for N = 8 and varying values of u is shown in Figure Figure 3.3. In accordance with our previous observation, the first the topleft waveform (u = 0) renders a constant (DC) value, whereas, all other waveforms (u = 1,2,...,7) give waveforms at progressively increasing frequencies [48].

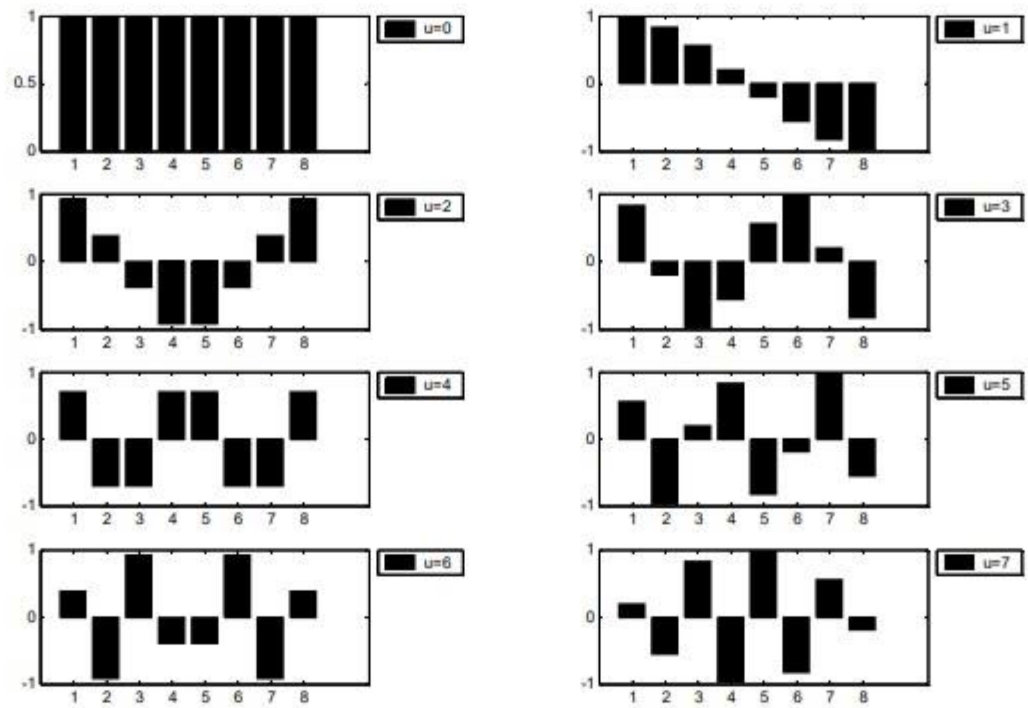

Figure 3.3: One dimensional cosine basis function (N=8).

- **The Two-Dimensional DCT**
  The objective is study the efficacy of DCT on images. This necessitates the extension of ideas presented in the last section to a two-dimensional space. The 2-D DCT is a direct extension of the 1-D case and is given by

$$F(u,v) = \alpha(u)\alpha(v) \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} A(x,y) \cos\left(\frac{u(2x+1)\Pi}{2N}\right) \cos\left(\frac{v(2y+1)\Pi}{2N}\right) \quad (3.14)$$

for u ,v = 0,1,2,...,N-1 and _(u) , _(v) are defined in Equation 3.12. The inverse transform is defined as

$$A(x,y) = \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} \alpha(u)\alpha(v) F(u,v) \cos\left(\frac{u(2x+1)\Pi}{2N}\right) \cos\left(\frac{v(2y+1)\Pi}{2N}\right) \quad (3.15)$$

for x,y = 0,1,2,...,N-1

The 2-D basis functions can be generated by multiplying the horizontally oriented 1-D basis functions (shown in Figure 3.4) with vertically oriented set of the same functions [48]. The basis functions for N = 8 are shown in. Again, it can be noted that the basis functions exhibit a progressive increase in frequency both in the vertical and horizontal direction. The top left basis function of results from multiplication of the DC component in Figure 3 with its transpose. Hence, this function assumes a constant value and is referred to as the DC coefficient.

