Republic of Iraq Ministry of Higher Education and Scientific Research Babylon University College of Engineering

Employing Efficient Techniques of DWT/PCA for Face Recognition

A Project

Submitted to the College of Engineering / University

Babylon in partial fulfillment of the requirements for the degree of

Higher Diploma in Engineering/Electrical Engineering/Electronics and Communications

Prepared by

Hawraa Laith Falah Hassan

Supervised by

Dr. Ahmed Q. Aldhahab

2022 A.D **1443 A.H**

بسم الله الرحمن الرحيم

وما توفيقي إلا بالله عليه توكلت

وإليه أنيب

صدق الله العلي العظيم

سورة هود، الآية (٨٨)

Dedication

This thesis work is dedicated to my Family, those have been a constant source of support and encouragement during the challenges of graduate school and life. I am truly thankful for having you in my life. This work is also dedicated to my parents, who have always loved me unconditionally and whose good examples have taught me to work hard for the things that I aspire to achieve.

Hawraa Laith Falah

2022

Acknowledgements

In the name of Allah, the Most Gracious the Most Merciful
Alhamdulillah, all praises and thanks be to Allah for blessing me to accomplish this work despite all the hardships.

I would like to express my deepest sincere thanks of gratitude to my supervisor whom I'm very proud to work with Dr.Ahmed Q. Aldhahab for his great ideas and valuable suggestion in accomplishing this project, and for his continuous Scientific, moral support to finalize this work.

Also, I wish to convey my heartfelt thanks to my family and my husband and my friends for their elevating inspiration and kind encouragement along the period of my project.

Last but not least I spread my deepest thanks to all the teachers who helped me and I would like to respect the immortal favour they offered me.

Hawraa Laith Falah

2022

Abstract

A new approach for facial recognition employing Two-Dimensional Discrete Wavelet Transform (2D DWT) and 2D Principal Component analysis (2D PCA) is proposed. Preprocessing, Feature Extraction, and Classification are the three main phases in this system. A cropping technique and selecting an appropriate dimensions are employed in the preprocessing step. In the feature extraction, 2-levels of 2D DWT decompositions are applied to the processed images for dimensionality reduction and feature extraction. The resultant 2D DWT features are further compacted using 2D PCA. The optimization techniques, Genetic and Ant Colony Optimization (ACO), are utilized to improve the system accuracy. The proposed algorithm is evaluated using four databases, namely, ORL, YALE, FERET, and Georgia Tech. that have different facial variations, such as facial expressions, illuminations, rotation, etc. Then, the results are analyzed using K-fold Cross Validation (CV). The results show that the proposed approach improves the recognition rates and reduces the storage requirements compared with existing methods.

Table of Contents

List of Figures

List of Tables

List of abbreviation

Abbreviation	**Definition**
DNA	Deoxyribonucleic Acid
GIF	Graphics Interchange Format
DARPA	Defense Advanced Research Projects Agency
NIST	National Institute of Standards and Technology's
JPG	Joint Photographic Group
2D	Two Dimension
3D	Three Dimension
DFT	Discrete Fourier Transform
DTFT	Discrete Time Fourier Transform
FFT	Finite Fourier Transform
WFT	Windowed Fourier Transform
STFT	Short-Time Fourier Transform
CWT	Continuous Wavelet Transform
ICWT	Inverse Continuous Wavelet Transformation

DWT	Discrete Wavelet Transform
2D DWT	Two Dimension Discrete Wavelet Transform
FWT	Fast Wavelet Transform
FFT	Fast Fourier Transform
PCA	Principal component analysis
2D PCA	Two Dimension Principal component analysis
ACO	Ant Colony Optimization
GA	Genetic Algorithm
CV	Cross Validation

CHAPTER ONE

INTRODUCTION

CHAPTER ONE

INTRODUCTION

1.1 Introduction

Biometrics is a term that describes the direct relationship with human characteristics. This term is derived from “bio” and “metrics” which means life and measure respectively.

There are three kinds of biometrics which are (Recognition [1-3], identification [4] and verification [5]). Distinguishable and measurable characteristics are biometric identifiers that are used to describe and name individuals.

“Who are you” is one of the key questions in identification systems that biometric systems compete to answer accurately. Identification using facial recognition systems can depend on a number of parameters such as the distance between the eyes, the shape of the nose ... etc. Distinctive identifiers of people can be abbreviated without limiting them to the following: “fingerprints, hand geometry, earlobe geometry, retina and iris patterns, voice waves, DNA, signatures, palm print, face, etc.”.

There are different fields in biometric system such as “sensor design, computer vision, image processing, pattern recognition, machine learning, signal processing, and information fusion”. Traditional systems for identifying people, such as personal identities, passwords or keys, are less accurate and more dangerous compared to systems that rely on biometric identifiers for individuals. Biometric systems have a natural protection against traditional risks such as theft and counterfeiting [6,7].

1.2 Categorization

It can be said that facial recognition system can work in one or both of the following two modes. The first is face verification and authentication while the second mode involves facial recognition and verification through one-to-one matching. This system has many applications, such as employee or student attendance systems in private and government companies.

Facial recognition involves comparing the face to be tested with a large number of faces stored in the database [8]. The purpose of this comparison is to identify the faces that belong to the database that most closely match the face whose identity is to be revealed. On the other hand, there are other applications that not only require the most congruent face, but also require all faces whose degree of congruence exceeds a certain limit [9]. Such it is required in the case of searching for a specific person within a video or a public camera.

Face recognition systems depend on a number of influential parameters such as intensity of illumination [10], age, facial expression,...etc. [11-13]. Depending on the aforementioned parameters, the applications of facial recognition systems can be divided into two main categories [14], depending on the nature of the user: the cooperative user and the non-cooperative user.

In cases that involve a cooperating person, the process of facial recognition is easier, because the person is standing in front of the camera and with a good posture and direction. A common example of this type of facial recognition systems is attendance systems and access to smart devices based on the face [15].

In the case of recognizing the faces of uncooperative people [16,17], the process is more difficult than the first case for various reasons that can be summarized without being limited to the following:

1. The distance between the camera and the person to be identified
2. Difficulty getting a good front picture

A common example of this situation, a person or group of people through a surveillance camera or video

There are other cases that represent a real reality that is not considered one of the aforementioned types. In these cases, the person is cooperating, but he/she is unable to fulfill all the conditions of the system for the cooperating people, such as the inability to approach the camera to a short distance or the difficulty of placing the face in the required position ... etc.

In all cases, the intensity and direction of the light is the biggest challenge for most applications of facial recognition systems [18].

1.3 Databases

There are four types of Datasets used in this work in order to demonstrate each of them in details

1.3.1 ORL (Our Database of Faces) Database

This set contains 400 images divided as follows. Ten photos per person. The number of people is 40. The main factor for different images of the same person is different lighting. In addition to facial expressions and the presence or absence of glasses. The background of all the images is the same. All images are the same size, 92*112 in grayscale. Figure 1.1 shows an example of this database [19].

Figure 1.1. An example of ORL database

1.3.2 YALE Database

There are 15 persons represented by 165 grayscale pictures in GIF format. Each person has 11 photos. The pictures differ from each other through facial expressions or some formations, in addition to the difference and angle of the light source. All images are the same size, 320*243 in grayscale. Figure 1.2 shows an example of this database [20].

Figure 1.2. An example of YALE database

1.3.3 FERET Database

This database is a collection of images employed in the Face Recognition Technology (FERET) initiative to evaluate facial recognition systems. It was created in 1993. This facial recognition technology database is a standard face image database that researchers may use to design algorithms and publish findings. The usage of a shared database also enables researchers to evaluate the efficacy of various methodologies and assess their strengths and shortcomings.

Between December 1993 and August 1996, the database's face photographs were gathered, provide 14126 photos to 1199 person. A high-resolution was provided by the DARPA which mean "Defense Advanced Research Projects Agency" in 2003. There are 2,413 still facial photos in the dataset, representing 856 people. Many Institutes manages this database, which has been utilized by over 460 research groups (NIST). All images are the same size, 256*384 in grayscale. Figure 1.3 shows an example of this database [21,22].

Figure 1.3. An example of FERET database

1.3.4 Georgia Tech Database

There are 50persons represented by 750 colored pictures in jpg format. Each person has 15 photos. The pictures show frontal and/or tilted faces with different facial expressions, lighting conditions and scale. All images are the same size, 241*181. Figure 1.4 shows an example of this database [23].

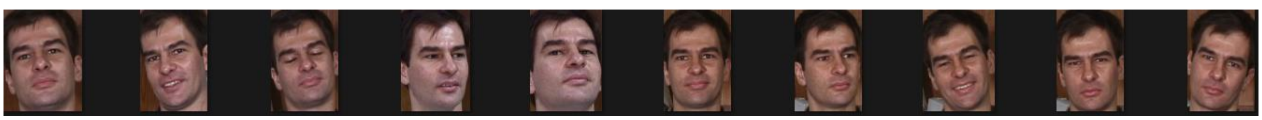


Figure 1.4. An example of Georgia Tech. database

1.4 Recognition

The task of recognizing faces is a challenge to digital systems, despite it being a simple and routine process for people. As each person is able to distinguish different faces with very high accuracy on a daily and routine basis [24]. The need for digital image processing arose when the prices of processors such as personal computers dropped. Where it was used in its infancy for biometric authentication.

Facial recognition systems can be defined simply as follows: They are systems capable of identifying a person by measuring his parameters based on a digital image [25,26]. The system saves these parameters mathematically and uses them for comparison when needed. Common examples of these systems include face recognition systems, fingerprint recognition systems, retina scanning recognition systems, and voice-based recognition systems.

Recognizing people through a digital image of the face has an advantage over other identification systems, as it is able to identify without direct contact with the person, in addition to the possibility of capturing the image from a distance and secretly. The importance of facial recognition has increased dramatically recently as a result of the increase in devices equipped with digital cameras. Where applications have increased dramatically due to the availability of very large amounts of different images on the web and the actual need to increase the level of security[27].

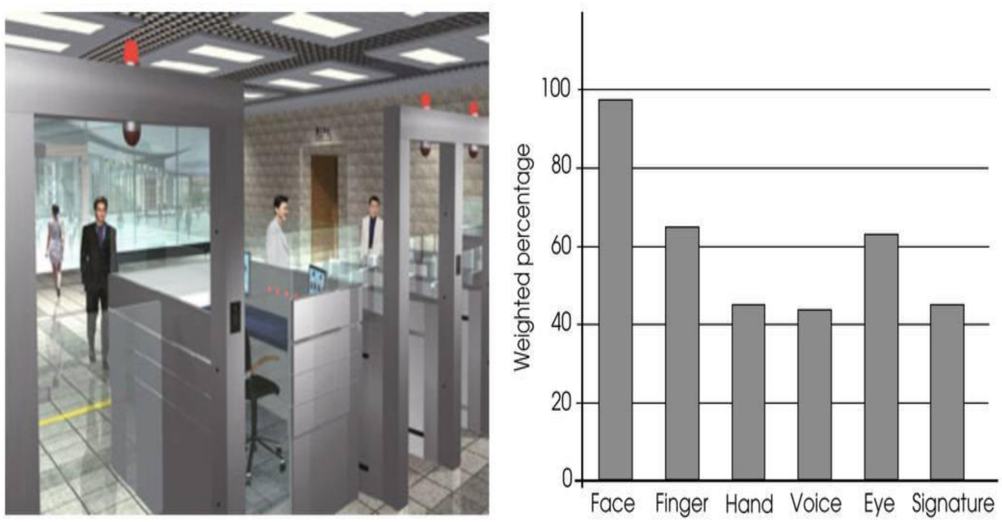


Figure 1.5 (a) Implementation of identification systems for passport verification. (b) relation of weighted percentage to the type of identification system

The significant development that has occurred in the field of identifying people based on a digital image of the face, especially in restricted environments, has increased dramatically since the advent of Eigenface [28]. Restricted environments can be defined as the environment in which the following conditions are met: adequate lighting, orientation of the face against the camera, a small distance between the face and the camera, and facial expression. The greater accuracy of these systems may outperform the natural identification of people by individuals. On the other hand, the challenges are still great when the images are under non-standard conditions[28].

1.4.1 Face Recognition Processing Workflow

Visual patterns are a convenient definition of the problem of recognizing people or faces. The face is represented by 3D (color images) or 2D (grayscale) matrices. Images consisting of two-dimensional arrays are more commonly used than three-dimensional images. These images are simpler and less expensive in terms of computational complications. While 3D image applications are increasing in cases where high security is required [29]. The facial recognition system in general consists of four main units, as shown in the figure 1.6 below.

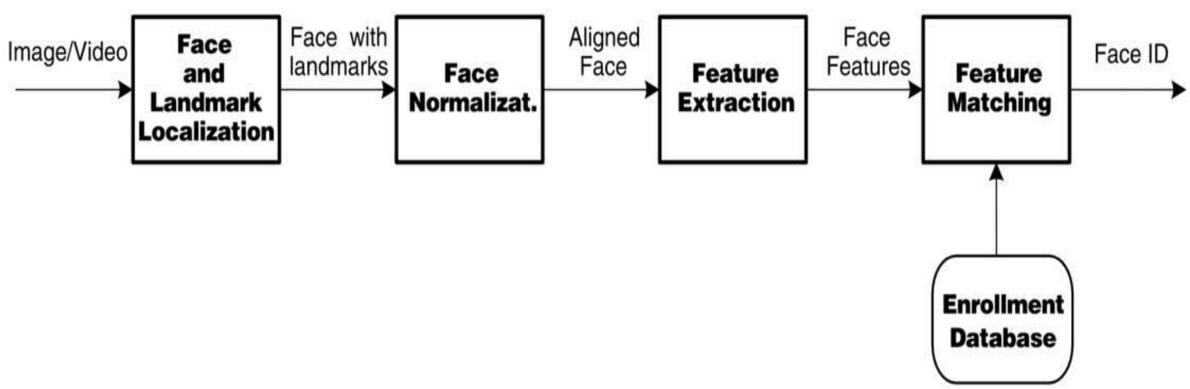


Figure 1.6 The basic units of facial recognition systems

The *face detection* process is of great importance in increasing the accuracy of face recognition [30]. By using this technique, all the non-essential components in the image (the background) are ignored and only the face area is cut out in preparation for the next stage which involves extracting the facial features (such as eyes, nose, mouth and other facial features) from the selected segment only. This stage is done through the unit for identifying features and aligning the face [30].

The process of normalizing the face geometrically is very important. It is expected that modern methods capable of detecting the face will develop under complex and varying environmental conditions [31]. The normalization process is based on converting the face to a standard frame, as well as optical normalization based on the intensity of lighting.

The most important features in facial images are the features that differ from one person to another so that they can be used to distinguish between people. The main task of the feature extraction system is to find enough geometric and photometric features to distinguish people from each other [30].

The matching process includes finding the largest possible match between the image to be identified and the images stored in the database. The result of the process is (matching or not matching). The main challenge in this process is to find the true measure of similarity by which the accuracy of facial recognition increases [32].

The accuracy of face recognition depends on many parameters. The process of extracting features from the face [29], as well as the process of normalization, in addition to all units of facial recognition systems, have a key role in changing the accuracy of facial recognition. Recognizing difficult patterns is the biggest challenge that contributes greatly to face recognition [29].

1.4.2 Face Recognition Types

Two dimensional, System The basis of face recognition techniques is to use only 2D to compare the image to be detected and the corresponding images stored in the database. The accuracy of these systems is rather low [33]. There are many limitations. As for this technique, such as the direct image of the individual, as well as the intensity of illumination, which should not differ much from the intensity of illumination in the corresponding image in the database. The above things, as well as the natural changes in the individual, such as the length of the hair, all affect the accuracy of the definition[33,34].

The adoption of 3D patterns to recognize faces has great advantages in terms of increasing accuracy, but in return, the complexity of the calculations increases. This type of system is less affected by the difference in light intensity than the two-dimensional systems[34].

1.4.3 Pros and Cons of the Face Recognition Systems

Facial recognition systems have many advantages and disadvantages. Face recognition systems were once a fictional system, but now they are more than just a fantasy. It has a huge impact in many areas of life. Most smart device manufacturers have become dependent on facial recognition systems to identify the owner of the device [35].

To identify a specific face, three basic stages are applied: the first stage includes detection of the face area, the second stage includes extracting characteristics from the face, and the last stage includes comparison with the characteristics stored in the database to reveal the identity of the person. There are many different applications for facial recognition. Therefore, there are many advantages and disadvantages depending on how it is used and the type of application [35,36]:

The pros can be summarized as bellow

1. Security Improvement

The use of facial recognition systems increases the security of the user (individual or company). Where all entrants to a specific area are identified and a warning is given in the event of unauthorized entry. Thus, the number of security personnel can be reduced.

2. Accuracy Improvement

The reliability of facial recognition systems is greatly increased as a result of technological development. Face recognition using 3D images and infrared cameras is rapidly increasing accuracy to the point where it is difficult to deceive, increasing safety.

3. Fully Automated

Previously, recognition systems alone were not able to guarantee high reliability. Therefore, it was necessary to have a guard to ensure security. This is no longer necessary. The system is able to manage everything with high reliability, which increases convenience and reduces costs.

The cons can be summarized as bellow

1. Data Storage

The database used in facial recognition systems includes many images. The memory used to store these images is not small, so this issue is one of the disadvantages of this technology. It always requires a large memory to store and process data. Big companies use many computers to process everything quickly.

2. Camera Angle

The angle of the camera has a huge impact on the accuracy of the entire face recognition system. Accurate facial recognition requires multiple angles of photos, including profile. The presence of any obstacles such as hats or facial hair will cause poor accuracy.

1.5 Applications of Face Recognition

1. Building insurance is one of the applications used by security companies.
2. Immigrant checkpoints use facial recognition systems to increase the accuracy of their work
3. Companies that have a large number of cars use these systems to identify drivers and secure vehicles
4. Used on trips to ensure that the right people ride the right buses with their designated drivers
5. It is used in the Internet of Things to recognize faces for enhanced security measures [37].

1.6 Thesis organization

The project is organized as follow

Chapter Two: includes an overview of the used transforms like DWT and 2D DWT. A comprehensive overview of PCA has been discussed additionally to 2D PCA.

Chapter Three: The third chapter discusses the system proposed. Which includes flowchart of the algorithm. It shows preprocessing steps, feature extraction and classification. In addition, it contains an explanation of training and testing modes.

Chapter Four: Chapter four shows project results based on k-Fold Cross Validation technique. The experimental results of ORL, YALE, FERET, and Georgia tech Databases is presented.

Chapter Five: In this chapter, the conclusion and future work are proposed.

CHAPTER TWO

THEORY

CHAPTER TWO
THEORY

This chapter deals with the main topics that will be relied upon in the subsequent chapters in order to achieve the goal of this work. As an example of the mentioned topics are Fourier transform (Discrete or Windowed) Fourier Transform, Wavelet transform, Discrete wavelet transform and Principal component analysis.

2.1 Transforms

In this section, Fourier transform, wavelet Fourier transform, and Discreate Fourier transform will be discussed as follow:

2.1.1 Fourier Transform

Fourier transform is a transformation used in mathematics to analyze functions of space and time into functions of a spatial or temporal nature, as an example of this is the use of notes that make up a musical chord in terms of frequency and special sizes. The Fourier transform is a mathematical operation and frequency domain representation that relates the time or space function to the frequency domain representation [38].

The Fourier transform outcomes in a complex value for frequency when the conversion of the time and it has two values, the first is the amount of this frequency is the absolute value of the transformed value and the second is its argument which represents the phase difference angle of the transformed value. The time domain is the domain of the original function because the Fourier transform does not depend solely on the time functions. There is a second theorem, the Fourier inversion theorem, where it synthesizes the original function using its own frequency domain, which is the inverse of the Fourier theorem [38].

Linear processes implemented in unique domain (frequency or time) have equivalent processes in another domain, that are sometimes at ease to implement. The multiplication by frequency corresponding to process of differentiation in time domain, so in frequency domain many differential equations are much easy to study.

In addition, normal multiplication in frequency domain performs convolution in time domain. When carrying out the chosen processes, the conversion of the consequence can be set back to time domain. The harmonic study is the logical reading for relationship among both time and frequency domains, containing the types of functions and processes which they are much easier in any one, and takes much deep contacts for several ranges of current mathematics [38].

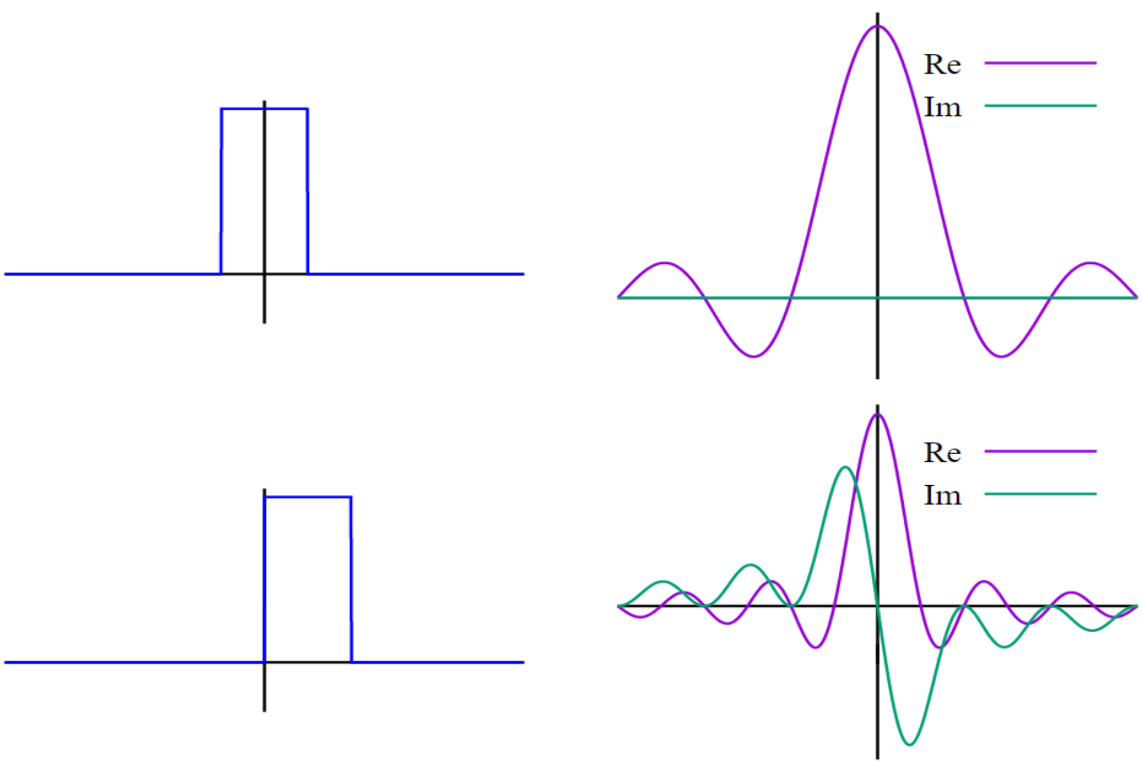


Figure 2.1 Fourier transform

In the figure 2.1 the first part the function of unit pulse f(t) and Fourier transformation value of it $\hat{f}(\omega)$, a function of frequency ω. Transformation delay in time domain is deduced as phase shifts which is complex value in frequency domain. The second part is shown called g(t) which is a unit pulse delayed, in addition to real or imaginary section of Fourier transform. This transform consists in a function which are Eigen functions for the set of transformations. The imaginary portion of $\hat{g}(\omega)$ is cancelled for the reason that a negative mark exponent which is used in Fourier transformation as equation (2.1), that is the defaulting as resultant from Fourier series, however the mark that does not problem for a convert that is not resulting to be reversed [38].

$$F(w) = \int_{-\infty}^{\infty} f(t)e^{-j\omega t}\, dt \tag{2.1}$$

2.1.2 Discrete Fourier Transform

The discrete Fourier transform (DFT) in mathematics transforms a limited order of function samples which have same space between them into an order which has same length and an order of same space samples of the discrete time Fourier transform (DTFT), that it has a function of frequency which has complex value. The period of the DTFT when it is sampled is reciprocal of the period of the input order.

A Fourier series which is called inverse DFT using the DTFT samples as factors of sinusoids that they are complex at the equivalent DTFT frequencies. It takes the similar sample values as the unique input order. The DFT is as a result held to be a frequency domain exemplification of the unique input order. If the unique order periods wholly the values that they are not zero of a function, its DTFT is endless (and intervallic), and the DFT offers unconnected samples of any sequence. If the unique order is one set of a intervallic function, the DFT offers all values of any DTFT cycle which are not zero [39].

The DFT is the best significant discrete convert, used to implement Fourier transformation in several applied uses. The function in digital signal processing is one amount or sign that differs with time, for example the pressure of the wave of a sound, temperature reading every day, and a signal of a radio all these signals are sampled for a limited period of time (mostly called window function).

Sometimes for image processing, samples represent pixel values in a column or row of a bitmap. DFT is always used in performing various mathematical operations such as multiplication of integer large numbers and convolutions, and it is also efficient in finding solutions to partial differential equations. As it contracts with a limited volume of data, it be able to be executed for Computers through numerical procedures or even devoted hardware. These executions frequently work more efficient algorithms of fast Fourier transform (FFT); besides this, the "DFT" or "FFT" are used frequently. In addition to present using, the "FFT" initialized can be used in the indefinite name "finite Fourier transform" [39].

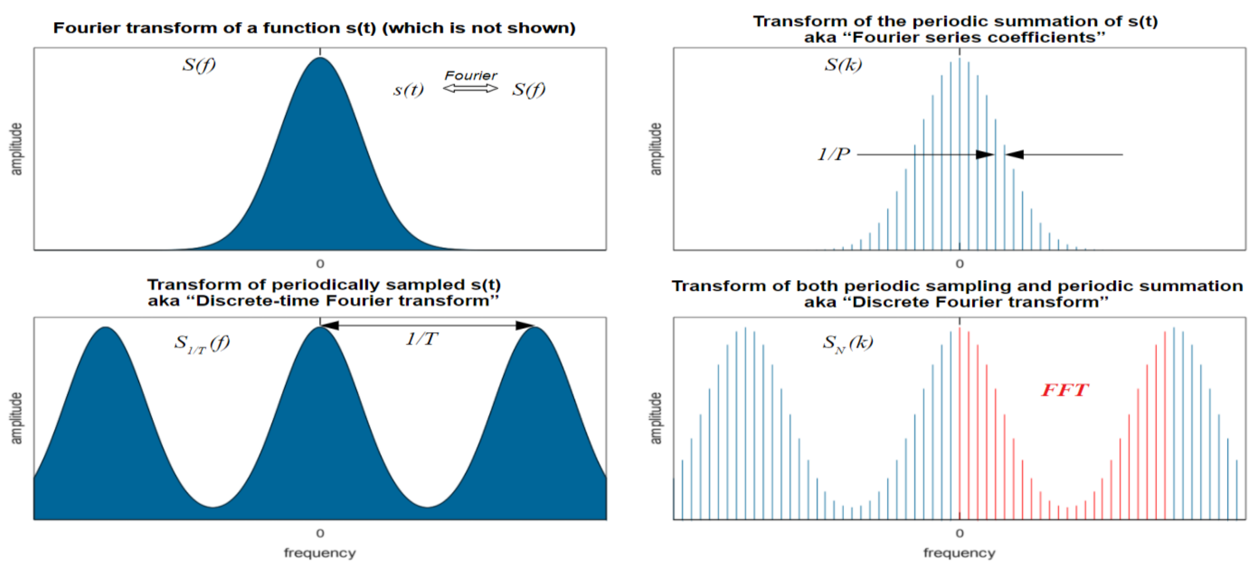


Figure 2.2 Discrete Fourier transform

The left parts of figure 2.2 represent Fourier transform and DTFT which it is periodic sum, while the right parts represent the spectral orders which are calculated from the first cycle from the periodic sum of the signal s(t) and the lower part from the first cycle of the periodic sum of s(nT) order. The function (2.2) are the integral of Fourier series and the sum of DFT respectively [39].

$$X(n) = \sum_{k=0}^{k-N-1} x(k)e^{-j\left(\frac{2\pi}{N}\right)nt} \quad where\ m = 0,1,2,3, \ldots\ldots N-1 \ \ldots\ldots..(2.2)$$

2.1.3 Windowed Fourier Transform (WFT)

The Fourier transform can represent a wave by representing periodic functions. We need to take part of the reference and focus Fourier analysis in it just to allow the descriptions to be local. To carry out this process, we can use the window function in which the part of the signal that is during the analysis is always saved and then we move it to another place to cover the other time parts we need.

Our method is called short-time Fourier transform or windowed transformation. Gabor has proposed this method, as he used a Gaussian function, so this transformation is called the Gabor transform, which is considered efficient because of its optimal features in most frequency and time domains. This localization is always simultaneous in the frequency and time bands and has a problem which is that it is restricted to a smaller limit called the Heisenberg uncertainty principle.

For the current state, like this function equal the result of multiplying the complex exponential by the window function, such as, g(t)e−i(ωt), the Gabor transform case after modulation and Gaussian, in this state the lower limit for the Heisenberg confidence concept touch. Localization properties of such equations in frequency or time domain shown in Figure 2.3, which shows those which compared to basic equations of the transformations that are mentioned in the previous time [40].

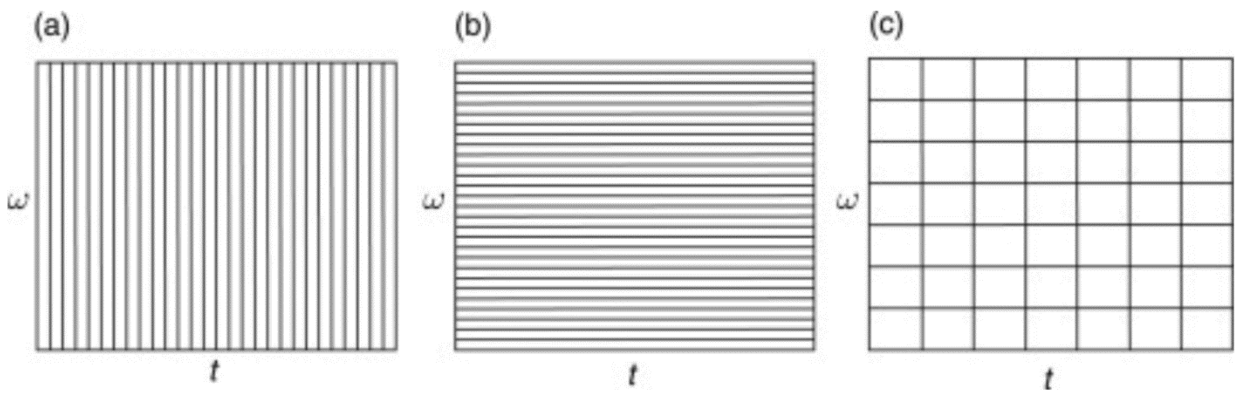


Figure 2.3. The localization in time and frequency space, of functions which are basically used for the so far mentioned transforms (Heisenberg boxes) a) Transform Dirac-δ (which are using for representing a time-domain signal); b) Fourier transform; while c) windowed Fourier transform

2.2 Wavelet Analysis

The results obtained from analyzing a variable signal using FT or STFT are not convincing. If we want more convincing results, the use of wavelet analysis would be better, since one of its advantages is that it is able to implement local analysis and is also able to search for aspects of the signal. Other analysis techniques usually lack such a feature, such as identifying breaking points and trends as well as interruptions, etc. It also allows the implementation of Multiring analysis, so it is considered better when compared with STFT [41].

We can freely choose the analysis function of the wavelet transform without using sinusoidal forms, so this is better when compared to Fourier analysis. To the purpose of distinguishing between many frequencies, the wave equation of ψ (t), which is a small in value wave, must be oscillating. The wavelet always contains the shape of the window and the analysis. Figure 2.4 shows an example of a potential wavelet which is called a Morlet wavelet. A set of wavelet functions with certain features has been developed for CWT [41,42].

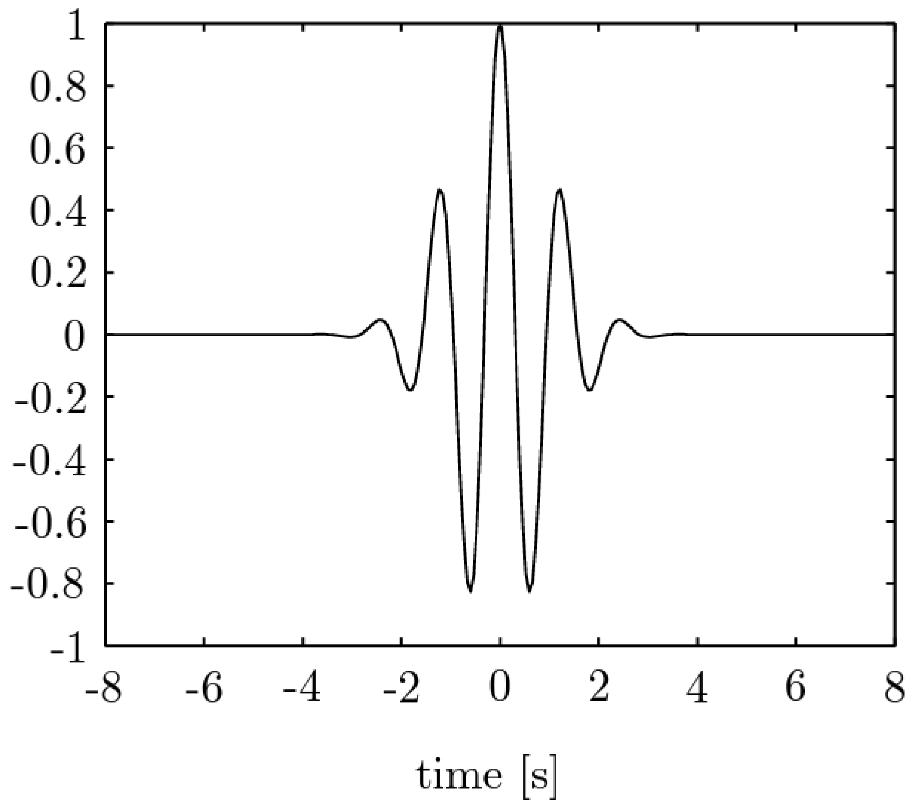


Figure 2.4: Morlet wavelet

A studying of the wave ψ(t) as wavelet for the mathematical procedure which is stated:

1) Wavelet has limited energy

$$E = \int_{-\infty}^{\infty} |\psi(t)|^2 \, dt < \infty \tag{2.1}$$

E is the energy which it is square integral of magnitude for the studying wave ψ(t) which it is less infinity [42].

2) Ψ(f) it is Fourier transform for the wave ψ(t), hence it must satisfy the condition

$$C_{\psi} = \int_{0}^{\infty} \frac{|\psi(f)|^2}{f}\, df < \infty \tag{2.2}$$

Noted that the condition states that the component of frequency of wavelet which is not zero (Ψ (0) = 0), for example the wavelet mean ψ(t) equal zero. It is called constant of admissibility. The value of Cψ depend on wavelet that we will choose.

3) The Fourier transform Ψ(f) for complex wavelets must be vanish for frequencies which are negative and real.

Wavelet transform can be applied to continuous and discrete signals [41].