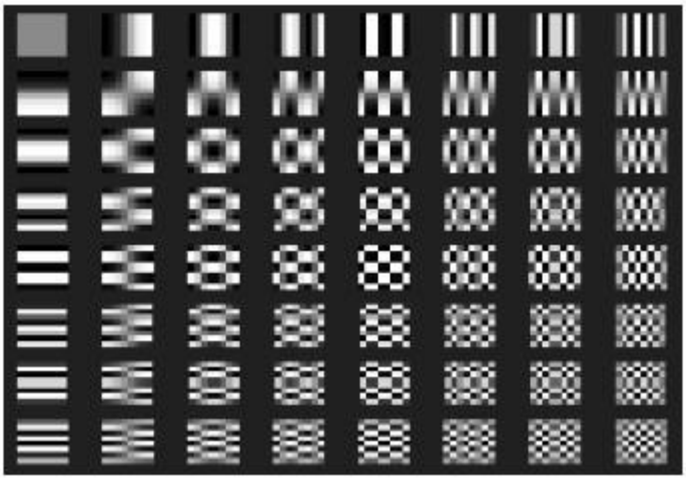

Figure 3.4: Two dimensional DCT basis functions (N = 8). Neutral gray represents zero,white represents positive amplitudes, and black represents negative amplitude [48].

4.2.2.2 Properties of DCT

In the previous section, it established a mathematical basis for DCT. However, the intuitive view of the image processing application was not presented. We outline (with examples) some of the characteristics of The DCT that are particularly valuable in image processing applications.

- **Decorrelation**

The principle advantage of image transformation is the removal of redundancy between neighboring pixels. This leads to uncorrelated transform coefficients which can be encoded independently. The normalized autocorrelation of the images before and after DCT is shown in Figure Figure 3.5. Clearly, the amplitude of the autocorrelation after the DCT operation is very small at all lags. Hence, it can be inferred that DCT exhibits excellent decorrelation properties.

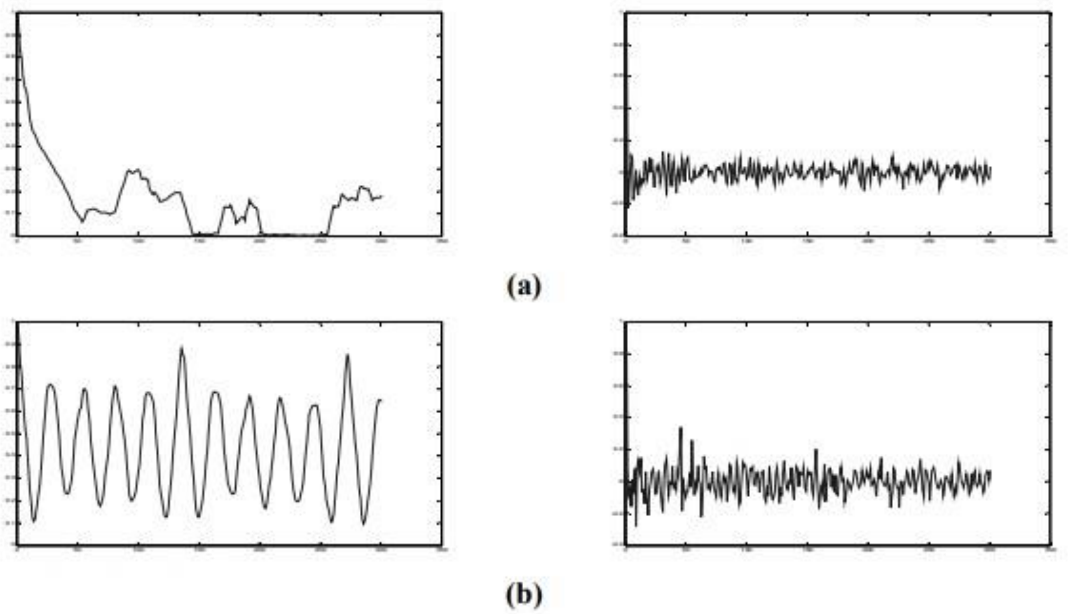

Figure 3.5: (a) Normalized autocorrelation of uncorrelated image before and after DCT-
(b) Normalized autocorrelation of correlated image before and after DCT.