2.2.1 Continuous Wavelet Transform

It is calculated using the following equation 2.3:

$$X_{wt(\tau,s)} = \frac{1}{\sqrt{|s|}} \int_{-\infty}^{\infty} x(t)\psi^* \left(\frac{t-\tau}{s}\right) dt \qquad 2.3$$

The converted signal XWT(τ,s) is of variables τ and s. The original wavelet is signified by ψ, the ∗ shows in complex wavelet we use the complex conjugate. The signal of energy is regulated at each balance when we divide the wavelet factors by 1/(s^0.5). It may confirm that wavelet has similar energy in each balance.

Original wavelet may be constricted or expanded using variation in the balance factor s. This difference with balance s varying the wavelet central frequency fc in addition to length of this window too. As a result, the balance s is used as an alternative of frequency for signifying outcomes of analysis of the wavelet. The factor of translation τ shows a position in time of wavelet, with variation of τ wavelet will be able to be moved by the signal. If we make constant balance s and varying translation τ, the scale rows of plane of time which need fill, changing a balance s with making the constant translation τ plugs the scale columns of scale plane. Components in XWT (τ, s) will named wavelet factors, any wavelet factor connected to balance frequency or a point in domain of time [43].

In other hand, the inverse continuous wavelet transformation (ICWT) is also the inverse transform of WT, like the case of the FT and STFT. Inverse continuous wavelet transformation (ICWT) calculated using the following:

$$x(t) = \frac{1}{C_{\psi}^{2}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} X_{wt(\tau,s)} \frac{1}{s^{2}} \psi\left(\frac{t-\tau}{s}\right) d\tau ds \tag{2.4}$$

Noted that must satisfy the wavelet second condition must have satisfied by constant of admissibility Cψ.

An equation 2.4 of wavelet has fc which is central frequency of it itself with any balance, balance s depends inversely with the frequency. The high balance matches to low frequency, give data for signal. Low gauges match high frequencies giving signal data detail.

The Heisenberg for inequality WT still holds, time bandwidth multiplying Δf Δt is not changed and lesser limited. Reducing the gauge s, for example the smaller window, growth resolution of time Δt, causing falling Δf resolution of frequency. That is indicating that the resolution of frequency Δf compared to frequency, for example analyzing of wavelet has constant relation to frequency resolution. Morlet wavelet is shown in Fig. 2.4 achieved using a window Gaussian, that fb is the bandwidth factor and fc is center frequency [43].

$$\psi(t) = g(t)e^{-j2\Pi f_c t} \ , \qquad g(t) = \sqrt{\Pi f_b}\, e^{\frac{t^2}{f_b}} \tag{2.5}$$

The fc and the fb of wavelet are the factors of tuning. From Morlet of wavelet, frequency in addition to scale coupled like

$$f = \frac{f_c}{s} \tag{2.6}$$

If we want to calculate the continuous wavelet transform, we may take discrete values for the scaling factor s and translation factor τ which is results in wavelet factors are named wavelet series. For study purposes, the discretization be able to be executed randomly, though it is required reconstruction.

The constant comparative resolution of frequency of study of wavelet also recognized like the property constant Q Which Q is factor quality of filter, which is called the center frequency fc that bandwidth fb divided it. For Q constant analysis (comparative resolution of frequency which is constant), a grid sample dyadic the climbing looks suitable. Grid dyadic create with music and hear of human too. Grid dyadic discretize the factor scale with logarithmic gauge. Time factor discretized by the scale factor. The dyadic grid is any easiest and efficient discretization means with applied purposes which results in production of a basis orthonormal wavelet.

The goal of the scaling function is to filter out conversion levels with all their values. The scaling function ensures that the entire spectrum is covered [44].

2.2.2 Discrete Wavelet Transforms

A discrete wavelet transform (DWT) for functional or numerical analysis is some wavelet transformation when the wavelets sampled discretely. Like other wavelet transformations, an advantage is resolution temporal as related to Fourier transforms: it captures both position data (in time) and frequency.

The original DWT was developed by mathematician Hungarian Alfréd Haar. An input characterized by list of 2^n numbers, Haar wavelet convert can be considered to couple input up values, keeping the variance and passthrough sum. This method is repetitive recursively, coupling up the summations to verify the next balance, that leads to 2^(n-1) variances and a last summation.

In general, Haar DWT states the desired possessions of wavelets. First of all, it can be executed in O(n) processes; the other, it captures temporal content, for example the periods when these frequencies happen in addition to notion of the frequency content for input, which it occurs by testing it at many gauges. For these both properties, the Fast wavelet transform (FWT) become the best alternative to the conventional fast Fourier transform (FFT).

According to the rate change operatives in the bank filter, the discrete WT is really very sensitive to placement the signal in time. Zhong and Mallat suggested a novel procedure for wavelet illustration of signal using to address the time varying problematic wavelet transforms, that it is invariant of time shifts. Due to this procedure, which is named a TI-DWT, the wavelet convert is determined for any point in period.

The discrete wavelet convert has a large number of uses in engineering, science, computer science and mathematics. In addition, signal coding using it to perform a discrete signal in a form which is more terminated, frequently like a preconditioning for compression of data. Applied uses can also be used in signal processing for gait analysis of accelerations, image processing in digital communications and others [45].

Depending on the number of decompositions C is related to the number of level (L) according to the following equation 2.7. Figures 2.5 shows Block diagram of filter analysis. Figure 2.6 shows a 3level filter bank

$$C = 3 * L + 1 \tag{2.7}$$

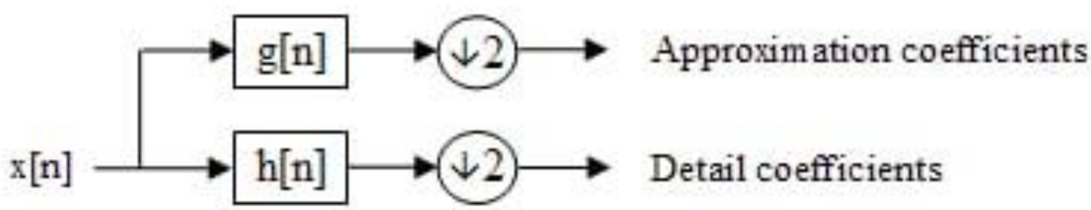


Figure 2.5. Block diagram of filter analysis

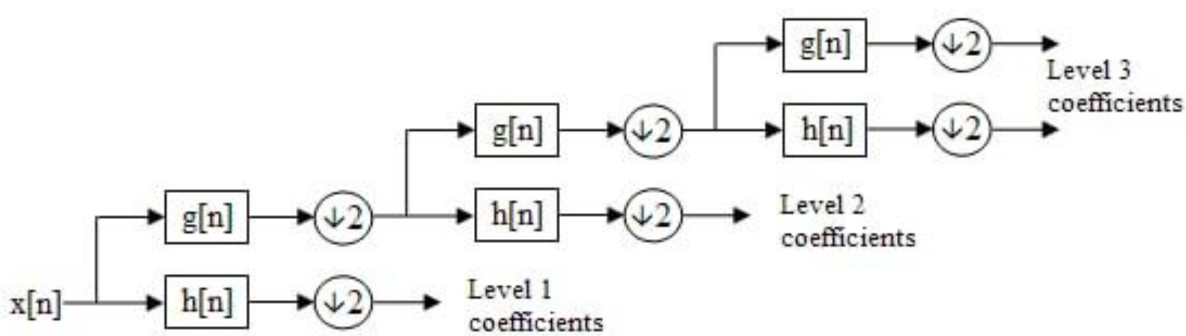


Figure 2.6. A Three level filter bank

The filter output of the low-pass filter g in the diagram above is then subsampled by 2 and further processed by passing it again through a new low pass filter g and a high- pass filter h with half the cut-off frequency of the previous one [45]. Figure 2.7 shows multi-level decomposition

$$y_{\mathrm{low}}[n] = \sum_{k=-\infty}^{\infty} x[k]g[2n-k]$$
$$y_{\mathrm{high}}[n] = \sum_{k=-\infty}^{\infty} x[k]h[2n-k] \qquad (2.8)$$

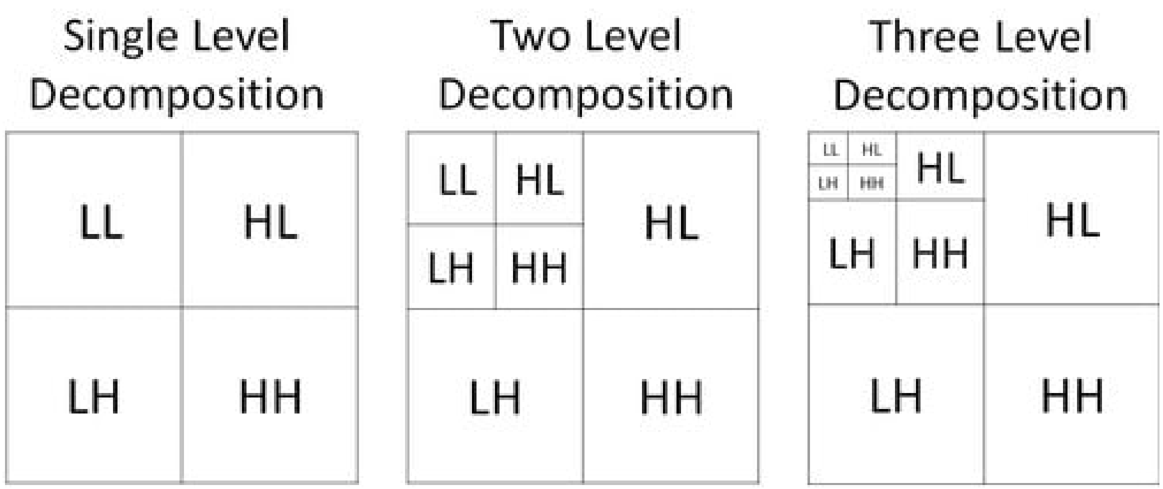


Figure 2.7 multi-level decomposition

2.3 2D Discrete Wavelet Transform

The DWT can be applied in two-dimensional information like an gray scale image based on the following Figure 2.8. The one-dimensional Discrete wavelet transform is applied two times. One of them in each row while the second one for each Column [46].

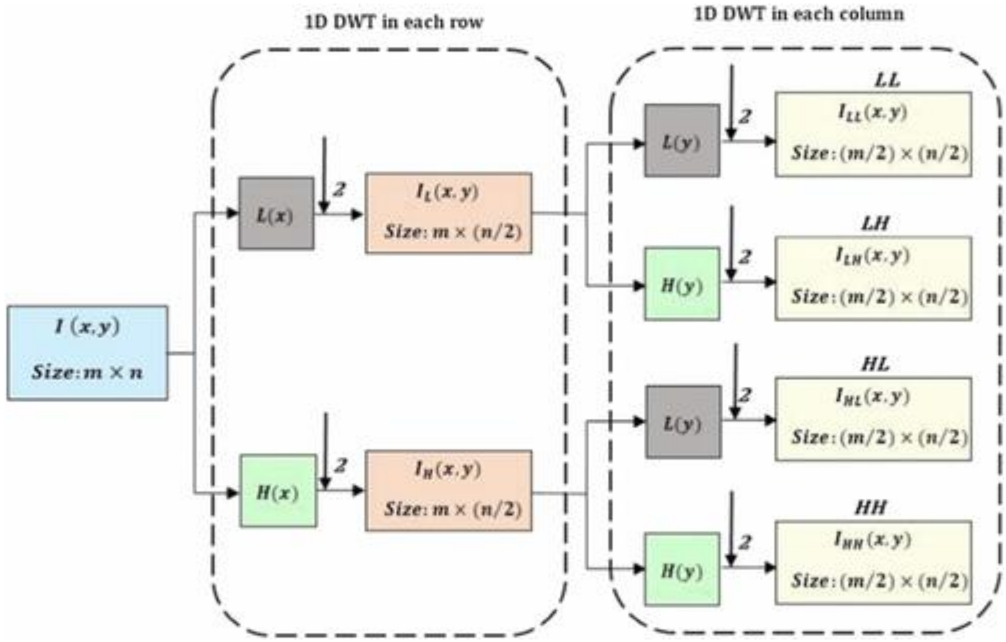


Figure 2.8. 2D DWT

2.3.1 Nyquist Sampling Theorem

Sampling is the basis on which to convert from continuous signals to discrete time signals. The sampling rate is not open but is limited by an important condition. The sampling rate must match the frequency of the continuous time signal to be converted into discrete time signals. This is to ensure that all information is handled without significant loss.

The sample rate that must be adopted to ensure the preservation of the data must exceed twice the frequency of the signal to be sampled. This rule is called Nyquist sampling theorem as the following equation 2.9 [47]. Figure 2.9. shows the Nyquist sampling theorem.

$$Nyquist\ rate = \frac{1}{2B} \qquad (2.9)$$

Where B represent the frequency range.

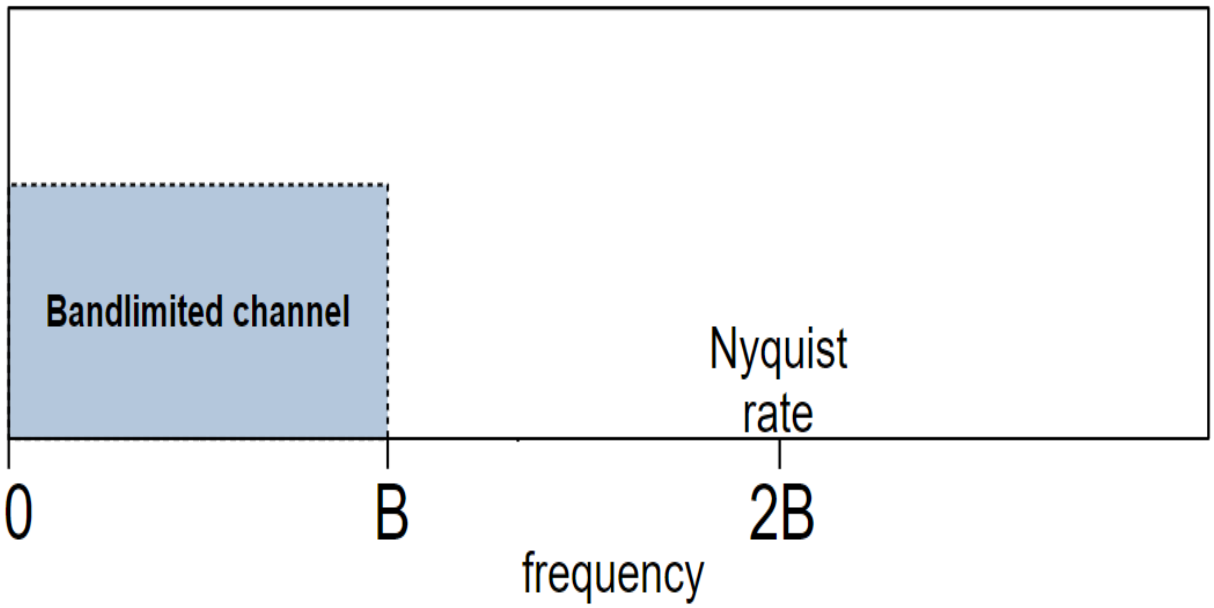


Figure 2.9. Nyquist sampling theorem

2.4 Types of Wavelet filter

There are multiple types of wavelet filters as follows

- Haar wavelet
- Gaussian wavelet of order 1
- Gaussian wavelet of order 3
- Morlet wavelet
- Mexican Hat
- Daubechies wavelet

Some of them explained in figure 2.10 Bellow. The Haar wavelet used in the work [48].

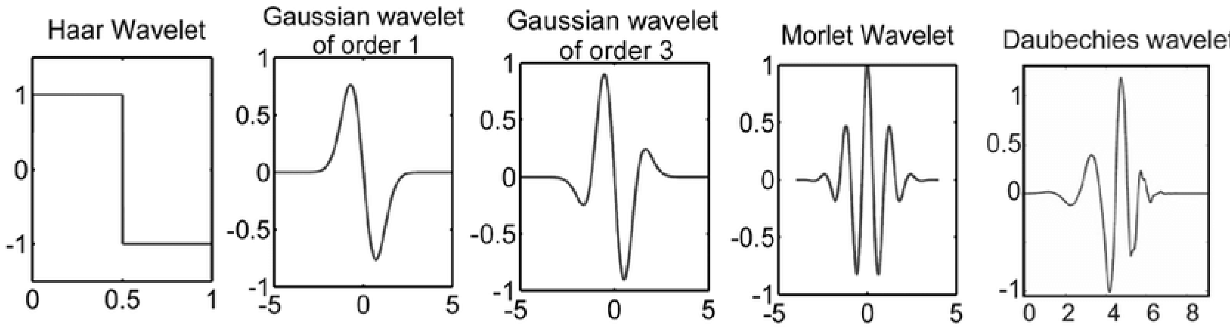


Figure 2.10. Types of Wavelet filter

2.5 Haar Wavelet

it is a transform which is based on the class of orthogonal, non-sinusodial, non-square functions whose elements are multiples of -1,+1,0 by the powers of 2^0.5 .
it is a computational efficient transfer function for N point whose vectors require 2(N-1) additions and N multiplications [48].

STEP 1: Now determine the value of N which represents the order of the system or the size of the kernel.
STEP 2: Determine the total number of bits required based upon the order of the function using the expression n=log N
STEP 3: Determine the values of the two specific terms p and q.

i. $P \in [0,n-1]$ $o \leq p \leq n-1$
ii. If $p = 0$ then the value of q can be either 0 or 1
iii. If $p \neq 0$ then $q \in [1,2^p$ $1 \leq q \leq 2^p$

STEP 4: Now determine the value of k which tells us the total number of rows in the kernel by the equation k = 2^p + q – 1.
STEP 5: Determine the value of the Z where $Z \in [0, 1/N, 2/N, \ldots, N-1/N]$
STEP 6: If $k = 0$ then $H(Z) = (1/N)^{0.5}$

$$\text{Else } H_k(z) = H_{pq}(z) = (1/N)\text{^}0.5 \begin{cases} +2^{p/2} : \frac{q-1}{2^p} \leq Z < \frac{q-0.5}{2^p} \\ -2^{p/2} : \frac{q-0.5}{2^p} \leq Z < \frac{q}{2^p} \\ 0: \quad \text{otherwise} \end{cases}$$

2.6 Principal Component Analysis (PCA)

It is a statistical method for detecting patterns in environments of high complexity and many dimensions. It can be used to detect similarities and differences between patterns. It translates information from a coordinate to the other, with first principal component containing the main pattern of locations. The information is projected along of the perpendicular to the first principal component axis second principal component axis. PCA is an useful technique for discovering patterns of similarity and dissimilarity in face recognition since spotting patterns in high-dimensional data is difficult, and photos are represented as points in high-dimensional space [49].

Eigen-face is a well known face recognition system. Authors of [11] discussed it in the work "Face Recognition Using Eigen Faces," which was published in 1991. Their approach is based on PCA, which reduces gray level pictures to a lower dimension subspace. The distance-based matching approach is used to recognize pictures between them. The new face will be classed as known if the gap between it and the faces in the training set is minimal and over a certain threshold. The new face would otherwise be labeled as unknown [49].

2.7 Two Dimensional PCA

Assume that X is a unitary column vector with n dimensions. The subsequent transformation function was used to translate the projecting picture A, a m*n random matrix, onto X.

$$Y = AX. \tag{2.10}$$

As a consequence, M dimensional vector is constructed, which referred to as the projection characteristic vector of picture A. What factors should we consider while choosing a decent vector X? the overall dispersion might be employed to determine the projection vector X's discriminating power. The trace of characteristics vectors covariance matrix may be employed for determine overall dispersion of the photo. following criterion used for that:

$$J(X) = tr(S_x). \tag{2.11}$$

Where:

S_x : represent the covariance matrix for train image.

J(X) : trace of Sx.

A practical meaning of increasing the criteria in equations (2.12) would be to identify a projection vector, leading in the maximum overall dispersion of the image. The following formula 3.3 denotes the covariance matrix Sx.

$$S_x = E(Y - EY)(Y - EY)^T = E\big(AX - E(AX)\big)\big(AX - E(AX)\big)^T \tag{2.12}$$

So,

$$tr\,(S_x) = X^T\,(A\,(E - EA)^T\,(A - EA\,))\,X \tag{2.13}$$

the following matrix can be defined as

$$G_t = E\,((A - EA)^T\,(A - EA)\,) \tag{2.14}$$

The structure of the matrices The photo scattering (covariance) matrices is known as Gt. Out of its description, it's simple to see that Gt is a n*n positive definite matrix. To use the train photo sets, we can assess Gt immediately. Assume there are M overall train photo. the jth train photo is represented as a m * n matrix Aj (j = 1, 2, 3,..., M), and the mean photo the train set is represented as A. The following formula 3.6 may then be used to calculate Gt [49,50].

$$G_t = \frac{1}{M}\sum_{J=1}^{M}(Aj - \bar{A})^T(Aj - \bar{A}) \qquad (2.15)$$

From equation 2.16 the following can be expressed

$$J\,(X) = X^T\,G_t X \qquad (2.16)$$

X represent unity column vector. The generalized whole dispersion criteria is the name of this criterion. X unity vector called the optimal projection axis.

The optimal projection axis X_{opt} is the unity vector lead to make J(x) max value, i.e., the eigenvector of equation 2.16 related to the biggest eigen value.

The difference between PCA and 2DPCA can be illustrated as shown in Figure 2.11 [50].

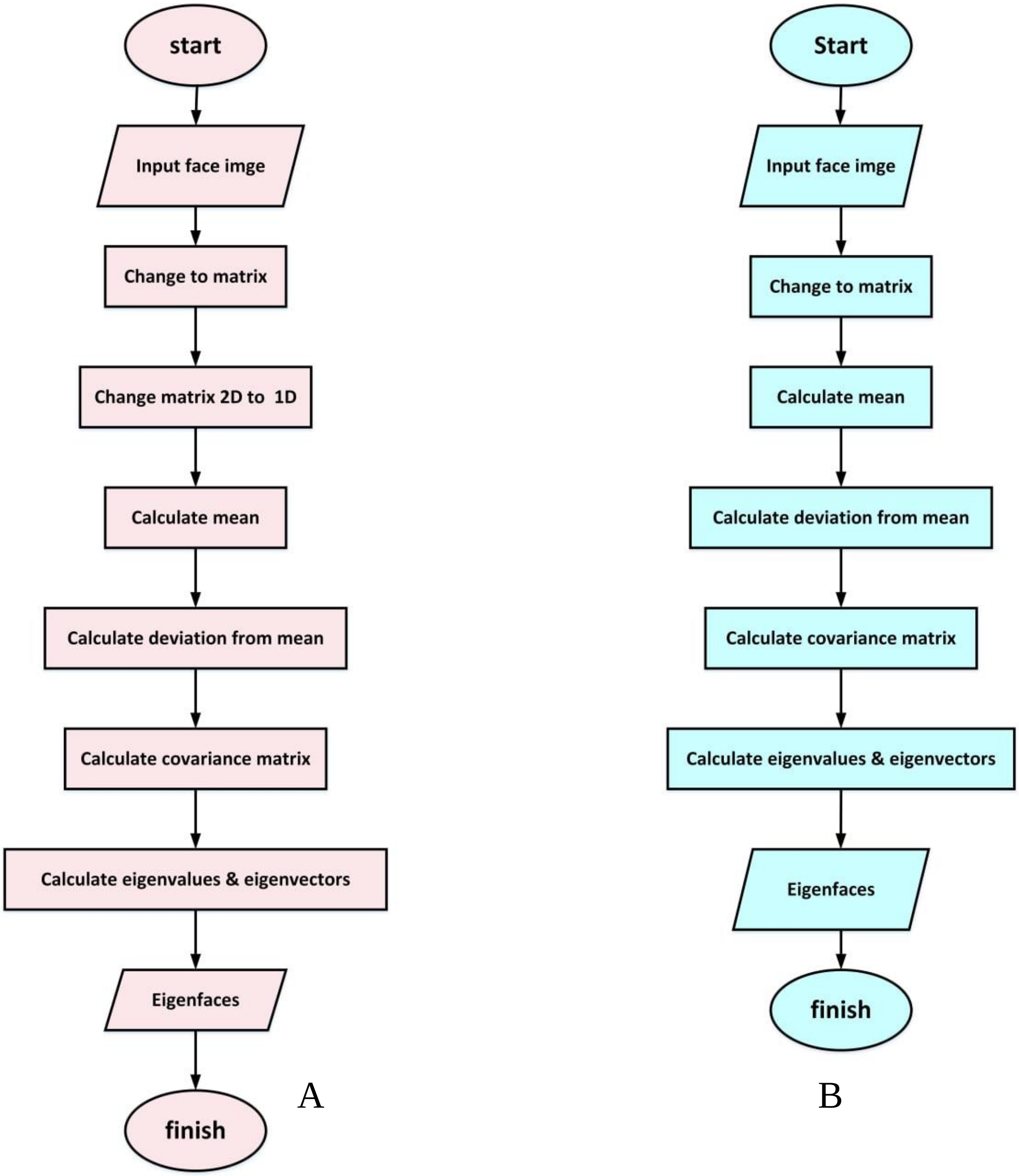


Figure 2.11 a- PCA flowchart, B- 2DPCA flowchart

2.8 Ant Colony Optimization Algorithm

Naturally, ants can find the shortest path between their nest and food site. It can find the shortest path even though there are some possible paths. The process is carried out using pheromone. When ants pass from a specific pathway, they put pheromone. Pheromone attracts other ants to pass the same path. Ants tend to choose a path that has a higher pheromone concentration. After a while, pheromone concentration becomes high on the shorter path while decreasing from wrong paths. The presence of an obstacle in the ants' path makes him choose his way randomly at first. All ants then move gradually to the shorter path, as shown in Figure 2.12 [51].

ACO follows the algorithm shown in Figure 2.13. Initially, configuration and configuration occur. During this step, the initial value of the pheromone is set. Then it enters a duplicate loop, ending when one of the stop conditions is reached. The amount of pheromone is continuously updated [51].

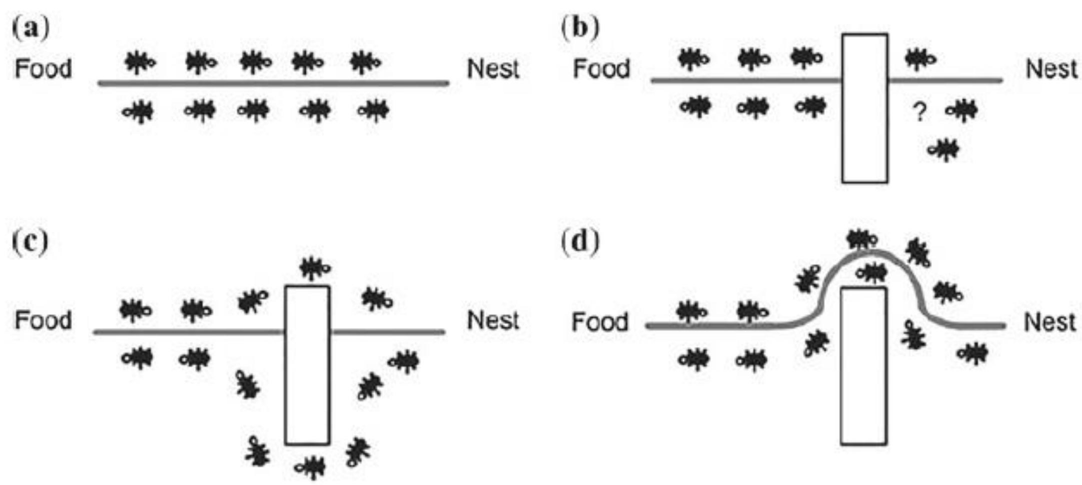


Figure 2.12. Ants approach to find the best path to the destination (food).

```
Set Parameters, initialize pheromones;
While (termination condition not met) do
Construct Ant colony;
Apply local search;
% Optional update pheromones;
End
```

Figure 2.13. The programming steps to obtain the optimal path according to ACO.

2.9 Genetic Optimization Algorithm

Genetic Optimization is a technique used in this algorithm to find the parameters required to optimize a particular objective. the idea of Genetic intelligence based on the concept of Darwin's theory of evolution. The Genetic algorithm works by simultaneously maintaining. Several candidate solutions in the explore space [52].

2.10 Euclidean Distance

The Euclidean distance among two points in Euclidean space is the length of a line segment among the two points. It can be determined from the coordinates of Cartesian the points using the theorem called Pythagorean theorem, thus occasionally being named the Pythagorean distance. The names derived from the Greek ancient mathematicians Euclid and Pythagoras, though Euclid didn't perform distances as numbers, and the linking from the Pythagorean formula to distance determination was not completed till 18th century [53]. Figure 2.14 shows the Euclidean distance. Figure 2.15. present the N dimensions Euclidean distance

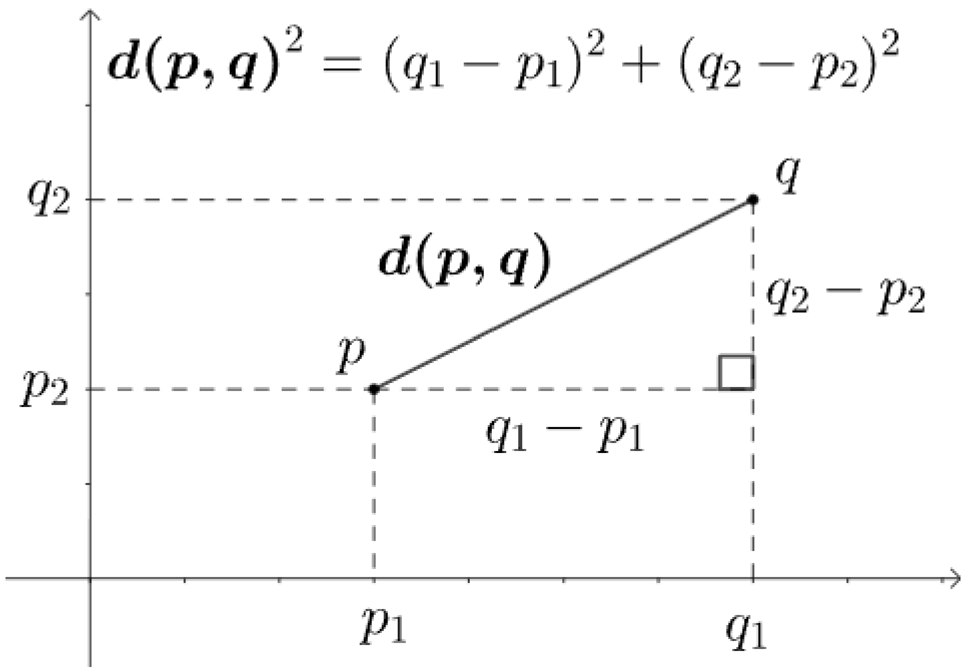


Figure 2.14: Euclidean distance

For the points assumed by Cartesian coordinates in n-dimensional Euclidean space, the distance is

$$d(p,q) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \cdots + (p_i - q_i)^2 + \cdots + (p_n - q_n)^2}. \quad (2.17)$$

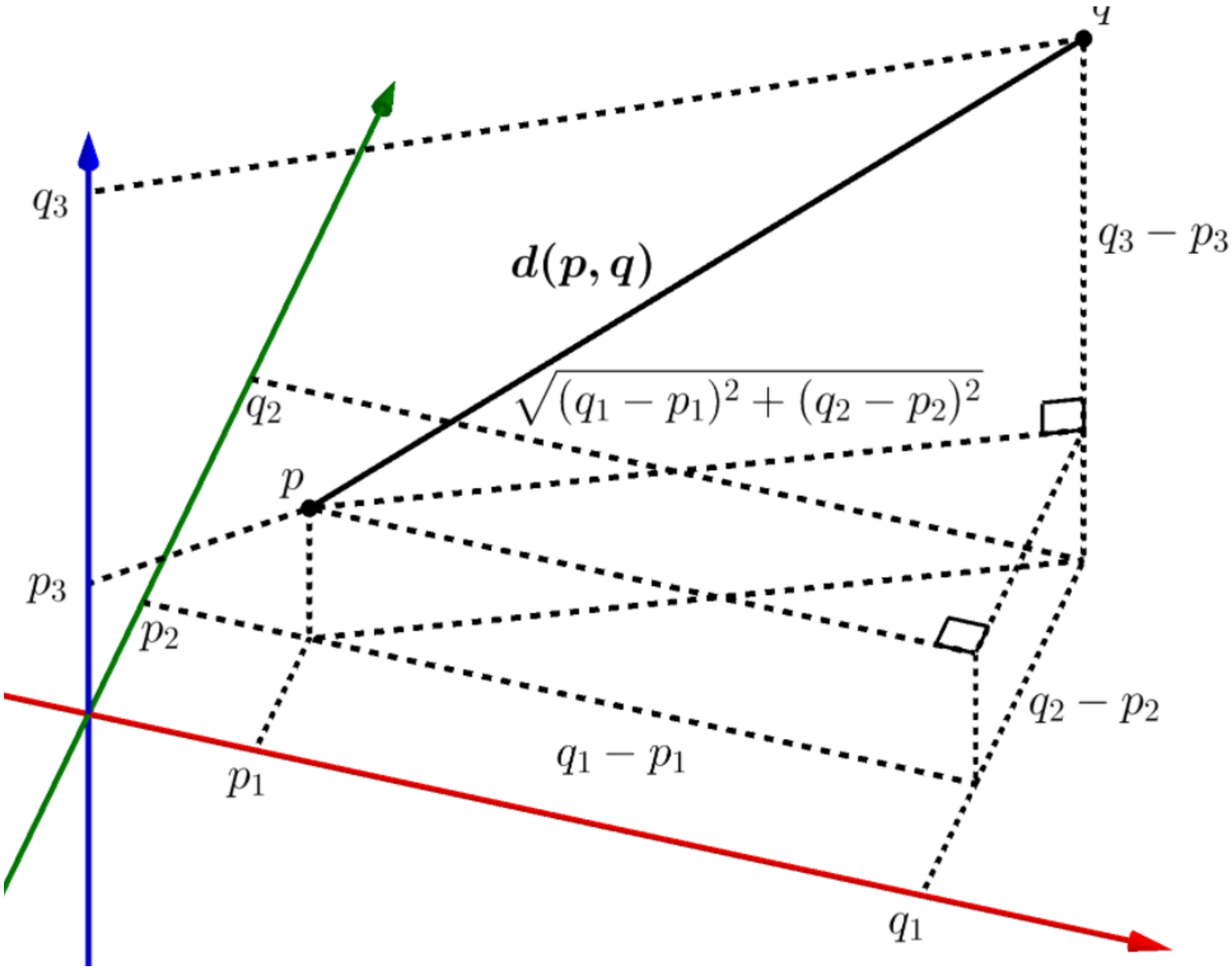


Figure 2.15. N dimensions Euclidean distance

2.11 K - Fold Cross Validation Technique

Cross-validation is primarily used in applied machine learning to estimate the skill of a machine learning model on unseen data. That is, to use a limited sample in order to estimate how the model is expected to perform in general when used to make predictions on data not used during the training of the model [54].

It is a popular method because it is simple to understand and because it generally results in a less biased or less optimistic estimate of the model skill than other methods, such as a simple train/test split.