- **Energy Compaction**

  Efficacy of a transformation scheme can be directly gauged by its ability to pack input data into as few coefficients as possible. This allows the quantizer to discard coefficients with relatively small amplitudes without introducing visual distortion in the reconstructed image. DCT exhibits excellent energy compaction for highly correlated images.

(a)

(b)

(c)

Figure 3.6: (a) Saturn and its DCT-(b)Baboon and its DCT-(c)Child and its DCT

- **Separability**

The DCT transform equation Equation 3.14 can be expressed as

$$F(u,v) = \alpha(u)\alpha(v)\sum_{x=0}^{N-1}\cos\left(\frac{u(2x+1)\Pi}{2N}\right)\sum_{y=0}^{N-1}A(x,y)\cos\left(\frac{v(2y+1)\Pi}{2N}\right) \quad (3.16)$$

for u ,v =0,1,2,...,N-1.

This property, known as separability, has the principle advantage that C(u, v) can be computed in two steps by successive 1-D operations on rows and columns of an image. This idea is graphically illustrated in Figure Figure 3.7. The arguments presented can be identically applied for the inverse DCT computation [49].

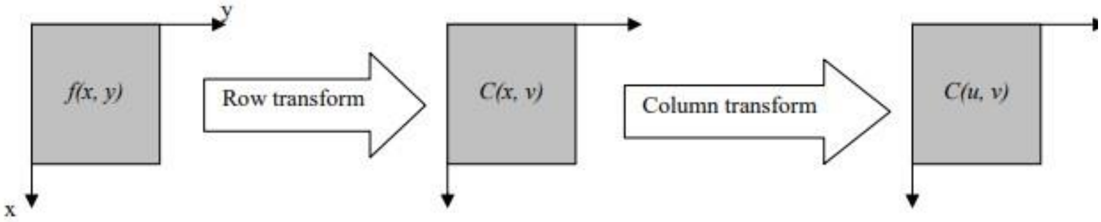

Figure 3.7: Computation of 2-D DCT using separability property

- **Symmetry**

  Another look at the row and column operations in Equation page 50 reveals that these operations are functionally identical. Such a transformation is called a symmetric transformation. A separable and symmetric transform can be expressed in the form [50].

$$T = kfk \tag{3.17}$$

where k is an N*N symmetric transformation matrix with entries a(i,j) , given by:

$$a(i,j) = \alpha(j) \sum_{j=0}^{N-1} \cos\left(\frac{i(2x+1)\Pi}{2N}\right) \tag{3.18}$$

and f is the N*N image matrix. This is an extremely useful property since it implies that the transformation matrix can be pre-computed offline and then applied to the image thereby providing orders of magnitude improvement in computation efficiency.

- **Orthogonality**

  In order to extend ideas presented in the preceding section, let us denote the inverse transformation of Equation 3.17 as :

$$T = k^{-1} f k^{-1} \tag{3.19}$$

As discussed previously, DCT basis functions are orthogonal. Thus, the inverse transformation matrix of k is equal to its transpose i.e.

$k^{-1}=k^{T}$..Therefore, and in addition to its decorrelation characteristics, this property renders some reduction in the pre-computation complexity.

4.2.2.3 DCT versus DFT and KLT

- DCT = Discrete Cosine Transform
- DFT = Discrete Fourier Transform
- KLT = Kamunen-Loeve Transform

At this point it is important to mention the superiority of DCT over other image transforms. More specifically, we compare DCT with two linear transforms: 1) The Karenina-Loeve Transform (KLT). 2) Discrete Fourier Transform (DFT).

The KLT is a linear transform where the basis functions are taken from the statistical properties of the image data, and can thus be adaptive. It is optimal in the sense of energy compaction, i.e., it places as much energy as possible in as few coefficients as possible. However, the KLT transformation kernel is generally not separable, and thus the full matrix multiplication must be performed. In other words, KLT is data dependent and, therefore, without a fast (FFT-like) pre-computation transform. Derivation of the respective basis for each image sub-block requires unreasonable computational resources. Although, some fast KLT algorithms have been suggested, nevertheless the overall complexity of KLT is significantly higher than the respective DCT and DFT algorithms.