The general procedure is as follows:

1. Shuffle the dataset randomly.
2. Split the dataset into k groups
3. For each unique group:
   a. Take the group as a hold out or test data set
   b. Take the remaining groups as a training data set
   c. Fit a model on the training set and evaluate it on the test set
   d. Retain the evaluation score and discard the model
4. Summarize the skill of the model using the sample of model evaluation scores

CHAPTER THREE

PROPOSED ALGORITHM

CHAPTER THREE

PROPOSED ALGORITHM

In this chapter, the proposed system is presented. Also, the steps of implementing each technique and transform employed in the proposed system are discussed. The purpose of using PCA and 2D-DWT is illustrated in this chapter. Furthermore, the question of “why does someone need dimensionality reduction in every recognition system that dealing with big data?” will be answered. Moreover, the objective of using optimizations, techniques are illustrated. The optimization technique is employed to enhance the system performance in term of storge requirements and recognition rates.

3.1 The Hardness of computer vision

Computer vision is extremely tough. Since the computer's camera and human's eye are different. The color intensity, which is converted into a number that varies from 0 to 255, is the key factor that the camera uses to form the image. Figure 3.1 depicts the difference between the picture that a person sees and the image that a machine sees for the identical shot.

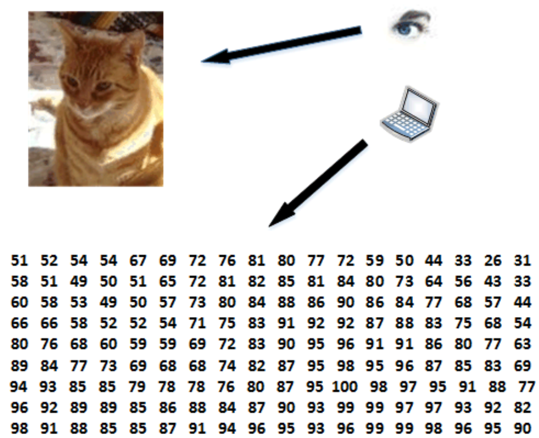


Figure 3.1 Human eyes VS Computer vision

The system of human vision readily recognizes the objects such as a cat. Human is able to distinguish between forms and people simply through continuous development and accumulated experience for many years. In addition, the human visual system collects three-dimensional objects with contextual qualities including depth, color, form, and appearance. All the features seen by the human eye disappear as soon as the image is sent to the computer in digital form. Therefore, the proposed system and similar systems are responsible for the process of distinguishing and identifying people based on the characteristics that are extracted. Therefore, computer vision is a very complex system.

3.2 Proposed system

For every recognition system, the proposed system can be divided into two modes, training mode shown in Figure 3.2.a and testing mode illustrated in Figure 3.2.b., In these two modes, part of each database will be used in the training mode and the other part will be used in the testing mode. There are three main steps for the proposed system, which are

1. Preprocessing
2. Feature extraction
3. Classification

The system is evaluated using k-Fold cross validation technique (CV).

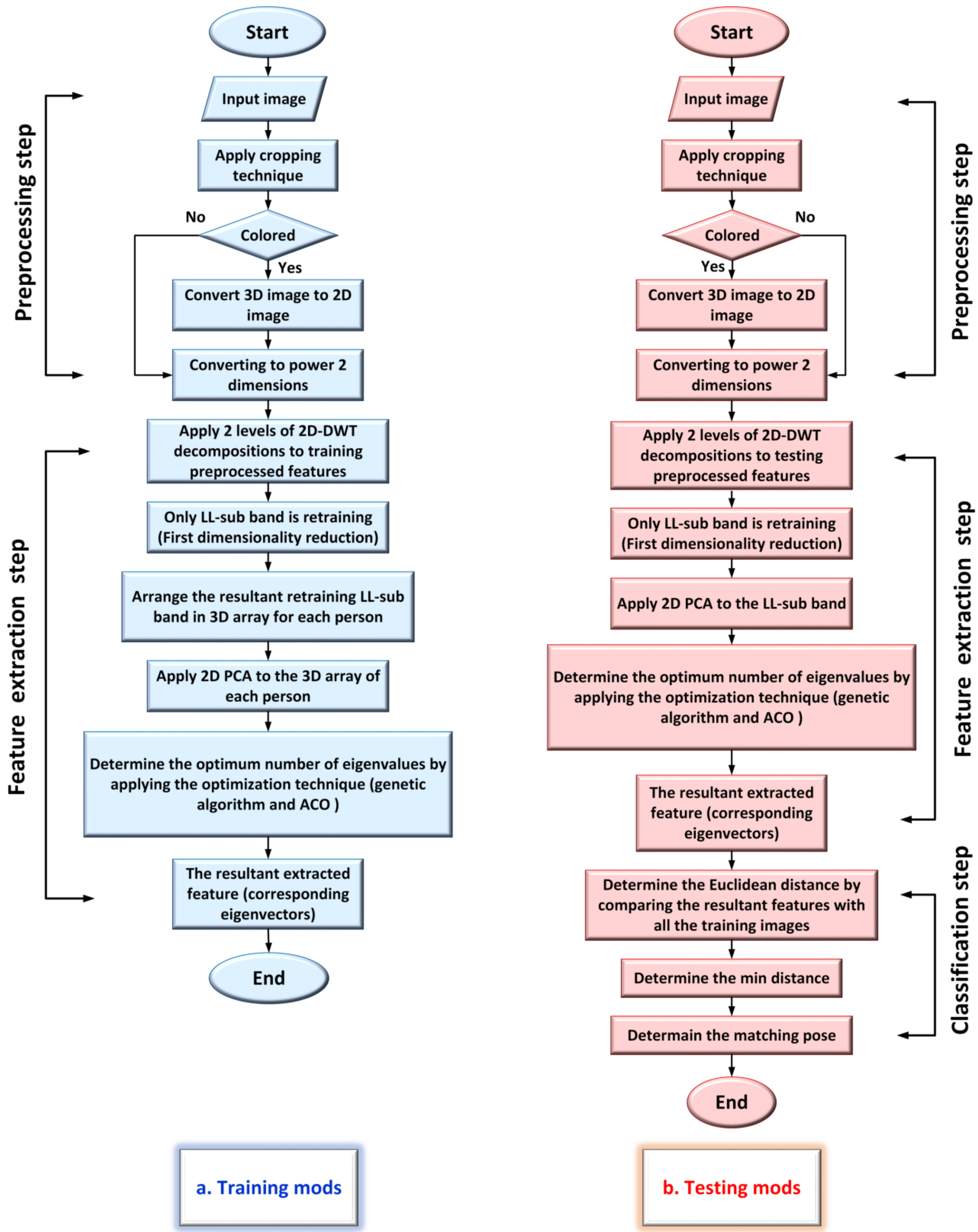


Figure 3.2 Flowchart of the proposed system

Fig.3.2.a the training mode Fig.3.2.b the testing mode

3.3 Training Mode

3.3.1 Preprocessing Steps

The preprocessing steps are applied to each pose in the databases used in this project. The preprocessing steps are:

1. Convert all colored poses of all persons in the databases used to a gray scale image to reduce the complexity.
2. Apply cropping technique in order to deal with the characteristics of facial parts instead of dealing with the whole image that containing unnecessary information, such as background, shadow, etc. Example of applying cropping is illustrated in Figure 3.3.

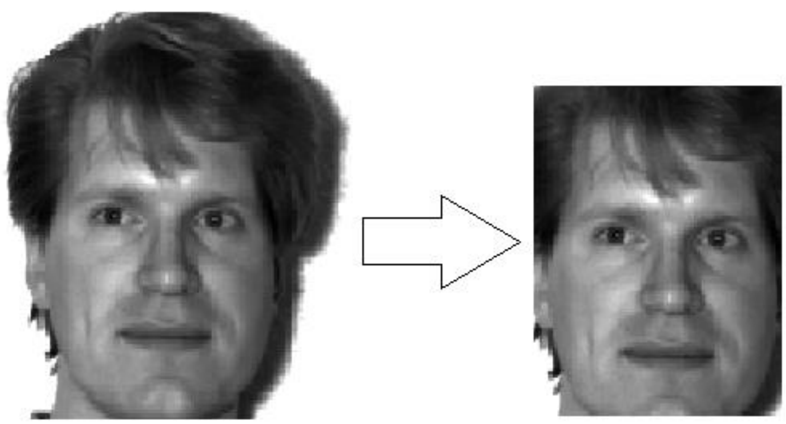


Figure 3.3 cropping

3. Resizing all poses of all persons in all databases used in this project to power of 2 dimensions, since the algorithm used, which is 2D DWT, required the image to be power of two dimensions. Figure 3.4 shows an example of converting an image to power of two dimensions. Table 3.1 shows the proposed dimensions of poses in different databases used. These dimensions are chosen after examining all the dimensions of all poses in the databases.

4. Arranging the data in a 3D array in order to reduce the computational time when execution the program.

Table 3.1: The Dimensions of The Databases

Databases	**Number**		**Actual Size**	**Dimensions After Cropping**	**Proposed Dimensions**
	Subjects	poses			
ORL	40	10	112*92	112*92	128*128
YALE	15	10	243*320	195*150	128*128
FERET	200	11	384*256	220*170	256*256
Georgia Tech.	50	15	241*181	241*181	256*256

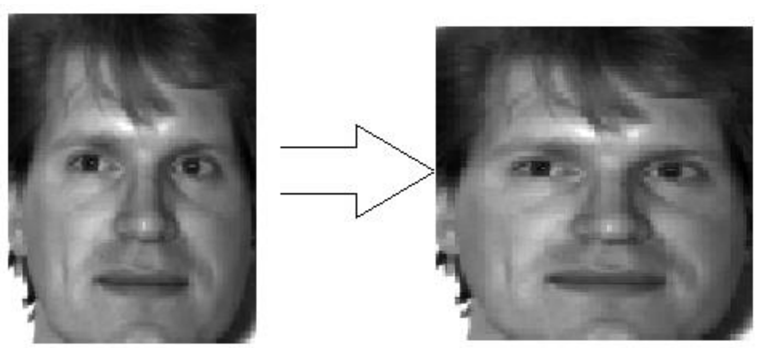


Figure 3.4 Image resizing from ‘195*150’ to ‘128*128 Pixels

3.3.2 Feature Extraction

The goal of this stage is to extract the distinctive characteristics from each pose, reduce dimension, and better representing the facial images. It is one of the most important steps in any recognition system. it consists of several steps as shown in Figure 3.2, which are:

1. The first step in the feature extraction is applying 2D DWT to each pose in the databases used. The resulted features or matrix is divided into four sub-bands as shown in Figure 3.5.

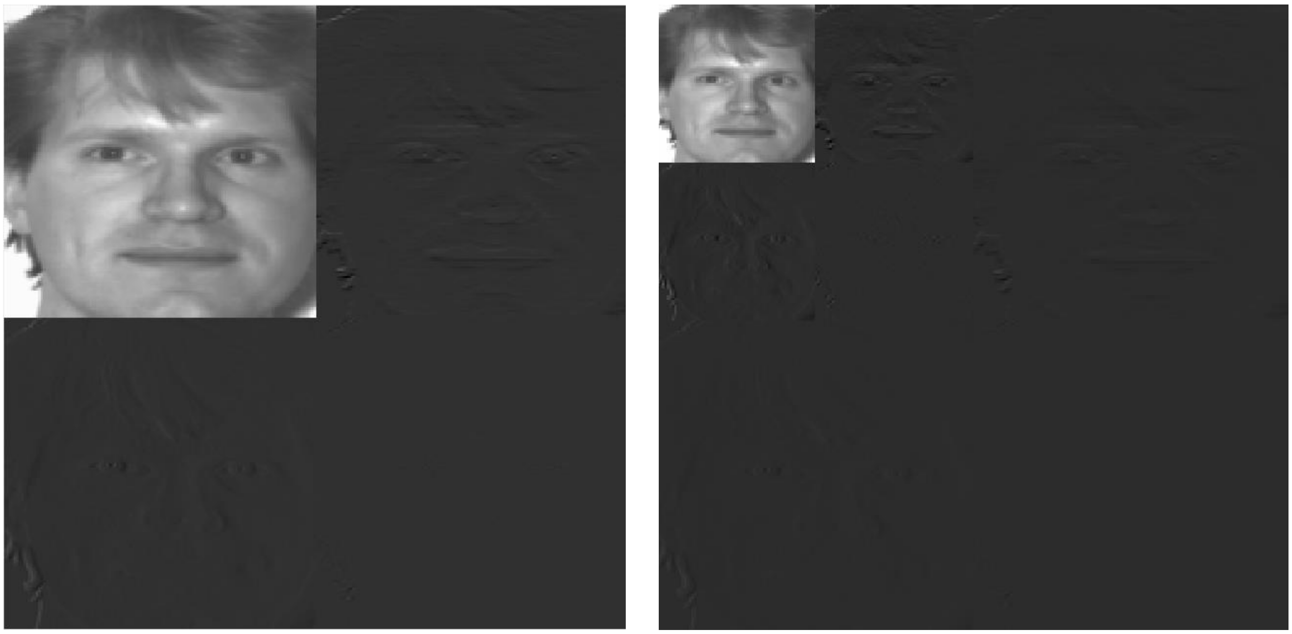
These sub-bands are

a) LL sub-band that is corresponding to low-low frequency sub-band. It is related to low pass-low pass filter. It is also called the approximation sub-band. This sub-band contain the low frequency component of the input image.
b) LH-sub-band is related to low pass-high pass filter.
c) HL-sub-band is related to high pass-low pass filter.
d) HH-sub-band is related to high pass-high pass filter. This is called a diagonal information of an image.

Since most of the information localized in LL sub-band as shown in Figure 3.5a; therefore, the LL sub-band is remaining and all other sub-bands are eliminated. In this project, 2-levels of 2D DWT decompositions are employed. Hence, the second level of 2D DWT is applied to the resultant of LL sub-band from the first level. The resultant features or matrix, shown in Figure 3.5.a, is also divided into four sub-bands. Only LL sub-band is retained, and all other sub-bands are deleted. As a results, the dimensions of the resultant features of each database used are shown in Table 3.2

Table 3.2: The dimensions of the resultant features

Databases	Original pose dimensions	Dimensions after cropping	Proposed dimensions	The resultant features of LL sub-band of 1-level of 2D DWT	The resultant features of LL sub-band of 2-levels of 2D DWT
ORL	112*92	112*92	128*128	64*64	32*32
YALE	243*320	160*150	128*128	64*64	32*32
FERET	384*256	220*170	256*256	128*128	64*64
Georgia Tech.	241*181	241*181	256*256	128*128	64*64



A- 1-level of 2D DWT B-2-levels of 2D DWT

Figure 3.5 Applying 2D DWT to pose 1 of person 1of YALE database

2. Arrange the LL sub-bands in a 3D array as shown in Figure 3.6.

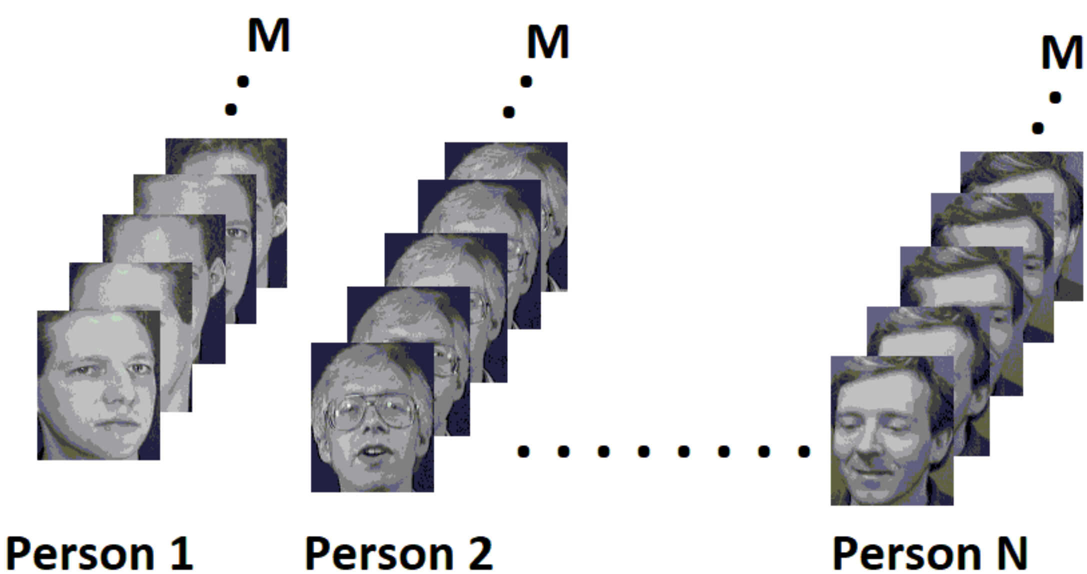


Figure 3.6. Arrangement of the LL sub-bands in a 3D array

3. Compute the total mean of all the training poses of all training persons. For example, in the ORL database, the total mean determined based on 200 poses (all training poses). The total mean of YALE database is shown in Figure 3.7. The total mean can be expressed as:

$$Total\ mean = \frac{1}{N*M}\sum_{i=1}^{N}\sum_{j=1}^{M}P(i,j) \qquad (3.1)$$

Where:

N: Number of persons.

M: Number of poses.

Figure 3.7 The total mean of the YALE database

4. Remove the mean image from every pose of each person in the databases used. Figure 3.8 shown an example of removing the mean image from the input one. This step can be expressed as.

The resultant image= input Image–total mean image (3.2)

Figure 3.8 (Image – total mean image = The resultant image).

5. Calculate the covariance matrix to the resultant image from step 4.
6. Determine the optimal number of eigenvectors that is correspond to maximum eigenvalues based on optimization algorithm (Ant colony algorithm and Genetic algorithm). Each optimization algorithm used in this work has different procedure demonstrated in Chapter Two. To make the optimization algorithms in their general sense integrated with the proposed system, it is necessary to create a fitness function. The fitness function in this project is the function responsible for calculating the accuracy in equation 3.4. Table 3.3 shows the feature matrix after optimization.

Table 3.3: The feature matrix after optimization

Databases	Dimensions of the final features extracted using GA optimization	Dimensions of the final features extracted employing ACO optimization
ORL	6*32	6*32
YALE	6*32	6*32
FERET	6*64	6*64
Georgia Tech.	6*64	6*64

As a consequence, the resultant dimensionality reduction can be expressed as:

$$\text{Dimensionality Reduction} = (1 - \left(\frac{\text{Dimensions of the final feature extracted}}{\text{Dimensions of the input pose}}\right)) * 100\%$$

(3.3)

Table 3.4 is shown the percentage of the dimensionality reduction for the four different databases.

Table 3.4: The feature matrix after optimization

Databases	Dimensions of the input pose after cropping	Dimensions of the final features	Dimensionality Reduction
ORL	112*92	6*32	98.13%
YALE	243*320	6*32	99.7%
FERET	220*170	6*64	98.9%
Georgia Tech.	241*181	6*64	99.1%

3.4 Testing Mode

In testing mode, the same preprocessing and feature extraction steps are applied to the test pose except converting the images to 3D array for each person. As shown in Figure 3.2.b

3.4.1 Classification

The Testing mode and Classification procedure can be demonstrated in the following steps, which is based on 2DPCA algorithm.

1. Compute the Euclidian distances between test feature matrix and all training feature matrices. Eq 3.4 presents the Euclidian distance between two feature matrices

$$d(p,q) = \sqrt{(p1 - q1)^2 + (p2 - q2)^2 + \cdots + (pj - qi)^2 + \cdots + (pm - qn)^2} \tag{3.4}$$

Where:

P: Training feature matrix

Q: Testing feature matrix

n: Number of training poses

m= Number of testing poses

2. Recognize the person based on the minimum Euclidian distance.
3. Calculate the recognition rate based on Eq. 3.5

$$\text{Recognition rate} = \frac{No.of\ Correctly\ identified\ poses}{Total\ number\ of\ testing\ poses} * 100\ \% \tag{3.5}$$

CHAPTER FOUR

RESULTS AND DISCUSSION

CHAPTER FOUR
RESULTS AND DISCUSSION

In this chapter, the experimental results of the system proposed are presented and discussed for different databases; namely, ORL, FERET, YALE and Georgia Tech. As mentioned, before that those databases contain different facial variations, such as illumination, facial expression, gesture, etc. To analyze the results, K - Fold CV technique is used. The results of the two-optimization algorithm are presented, which are genetic optimization algorithm and ant colony optimization algorithm.

4.1 K - Fold Cross Validation Technique Results

This technique is used in order to obtain the best possible evaluation of the performance of the proposed system. This technique was applied to all databases used in this work. For example, in ORL database that has 10 different poses, when k=2, number of training and testing poses are 200 while when k=3, the total number of training poses are 280,280,240. Therefore, number of training poses is increased when k is increased. Table 4.1 and Table 4.2 show an example of how k-fold CV is working.

Table 4.1. k-fold CV when k=2

Case 1	Case 2
Train	Test
Test	Train

Table 4.2. k-fold CV when k=3

Case 1	Case 2	Case 3
Test	Train	Train
Train	Test	Train
Train	Train	Test

4.2 The Experimental Results

In this section, the Experimental results of the proposed system is presented. The system proposed is tested based on four different databases; namely, ORL, YALE, FERET, and Georgia Tech. K - Fold CV technique is used to analyze the results of the proposed system.

To find the exact match between the testing face and training data, the Euclidean distance is employed to measure the minimum distance between the testing image and the training databases based on the following equation 4.1.

$$d(T_i, R_j) = \sum_{k=1}^{d} \parallel Y_k^i - Y_k^j \parallel \qquad (4.1)$$

Where

$T_i = test\ image\ i\ And\ \ R_j = train\ image\ j.$

$\parallel Y_k^i - Y_k^j \parallel$ Represents the distance between two components $Y_k^i\ and\ Y_k^j$.

Y_k^i Component k (Eigenvectors) of the i^{th} Test face.

Y_k^j Component k (Eigenvector)s of the j^{th} train face.

d Number of element in Eigenvectors.

4.2.1 Experimental Results of The ORL Database

The ORL Database contains 400 images from 40 distinct subjects. All poses of the ORL database were taken at different times, varying in lighting, different facial expressions (open / closed eyes, smiling / not smiling), and facial details (glasses / no glasses). The experimental results of the ORL database are shown in the Table 4.3.

Table 4.3 Accuracies based on even and odd training images

Database	Training images	Accuracy without optimization	Accuracy with GA optimization	Accuracy with ACO optimization
ORL	Even	93 %	97.5 %	97.5 %
	Odd	93 %	97.5 %	97.5 %

As shown in Table 4.3, there are two ways to evaluate the proposed system. The first one is by selecting the even poses for training and the odd ones for testing. The second way is by choosing the odd poses for training and the even ones for testing. Another way for evaluating the system proposed is by using the k-fold CV as shown in Table 4.4.

Table 4.4 The Experimental results of the ORL Database

<table>
<tr><th>K fold</th><th>train images</th><th>Test images</th><th>Accuracy without optimization</th><th>Accuracy with GA optimization</th><th>Accuracy with ACO optimization</th><th>2D-TD-MuPCA [R.Chehata, et al. 2019]</th></tr>
<tr><td rowspan="3">K=2</td><td>1 To 5</td><td>6 To 10</td><td>91 %</td><td>97.5 %</td><td>97.5 %</td><td rowspan="3">96.38%</td></tr>
<tr><td>6 To 10</td><td>1 To 5</td><td>91 %</td><td>97.5 %</td><td>97.5 %</td></tr>
<tr><td colspan="2">Average</td><td>91%</td><td>97.5%</td><td>97.5%</td></tr>
<tr><td rowspan="4">K=3</td><td>4 to 10</td><td>1 to 3</td><td>93 %</td><td>97.5 %</td><td>97.5 %</td><td rowspan="4">98.1%</td></tr>
<tr><td>1 to 3 and 8 to 10</td><td>4 to7</td><td>94 %</td><td>97.4 %</td><td>97.4 %</td></tr>
<tr><td>1 to 7</td><td>8 to 10</td><td>92 %</td><td>97.9 %</td><td>97.9 %</td></tr>
<tr><td colspan="2">Average</td><td>93%</td><td>97.6%</td><td>97.6%</td></tr>
</table>

It is obvious from Table 4.4 that the number of the training images has a significant effect on the performance of the proposed system. As the number of the training images increased (k>2), the accuracy is increased. The average, see Table 4.4, is taking to ensure the consistency. Figure 4.1 and 4.2 are shown the experimental results of the ORL database at k=2 and k=3. Figure 4.3 illustrates the recognition rates of each pose of person number 26 of the ORL database at k=2. As shown in Table 4.4, the results obtained by the system proposed are outperform those results accomplished by [55].

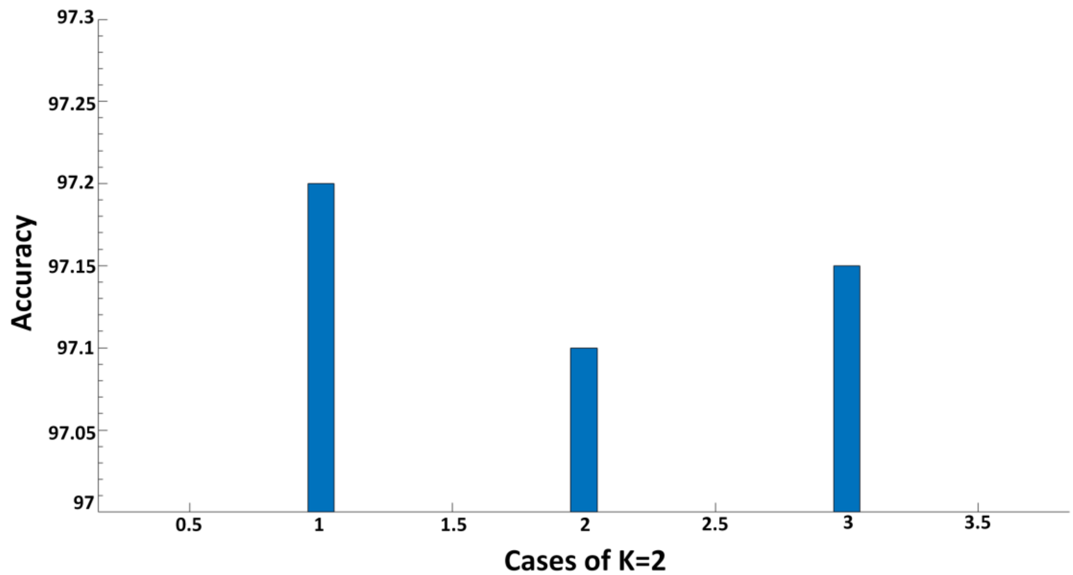


Figure 4.1 The experimental results of the ORL database at k=2.

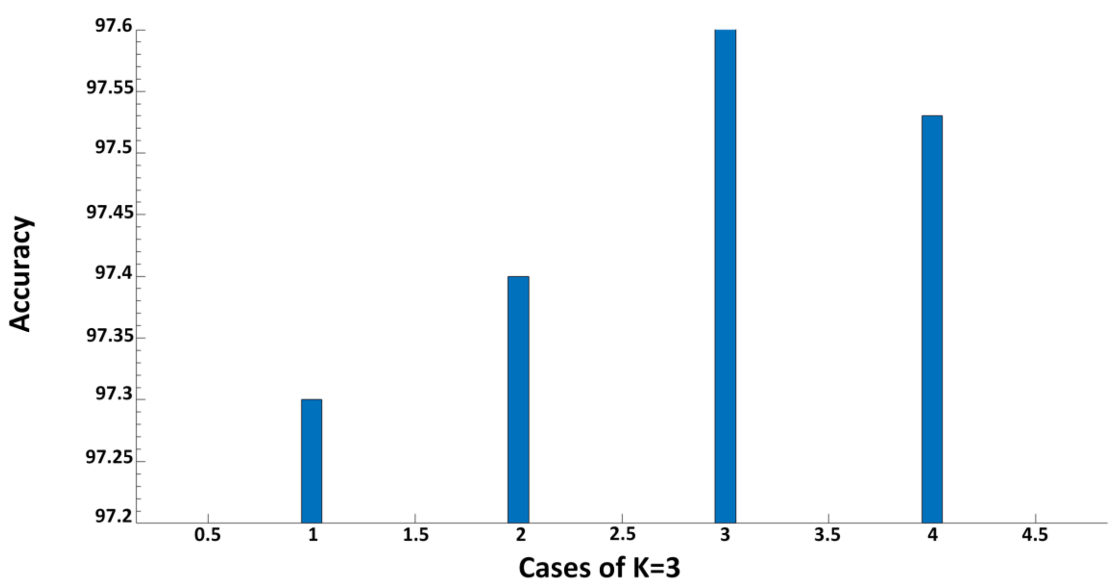


Figure 4.2 The experimental results of the ORL database at k=3.

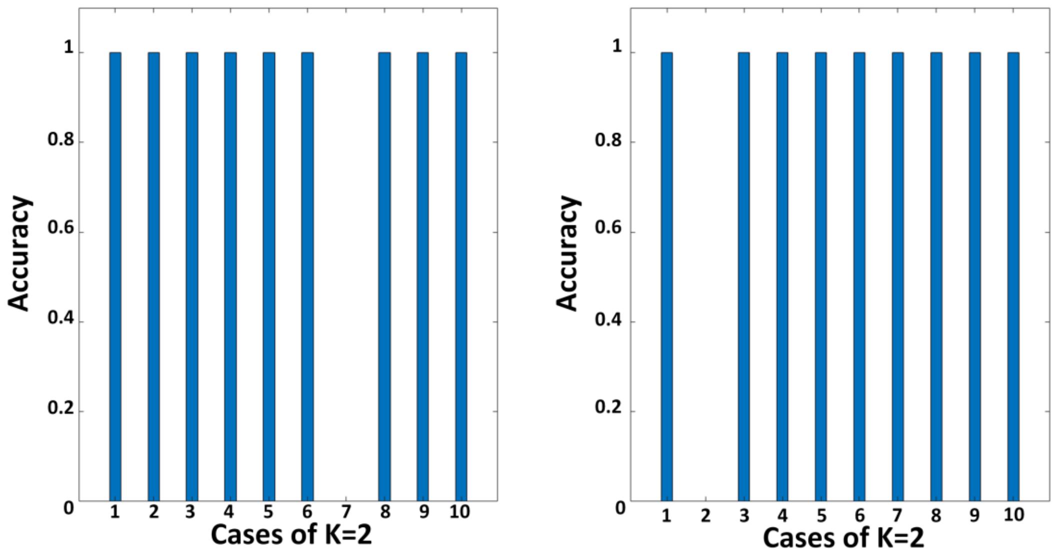


Figure 4.3 The accuracy for all poses of person 26 of the ORL database at k=2

4.2.2 Experimental Results of The YALE database

The database contains 15 different subjects, each with 11 different poses. This database has different facial variation that are mentioned in chapter one. The Experimental results of the YALE database are shown in Table 4.5 and 4.6.

The recognition rates of the YALE database based on even and odd selection are shown in Table 4.5.

Table 4.5 Accuracies based on even and odd training images

Database	Training images	Accuracy without optimization	Accuracy with GA optimization	Accuracy with ACO optimization
YALE	Even	93.3 %	97.78 %	97.78 %
	Odd	93.3 %	97.3 %	97.3 %

K-fold (CV) is also used to evaluate the system proposed and the experimental results are demonstrated in Table 4.6. As shown in Table 4.6, the recognition rates achieved by the system proposed are higher than those obtained by [55].

Table 4.6 The Experimental results of The YALE Database

K fold	train images	Test images	Accuracy without optimization	Accuracy with GA optimization	Accuracy with ACO optimization	2D-TD-MuPCA [R.Chehata, et al. 2019]
K=2	1 To 5	6 To11	90.1 %	97.87 %	97.87 %	90.29%
	6 To11	1 To 5	90.7 %	97.3 %	97.3 %	
	Average		90.4%	97.58%	97.58%	
K=3	4 to 11	1 to 3	93 %	97.8 %	97.8 %	90.7%
	1 to 3 & 8 to11	4 to 7	94%	97.8 %	97.8 %	
	1 to 7	8 to 11	95 %	97.5 %	97.5 %	

	Average	94%	97.7%	97.7%	

The system proposed exhibits the same behavior as before. As k is increased, the results are also increased (improved). Figures 4.4 and 4.5 are shown the experimental results of the YALE database at k=2 and k=3, respectively. Figure 4.6 shows the results that obtained for each pose of person 7 of the YALE database at k=2. The results obtained by the system proposed shown in Table 4.6 are higher than those achieved by [55].

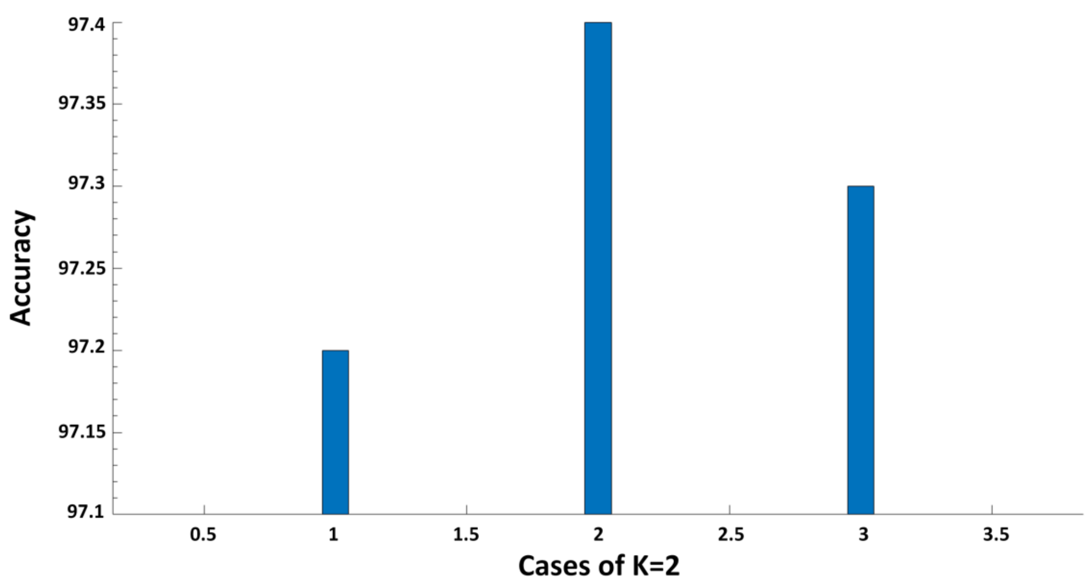


Figure 4.4 The experimental results of the YALE database at k=2.

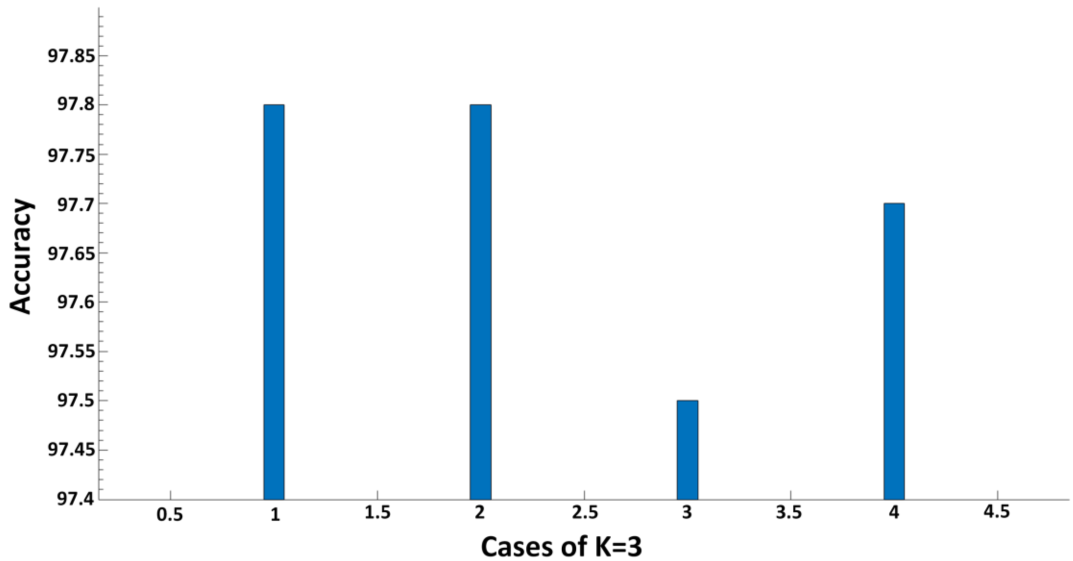


Figure 4.5 The experimental results of the YALE database at k=3.