4.3 voice recognition

4.3.1 Linear Predictive Coding (LPC)

Method A signal processing is an activity to extract a signal information. Linear Predictive Coding (LPC) is a powerful speech analysis technique and facilitating a features extraction which has a good quality and efficient result for computing. In 1978, LPC uses to make a speech synthesis. LPC doing an analysis with predicting a formant decided a formant from signal called inverse filtering, then estimated an intensity and frequency from residue speech signal. Because speech signal has many variations depending on a time, the estimation will do to cut a signal called frame. Procedure to get an LPC coefficient shown in Figure 1.

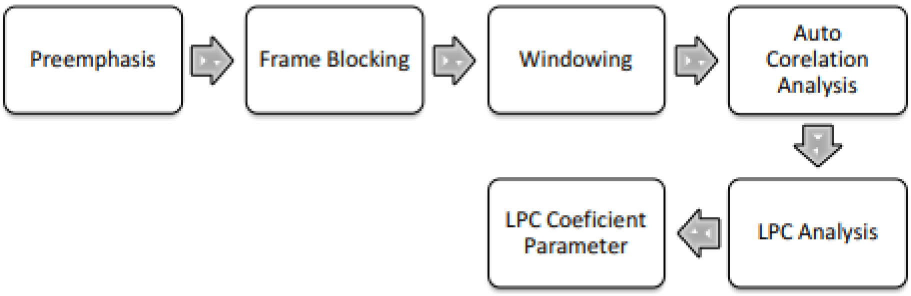

Figure 1. LPC Method Diagram

4.3.1.1 Preemphasis.

On processing of speech signal, preemphasis filter needed after sampling process. The filtering purpose is to get a smooth spectral shape of the speech signal. A spectral which have a high value for the low-frequency field and decrease for field frequency higher than 2000 Hz. Preemphasis filter based on the relation of input/output on time domain which is shown by the equation (1),

$$Y(n) = x(n) - ax(n-1) \quad (1)$$

a is a constant of preemphasis filter, ordinary have $0.9 < a < 1.0$.

4.3.1.2 Frame Blocking.

Frame Blocking. On this process, segmented of speech signal become some frame which overlaps. So that no signal is lost (deletion).

4.3.1.3 Windowing.

Analog signal which converts become digital signal read frame by frame and each frame is windowing with the certain window function. This windowing process purpose to minimize discontinue signal from initial to end of each frame. If window as $w(n)$, $0 \leq n \leq N - 1$, when N is total of sample of each frame, thus result of windowing is a signal:

$$Y1(n)=x1(n)w(n),0<n<N-1 \quad (2)$$

Where $w(n)$ usually use window Hamming with the form:

$$w(n) = 0.54 - 0.46.cos\left(\frac{2\pi n}{N-1}\right), 0 \leq n \leq N - 1 \,. \quad (3)$$

4.3.1.4 Autocorellation Analysis.

The next step is autocorrelation analysis toward each frame result by windowing $y1(n)$ with equation (4),

$$r_1(m) = \sum_{n=0}^{n-1-m} y_1(n)y_1(n+m), \; m = 0,1,2..,p \,, \quad (4)$$

Where *p* is ordered from LPC. LPC order which usually used is between 8 until 16.

4.3.1.5 LPC Analysis.

This step will convert each frame from p+1 autocorrelation become compilation of "LPC parameter" αm = α Pm D for m = 1, 2... p. This compilation becomes LPC coefficient or become other LPC transformation. The formal method to change autocorrelation coefficient become parameter LPC compilation called Durbin method, the form as:

$$E^0 = r(0) \quad (5)$$

$$k_m = \frac{\left\{r(m) - \sum_{j=1}^{m-1} \alpha_j^{(m-1)} r(|m-j|)\right\}}{E^{(m-1)}}, \; 1 \leq m \leq p \,, \quad (6)$$

$$\alpha_m^m = k_m \,, \quad (7)$$

$$\alpha_j^m = \alpha_j^{(m-1)} - k_m \alpha_{m-j}^{(m-1)}, \; 1 \leq j \leq m-1 \,, \quad (8)$$

$$E^m = (1 - k_m^2)E^{(m-1)} \,. \quad (9)$$

With $r(0)$ is a result of autocorrelation, $E(m)$ is an error, k_m is rebound of the coefficient, α_{mj} is prediction coefficient for $1 \leq j \leq m$.