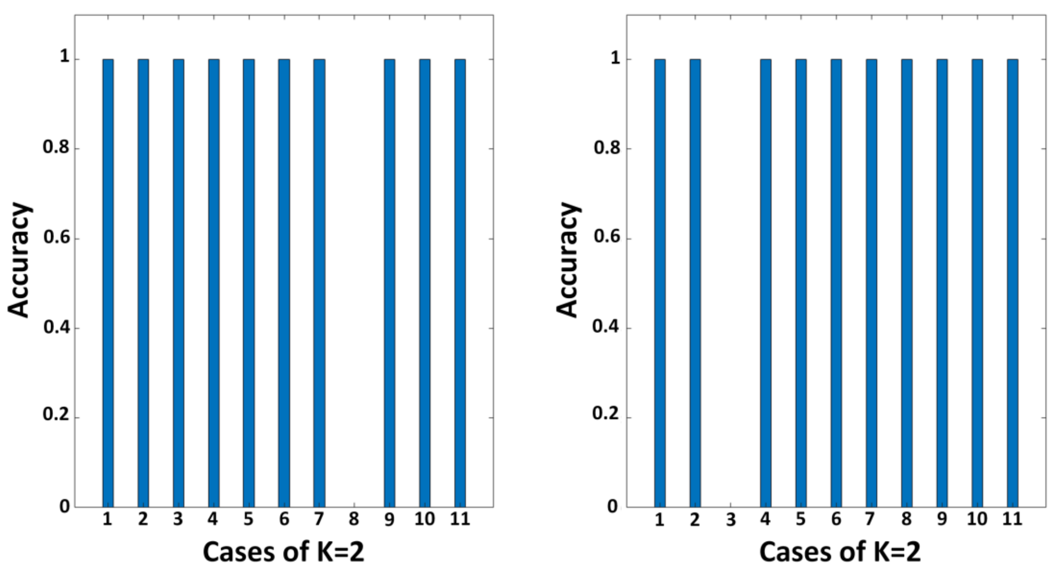


Figure 4.6 The accuracy for all poses of person 7 of the YALE database at k=2

4.2.3 Experimental Results of The FERET database

This database is a collection of images employed in the Face Recognition Technology (FERET) initiative to evaluate facial recognition systems. It was founded in 1993. The FETER database is a standard face image database that researchers may use to design algorithms and publish findings. The images for each person are labeled from 1 to 11.

The recognition rates of the FERET database based on even and odd selection are shown in Table 4.7.

Table 4.7 Accuracies based on even and odd training images

Database	Training images	Accuracy without optimization	Accuracy with GA optimization	Accuracy with ACO optimization
FERET	Even	93 %	98 %	98 %
	Odd	93 %	98 %	98 %

The experimental results of the FERET database are illustrated using the k-fold CV as shown in Table 4.8.

Table 4.8 The Experimental results of the FERET Database

<table>
<tr><th>K fold</th><th>train images</th><th>Test images</th><th>Accuracy without optimization</th><th>Accuracy with GA optimization</th><th>Accuracy with ACO optimization</th><th>2D-TD-MuPCA [R.Chehata, et al. 2019]</th></tr>
<tr><td rowspan="3">K=2</td><td>1 To 5</td><td>6 To 11</td><td>94.4 %</td><td>98.2 %</td><td>98.2 %</td><td rowspan="3">90.75%</td></tr>
<tr><td>6 To 11</td><td>1 To 5</td><td>94.27 %</td><td>98.6 %</td><td>98.6 %</td></tr>
<tr><td colspan="2">Average</td><td>94.3%</td><td>98.4%</td><td>98.4%</td></tr>
<tr><td rowspan="4">K=3</td><td>4 to 11</td><td>1 to 3</td><td>90 %</td><td>98.9 %</td><td>98.9 %</td><td rowspan="4">96.5%</td></tr>
<tr><td>1 to 3 & 8 to 11</td><td>4 to 7</td><td>90.2 %</td><td>98.8 %</td><td>98.8 %</td></tr>
<tr><td>1 to 7</td><td>8 to 11</td><td>91.6 %</td><td>98.2 %</td><td>98.2 %</td></tr>
<tr><td colspan="2">Average</td><td>90.6%</td><td>98.6%</td><td>98.6%</td></tr>
</table>

The system proposed shows the same performance compared to its performance with other databases. The results accomplished by proposed system demonstrated in Table 4.8 are higher than those results achieved by [55]. As mentioned before, the value of k has a direct effect on the system performance. Figures 4.7 and 4.8 are demonstrating the experimental results of the FERET database at k=2 and k=3, respectively. Figure 4.9 illustrates the results that obtained for each pose of person 12 of FERET database at k=2.

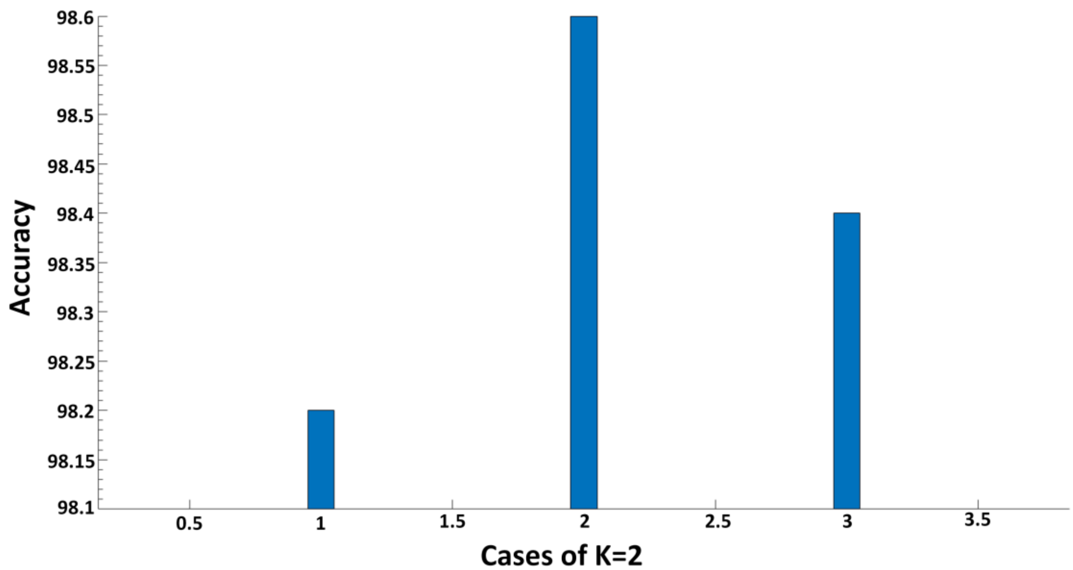


Figure 4.7 The experimental results of the FERET database at k=2.

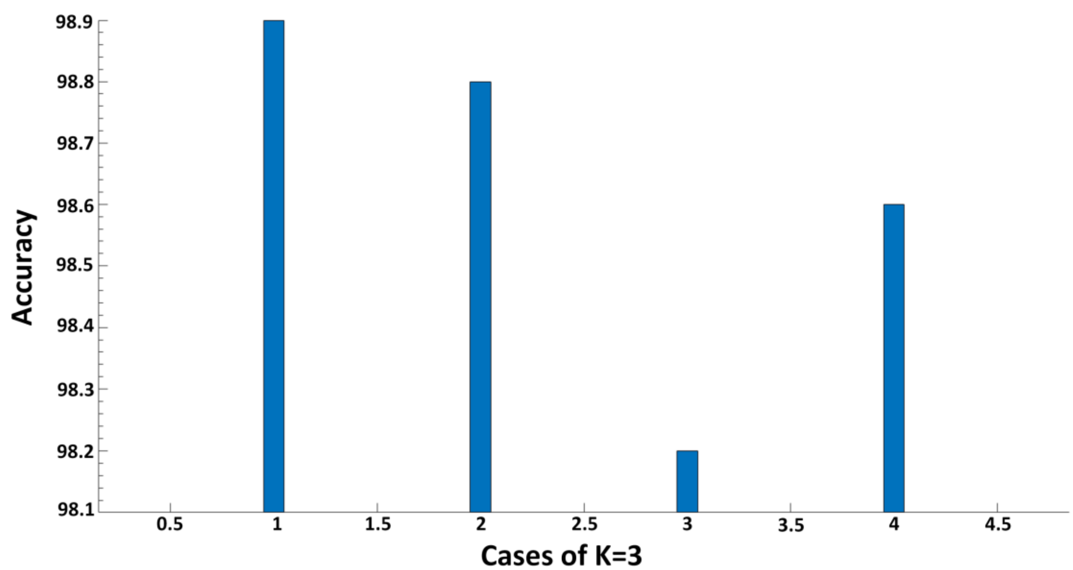


Figure 4.8 The experimental results of the FERET database at k=3.

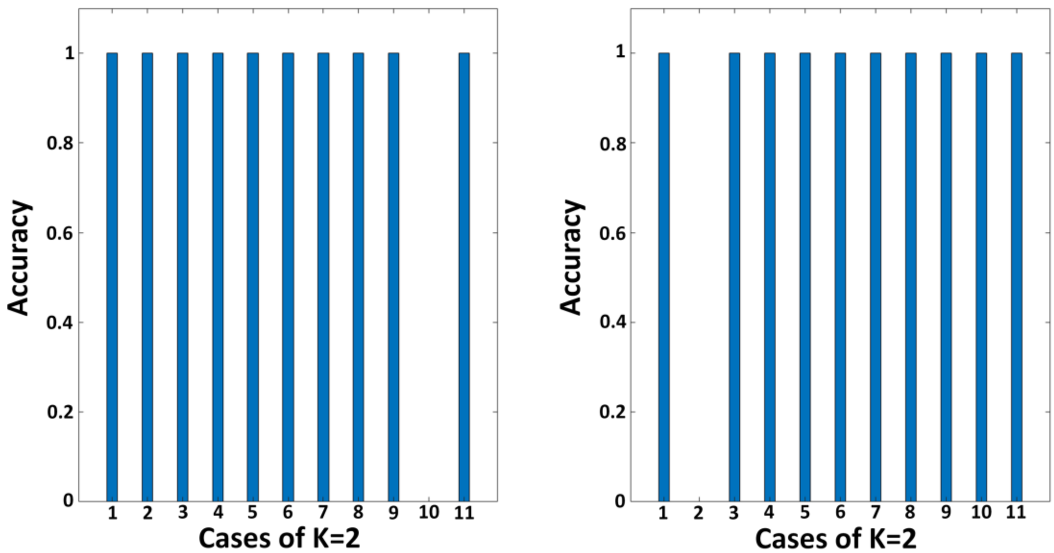


Figure 4.9 The accuracy for all poses of person 12 of the FERET database at k=2.

4.2.4 Experimental Results of The Georgia Tech. Database

This database contains 50 subjects, each with 15 different poses. These poses have frontal and/or tilted faces with different facial expressions, lighting conditions, and scale. All images have the same dimensions, which are 241*181.

The recognition rates of the Georgia Tech. database based on even and odd selection are reported in Table 4.9.

Table 4.9 Accuracies based on even and odd training images

<table>
<tr><th>Database</th><th>Training images</th><th>Accuracy without optimization</th><th>Accuracy with GA optimization</th><th>Accuracy with ACO optimization</th></tr>
<tr><td rowspan="2">Georgia Tech</td><td>Even</td><td>90%</td><td>94.5%</td><td>94.5%</td></tr>
<tr><td>Odd</td><td>90%</td><td>94.5%</td><td>94.5%</td></tr>
</table>

The excremental results of the k-fold CV are presented in Table 4.10.

Table 4.10 The Experimental results of the Georgia Tech Database

<table>
<tr><th>K fold</th><th>train images</th><th>Test images</th><th>Accuracy without optimization</th><th>Accuracy with GA optimization</th><th>Accuracy with ACO optimization</th><th>2D-TD-MuPCA [R.Chehata, et al. 2019]</th></tr>
<tr><td rowspan="3">K=2</td><td>1 To 7</td><td>8 To 15</td><td>90.4 %</td><td>94.2 %</td><td>94.2 %</td><td rowspan="7">92%</td></tr>
<tr><td>8 To 15</td><td>1 To 7</td><td>90.27 %</td><td>94.6 %</td><td>94.6 %</td></tr>
<tr><td colspan="2">Average</td><td>90.3%</td><td>94.4%</td><td>94.4%</td></tr>
<tr><td rowspan="4">K=3</td><td>6 to 15</td><td>1 to 5</td><td>90 %</td><td>94.9 %</td><td>94.9 %</td></tr>
<tr><td>1 to 5 & 11 to 15</td><td>6 to 10</td><td>90.2 %</td><td>94.8 %</td><td>94.8 %</td></tr>
<tr><td>1 to 10</td><td>11 to 15</td><td>90.6 %</td><td>94.2 %</td><td>94.2 %</td></tr>
<tr><td colspan="2">Average</td><td>90.26%</td><td>94.6%</td><td>94.6%</td></tr>
</table>

As clearly noticed from table 4.10 that the results of the system proposed exhibit the same altitude as before. As before, the performance of the system proposed is out performed the one that presented by [55]. Figures 4.10 and 4.11 are shown the experimental results of Georgia Tech database at k=2 and k=3, respectively. Figure 4.12 displays the results that obtained for each pose of person 2 of Georgia Tech database at k=2.

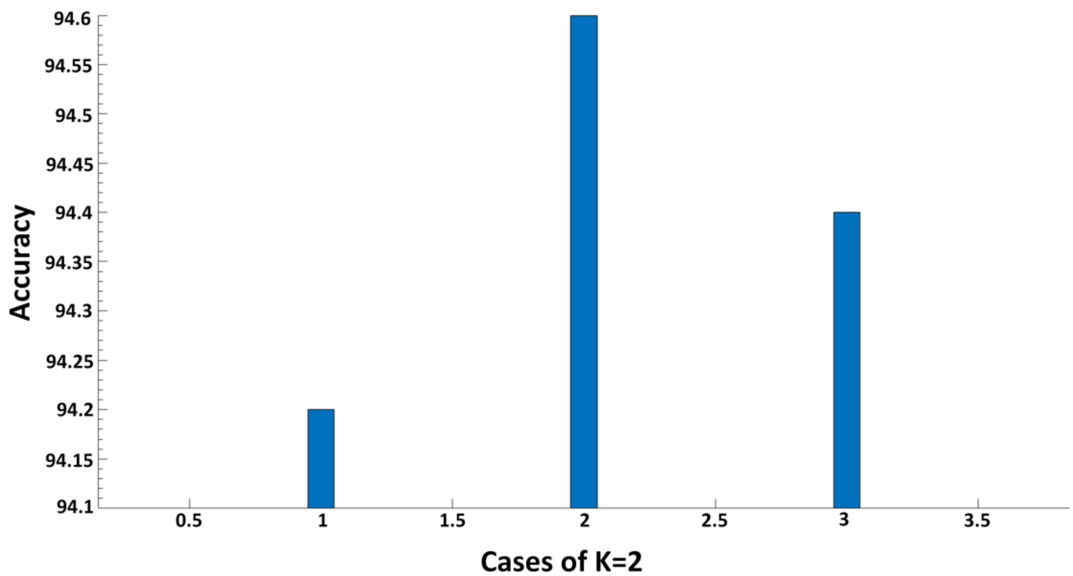


Figure 4.10 The experimental results of the Georgia Tech database at k=2

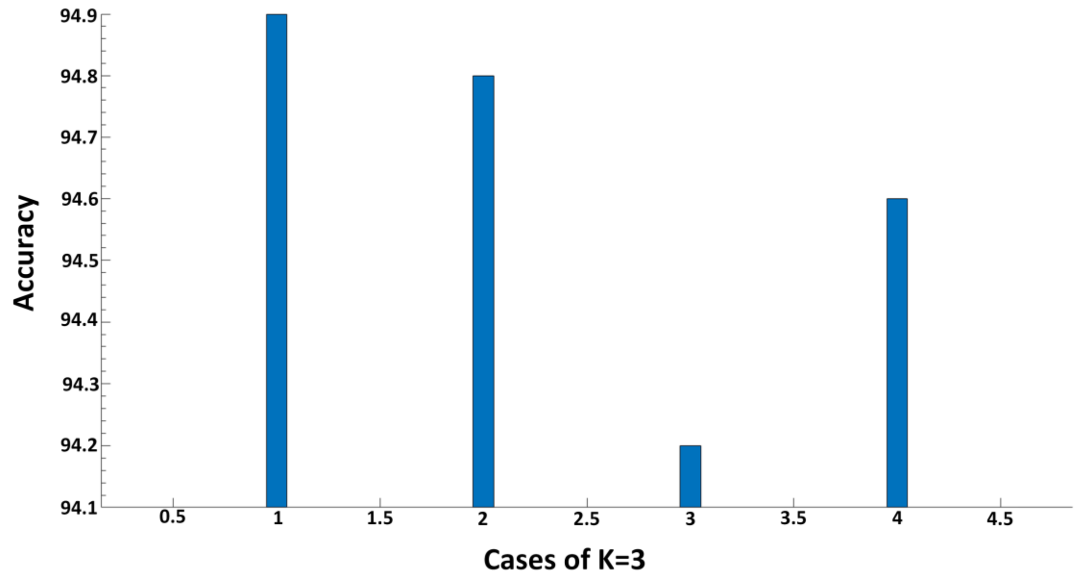


Figure 4.11 The experimental results of the Georgia Tech database at k=3.