4.3.1.6 LPC Parameter Conversion to LPC Coefficient.

LPC coefficient parameter am conversion to cepstral coefficient cm to get the best performance and endure to noise, as the questions:

$$c_m = a_m + \sum_{k=1}^{m-1} (k/m) c_k a_{m-k}\ , \quad 1 \le m \le p \tag{10}$$

$$c_m = \sum_{k=1}^{m-1} (k/m) c_k a_{m-k}\ , \quad m > p \tag{11}$$

Conclusion

The conception of a monitoring system based on the recognition of face, voice, and fingerprint aims to create a comprehensive and secure solution for identification and monitoring purposes. This integrated system combines three biometric modalities: face recognition, voice recognition, and fingerprint recognition.

Face recognition involves capturing and analyzing facial features to create unique biometric templates for each individual. Advanced algorithms, such as Principal Component Analysis (PCA), are utilized to extract key facial characteristics and match them against a database for identification.

Voice recognition relies on recording and analyzing individuals' voice patterns, including pitch, tone, and pronunciation. A reference model is created based on these unique voice characteristics, enabling accurate verification by comparing recorded voice samples with the reference models.

Fingerprint recognition, known for its high accuracy and reliability, captures and analyzes the unique patterns and ridges on an individual's fingertip. Personalized fingerprint templates are created for each person, allowing for real-time authentication and matching of fingerprints.

By integrating these three modalities, the monitoring system ensures a robust and comprehensive approach to identification and monitoring. It offers enhanced accuracy, reliability, and security compared to traditional systems that rely on a single identification method. This integrated system finds applications in various domains, including security systems, access control, attendance management, and law enforcement, providing both convenience and efficiency in managing and monitoring individuals in a given environment.

ملخص

يهدف مفهوم نظام المراقبة القائم على التعرف على الوجه والصوت وبصمة الإصبع إلى إنشاء حل شامل وآمن لأغراض تحديد الهوية والمراقبة. يجمع هذا النظام المتكامل بين ثلاث طرق بيومترية: التعرف على الوجوه والتعرف على الصوت والتعرف على بصمات الأصابع.

يتضمن التعرف على الوجه التقاط ميزات الوجه وتحليلها لإنشاء قوالب بيومترية فريدة لكل فرد. تُستخدم الخوارزميات المتقدمة ، مثل تحليل المكونات الرئيسية (PCA) ، لاستخراج خصائص الوجه الرئيسية ومطابقتها مع قاعدة بيانات لتحديد الهوية.

يعتمد التعرف على الصوت على تسجيل أنماط صوت الأفراد وتحليلها ، بما في ذلك درجة الصوت والنبرة والنطق. يتم إنشاء نموذج مرجعي بناءً على هذه الخصائص الصوتية الفريدة ، مما يتيح التحقق الدقيق من خلال مقارنة عينات الصوت المسجلة بالنماذج المرجعية.

التعرف على بصمات الأصابع ، المعروف بدقته العالية وموثوقيته ، يلتقط ويحلل الأنماط والتلال الفريدة على أطراف أصابع الفرد. يتم إنشاء قوالب بصمات الأصابع المخصصة لكل شخص ، مما يسمح بالمصادقة في الوقت الفعلي ومطابقة بصمات الأصابع.

من خلال دمج هذه الأساليب الثلاثة ، يضمن نظام المراقبة اتباع نهج قوي وشامل لتحديد الهوية والمراقبة. إنه يوفر دقة وموثوقية وأمان محسنين مقارنة بالأنظمة التقليدية التي تعتمد على طريقة تعريف واحدة. يجد هذا النظام المتكامل تطبيقات في مجالات مختلفة ، بما في ذلك أنظمة الأمان والتحكم في الوصول وإدارة الحضور وإنفاذ القانون ، مما يوفر الراحة والكفاءة في إدارة ومراقبة الأفراد في بيئة معينة.