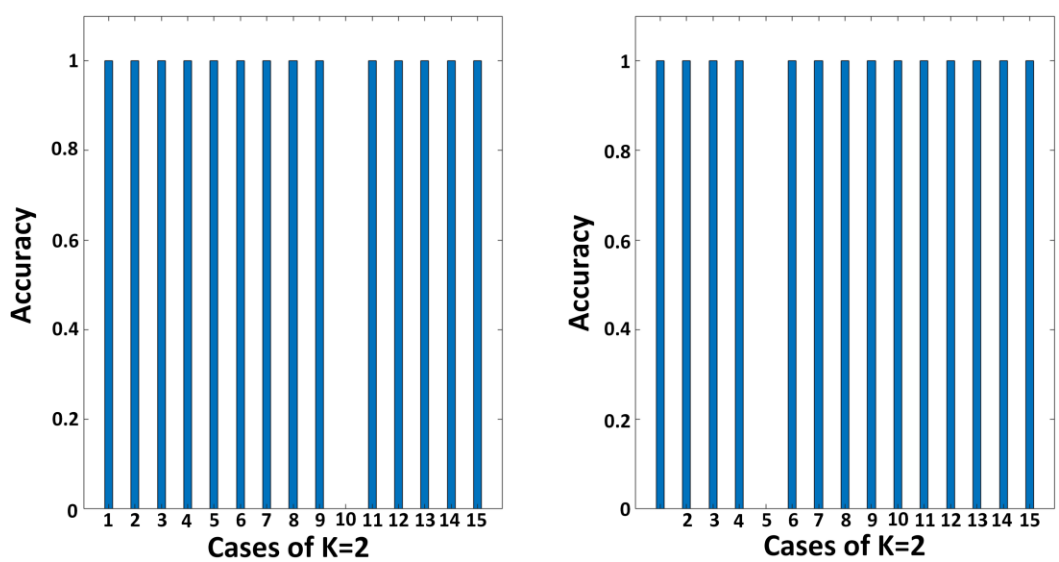


Figure 4.12 The accuracy for all poses of person 2 of the Georgia Tech database at k=2

Figure 4.13 demonstrates the relationship between the testing and the training poses. It is proven from the Figure 4.13 that the recognition rates are increased when the number of the training poses is increased for all types of the databases.

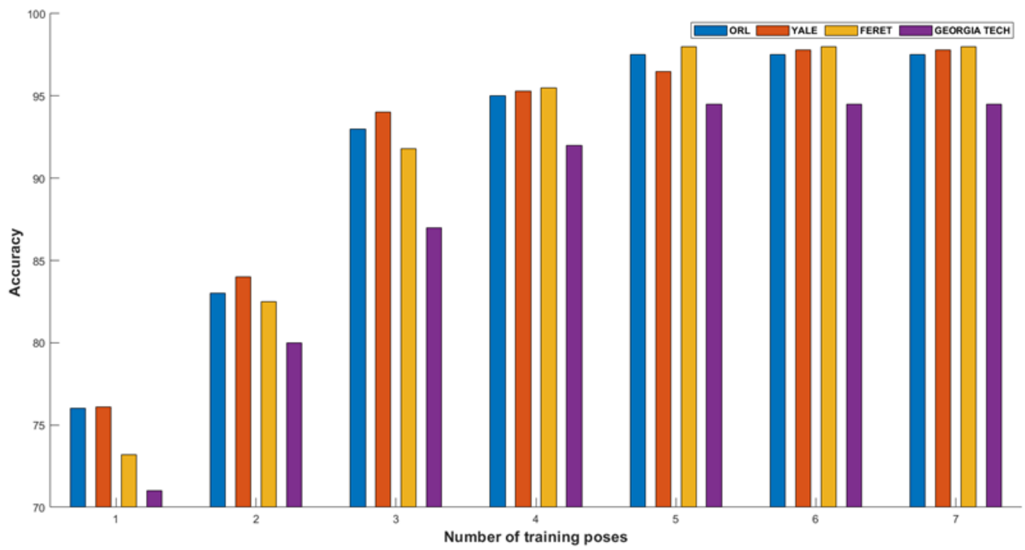


Figure 4.13 Relation between the accuracy and the number of the training poses for all databases.

CHAPTER FIVE

CONCLUSIONS AND FUTURE WORKS

CHAPTER FIVE

CONCLUSIONS AND FUTURE WORKS

5.1 Conclusions

- The proposed system depends on the use of the following two techniques together: Two dimensions Principal Component Analysis (2D-PCA) and Two dimensions Discrete Wavelet Transform (2D-DWT).
- LL- sub-band out of two levels of 2D-DWT decompositions were used to each pose.
- 2D PCA was used on LL sub-band of the second level of 2D DWT to get better facial representation.
- Results achieved, were improved using optimization techniques, Genetic and ACO algorithms.
- The recognition rates achieved were improved compared with those reported by [55]
- The K-fold CV technique was used for the fair analysis of the results.
- Recognition accuracies at k=2 are realized, were 97.5%, 97.58%, 98.4%, and 94.4% for ORL, YALE, FERET, and Georgia Tech. databases, respectively.
- Recognition accuracies at k=3 are realized, were 97.6%, 97.7%, 98.6%, and 94.6% for ORL, YALE, FERET, and Georgia Tech. databases, respectively.

5.2 Future Works

- Develop a system capable of calculating the angle of rotation of the face to apply the procedure of face recognition and compare it with faces that have an asymptotic angle in the database instead of comparing it with all faces. This proposal contributes to increasing the identification speed, which in turn is considered one of the important factors in these systems.

- The use of high dynamic range (HDR) technology can contribute to increased accuracy. So, it's possible to develop a recognition system based on HDR images

References

[1] Kumar, Santosh, and Sanjay Kumar Singh. "Cattle recognition: A new frontier in visual animal biometrics research." Proceedings of the national academy of sciences, india section A: physical sciences 90.4 (2020): 689-708.

[2] Winston, J. Jenkin, and D. Jude Hemanth. "A comprehensive review on iris image-based biometric system." Soft Computing 23.19 (2019): 9361-9384.

[3] Alay, Nada, and Heyam H. Al-Baity. "Deep learning approach for multimodal biometric recognition system based on fusion of iris, face, and finger vein traits." Sensors 20.19 (2020): 5523.

[4] Lin, Chenhao, and Ajay Kumar. "Matching contactless and contact-based conventional fingerprint images for biometrics identification." IEEE Transactions on Image Processing 27.4 (2018): 2008-2021.

[5] Qin, Xiaoqian, Dakun Liu, and Dong Wang. "A literature survey on kinship verification through facial images." Neurocomputing 377 (2020): 213-224.

[6] Choudhury, Ziaul Haque. "Encryption and encoding of facial images into quick response and high capacity color 2d code for biometric passport security system." arXiv preprint arXiv:2203.15738 (2022).

[7] Choudhury, Z. H. (2022). Encryption and encoding of facial images into quick response and high capacity color 2d code for biometric passport security system. arXiv preprint arXiv:2203.15738.

[8] Khan, S., Javed, M. H., Ahmed, E., Shah, S. A., & Ali, S. U. (2019, March). Facial recognition using convolutional neural networks and implementation on smart glasses. In 2019 International Conference on Information Science and Communication Technology (ICISCT) (pp. 1-6). IEEE.

[9] SEAL, Sayan, et al. An approach towards development of automated attendance system using face detection and recognition. In: 2020 11th IEEE Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON). IEEE, 2020. p. 0333-0340.

[10] Adjabi, Insaf, et al. "Past, present, and future of face recognition: A review." Electronics 9.8 (2020): 1188.

[11] Dino, Hivi Ismat, and Maiwan Bahjat Abdulrazzaq. "Facial expression classification based on SVM, KNN and MLP classifiers." 2019 International Conference on Advanced Science and Engineering (ICOASE). IEEE, 2019.
[12] OMOYIOLA, Bayo Olushola. Overview of biometric and facial recognition techniques. IOSR journal of computer engineering (IOSRJCE), 2018, 20.4: 1-5.
[13] Mehta, Dhwani, Mohammad Faridul Haque Siddiqui, and Ahmad Y. Javaid. "Facial emotion recognition: A survey and real-world user experiences in mixed reality." Sensors 18.2 (2018): 416.
[14] Ilyas, Naveed, Ahsan Shahzad, and Kiseon Kim. "Convolutional-neural network-based image crowd counting: Review, categorization, analysis, and performance evaluation." Sensors 20.1 (2019): 43.
[15] Setialana, P., Jati, H., Wardani, R., Indrihapsari, Y., & Norwawi, N. M. (2021). Intelligent Attendance System with Face Recognition using the Deep Convolutional Neural Network Method. In Journal of Physics: Conference Series (Vol. 1737, No. 1, p. 012031). IOP Publishing.
[16] JIGNESH CHOWDARY, G., et al. Face mask detection using transfer learning of inceptionv3. In: International Conference on Big Data Analytics. Springer, Cham, 2020. p. 81-90.
[17] Wang, Z., Wang, G., Huang, B., Xiong, Z., Hong, Q., Wu, H., ... & Liang, J. (2020). Masked face recognition dataset and application. arXiv preprint arXiv:2003.09093.
[18] Ramachandra, Raghavendra, and Christoph Busch. "Presentation attack detection methods for face recognition systems: A comprehensive survey." ACM Computing Surveys (CSUR) 50.1 (2017): 1-37.
[19] Saraswathi, M., and Dr S. Sivakumari. "Evaluation of PCA and LDA techniques for Face recognition using ORL face database." Int. J. Comput. Sci. Inf. Technol 6.1 (2015): 810-813.
[20] Sharif, Muhammad, et al. "Face Recognition: A Survey." Journal of Engineering Science & Technology Review 10.2 (2017).
[21] Zhao, XueMei, and ChengBing Wei. "A real-time face recognition system based on the improved LBPH algorithm." 2017 IEEE 2nd International Conference on Signal and Image Processing (ICSIP). IEEE, 2017.
[22] Hashemi, Seyyed Mohammad Reza, and Maryam Faridpour.

"Evaluation of the algorithms of face identification." 2015 2nd International Conference on Knowledge-Based Engineering and Innovation (KBEI). IEEE, 2015.

[23] Parivazhagan, A., and A. Brintha Therese. "Face detection cum recognition system using novel techniques for human authentication." International Journal of Biometrics 10.4 (2018): 315-333.

[24] Bendale, Abhijit, and Terrance Boult. "Towards open world recognition." Proceedings of the IEEE conference on computer vision and pattern recognition. 2015.

[25] Kortli, Yassin, et al. "Face recognition systems: A survey." Sensors 20.2 (2020): 342.

[26] Masi, Iacopo, et al. "Do we really need to collect millions of faces for effective face recognition?." European conference on computer vision. Springer, Cham, 2016.

[27] Botta, Alessio, et al. "Integration of cloud computing and internet of things: a survey." Future generation computer systems 56 (2016): 684-700.

[28] Li, Yuezun, Ming-Ching Chang, and Siwei Lyu. "In ictu oculi: Exposing ai generated fake face videos by detecting eye blinking." arXiv preprint arXiv:1806.02877 (2018).

[29] Jack, Rachael E., et al. "Four not six: Revealing culturally common facial expressions of emotion." Journal of Experimental Psychology: General 145.6 (2016): 708.

[30] Priadana, Adri, and Muhammad Habibi. "Face detection using haar cascades to filter selfie face image on instagram." 2019 International Conference of Artificial Intelligence and Information Technology (ICAIIT). IEEE, 2019.

[31] Wang, Wei, et al. "Recurrent face aging." Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.

[32] Deng, Jiankang, et al. "Arcface: Additive angular margin loss for deep face recognition." Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2019.

[33] Wang, Zhongyuan, et al. "Masked face recognition dataset and application." arXiv preprint arXiv:2003.09093 (2020).

[34] Singh, Shilpi, and S. V. A. V. Prasad. "Techniques and challenges of face recognition: A critical review." Procedia computer science 143

(2018): 536-543.
[35] Ding, Changxing, and Dacheng Tao. "A comprehensive survey on pose-invariant face recognition." ACM Transactions on intelligent systems and technology (TIST) 7.3 (2016): 1-42.
[36] Voulodimos, Athanasios, et al. "Deep learning for computer vision: A brief review." Computational intelligence and neuroscience 2018 (2018).
[37] Amos, Brandon, Bartosz Ludwiczuk, and Mahadev Satyanarayanan. "Openface: A general-purpose face recognition library with mobile applications." CMU School of Computer Science 6.2 (2016): 20.
[38] Dutta, Aastha. "Fourier transform infrared spectroscopy." Spectroscopic methods for nanomaterials characterization (2017): 73-93.
[39] Liao, Xin, Kaide Li, and Jiaojiao Yin. "Separable data hiding in encrypted image based on compressive sensing and discrete fourier transform." Multimedia Tools and Applications 76.20 (2017): 20739-20753.
[40] Ri, Shien, et al. "Comparative study of sampling moiré and windowed Fourier transform techniques for demodulation of a single-fringe pattern." Applied optics 57.36 (2018): 10402-10411.
[41] Prasad, Lakshman, and Sundararaja S. Iyengar. Wavelet analysis with applications to image processing. CRC press, 2020.
[42] Gogolewski, Damian. "Influence of the edge effect on the wavelet analysis process." Measurement 152 (2020): 107314.
[43] Daubechies, Ingrid, and Stephane Maes. "A nonlinear squeezing of the continuous wavelet transform based on auditory nerve models." Wavelets in medicine and biology. Routledge, 2017. 527-546.
[44] He, Runnan, et al. "Automatic detection of atrial fibrillation based on continuous wavelet transform and 2D convolutional neural networks." Frontiers in physiology 9 (2018): 1206.
[45] Chen, Duo, et al. "A high-performance seizure detection algorithm based on Discrete Wavelet Transform (DWT) and EEG." PloS one 12.3 (2017): e0173138.
[46] Zhang, Dengsheng. "Wavelet transform." Fundamentals of Image Data Mining. Springer, Cham, 2019. 35-44.
[47] Wang, Xia, et al. "A probabilistic multimodal optimization algorithm based on Buffon principle and Nyquist sampling theorem for noisy

environment." Applied Soft Computing 104 (2021): 107068.
[48] Ramadhan, Afrah, Firas Mahmood, and Atilla Elci. "Image denoising by median filter in wavelet domain." arXiv preprint arXiv:1703.06499 (2017).
[49] Kurita, Takio. "Principal component analysis (PCA)." Computer Vision: A Reference Guide (2019): 1-4.
[50] Xiao, Xiaolin, and Yicong Zhou. "Two-dimensional quaternion PCA and sparse PCA." IEEE transactions on neural networks and learning systems 30.7 (2018): 2028-2042.
[51] Dorigo, Marco, and Thomas Stützle. "Ant colony optimization: overview and recent advances." Handbook of metaheuristics (2019): 311-351.
[52] Tran-Ngoc, Hoa, et al. "Model updating for Nam O bridge using particle swarm optimization algorithm and genetic algorithm." Sensors 18.12 (2018): 4131.
[53] Draisma, Jan, et al. "The Euclidean distance degree of an algebraic variety." Foundations of computational mathematics 16.1 (2016): 99-149.
[54] Valavi, Roozbeh, et al. "blockCV: An r package for generating spatially or environmentally separated folds for k-fold cross-validation of species distribution models." bioRxiv (2018): 357798.

الخلاصة

تم اقتراح نهج جديد للتعرف على الوجه باستخدام التحويل المويجي المنفصل ثنائي الأبعاد (2DDWT) وتحليل المكون الرئيسي ثنائي الأبعاد (2DPCA).المعالجة المسبقة ، واستخراج الميزات ، والتصنيف هي المراحل الرئيسية الثلاث في هذا النظام. تم توضيف تقنية قص الصورة لابعاد مناسبة في هذه المرحلة. في استخراج الميزة ، يتم تطبيق مستويين من تحلل DWT ثنائي الأبعاد على الصور المعالجة لتقليل الأبعاد واستخراج الميزات. يتم ضغط ميزات D DWT٢ الناتجة باستخدام 2DPCA. يتم استخدام تقنيات التحسين ، التحسين الوراثي ومستعمرة النمل (ACO) ، لتحسين دقة النظام. يتم تقييم الخوارزمية المقترحة باستخدام أربع قواعد بيانات ، وهي ORL و YALE و FERET و Georgia Tech. التي لها اختلافات في الوجه ، مثل تعابير الوجه ، والإضاءة ، والدوران ، وما إلى ذلك ، ثم يتم تحليل النتائج باستخدام K-fold Cross Validation (CV). تظهر النتائج أن النهج المقترح يحسن معدلات التعرف ويقلل من متطلبات التخزين مقارنة بالطرق الحالية.

جمهورية العراق
وزارة التعليم العالي والبحث العلمي
جامعة بابل كلية الهندسة / قسم الهندسة الكهربائية

توظيف تقنية فعالة تعتمد على التحويل المويجي المنفصل و تحليل المكون الرئيسي لغرض التعرف على الوجو ه

رسالة

مقدمة الى كلية الهندسة

كجزء من متطلبات نيل درجة الدبلوم العالي في جامعة بابل

في كلية الهندسة/ قسم الهندسة الكهربائية/الكترونيك واتصالات

من قبل:

حوراء ليث فلاح حسن

إشراف:

د٠احمد قاسم الذهب

١٤٤٣هـ ٢٠٢٢ م