

# **Wavelet Based Face Recognition in the Presence of Illumination Variation**

**Pooya Ferdosipour**

Submitted to the  
Institute of Graduate Studies and Research  
in partial fulfillment of the requirements for the degree of

Master of Science  
in  
Electrical and Electronic Engineering

Eastern Mediterranean University  
January 2016  
Gazimağusa, North Cyprus

Approval of the Institute of Graduate Studies and Research

---

Prof. Dr. Cem Tanova  
Acting Director

I certify that this thesis satisfies the requirements as a thesis for the degree of Master of Science in Electrical and Electronic Engineering.

---

Prof. Dr. Hasan Demirel  
Chair, Department of Electrical  
and Electronic Engineering

We certify that we have read this thesis and that in our opinion it is fully adequate in scope and quality as a thesis for the degree of Master of Science in Electrical and Electronic Engineering

---

Prof. Dr. Şener Uysal  
Supervisor

---

Examining Committee

1. Prof. Dr. Hasan Demirel

2. Prof. Dr. Hüseyin Özkaramanli

3. Prof. Dr. Şener Uysal

## ABSTRACT

As a context of biometrics, significant advances have been made in face recognition during the recent decades. Face recognition is one of the most successful applications of image analysis. The accuracy of automated face recognition is greatly affected by varying in lighting between probe and train images. Difference in lighting condition is one of the difficulties in automated face recognition systems. Histogram equalization technique is widely used to diminish the desired effect of different illumination condition between probe and train images by normalizing variation in illumination. Experiments show that normalizing images that has good lighting condition could lead to an increase in recognition error.

Wavelet transform, that is well-known as a multiresolution method, is used in features extracting phase. The multiresolution property of wavelet transform is used in extracting feature leading us to have facial feature descriptors at different scales and frequencies. This thesis presents image quality based technique which is measured in terms of luminance to overcome the disadvantage of varying lighting condition to increase the accuracy of face recognition method. 10-fold cross variation is used to investigate the effect of data selection on classification algorithm. At the end, results are compared to investigate the best method for automated face recognition when illumination variation exists.

**Keywords:** Biometrics, illumination, Wavelets transform (WT), face recognition, feature extraction

## ÖZ

Son yıllarda, bir biyometri alanı olarak yüz tanıma konusunda kayda değer gelişmeler meydana gelmiştir. Yüz tanıma, görüntü işlemenin en başarılı uygulamalarından biridir. Otomatikleştirilmiş yüz tanımanın hassasiyeti, araştırma ile eğitime görüntüleri arasında ışıklandırma değişimlerinden büyük oranda etkilenmektedir. Işıklandırma koşullarındaki farklılıklar otomatikleştirilmiş yüz tanıma sistemlerinin zorluklarından biridir. Histogram eşitleme tekniği, aydınlatma farklılıkları normalleştirilerek araştırma ile eğitim görüntüleri arasındaki ışıklandırma farklılıklarının istenmeyen etkilerinin azaltılması için geniş çaplı bir kullanıma sahiptir. Yapılan deneyler iyi ışıklandırma koşullarına sahip olan normalleştirme görüntülerinin tanıma hatasının yükselmesine neden olabileceklerini göstermektedir.

Çokçözünürlüklü bir yöntem olarak bilinmekte olan Dalgacık Dönüşümü, özellik ayrıştırma aşamasında kullanılmaktadır. Özellik ayrıştırma aşamasında kullanılan dalgacık dönüşümünün çokçözünürlüklü olma özelliği, farklı ölçek ve aralıklarda yüz özellik tanımlayıcılarına sahip olmamıza olanak sağlamaktadır. Bu tez çalışması, yüz tanıma yönteminin hassasiyetinin artırılması amacıyla değişken ışıklandırma koşullarından kaynaklanan dezavantajları ortadan kaldırmak üzere parlaklık cinsinden ölçülen görüntü kalitesine dayalı bir teknik sunmaktadır. Veri seçiminin sınıflandırma algoritması üzerindeki etkisinin araştırılması için 10-katlı çapraz doğrulama kullanılmıştır. Son olarak, ışıklandırmada değişiklikler bulunduğunda otomatikleştirilmiş yüz tanıma için en iyi yöntemin seçilmesi amacıyla sonuçlar karşılaştırılmıştır.

**Anahtar Kelimeler:** Biyometri, aydınlatma, Wavelet dönüştürmek, yüz tanıma.

Dedicated to

A world without border, racism and war.

## **ACKNOWLEDGMENT**

It is a great privilege for me to have been associated with Prof. Dr. Sener Uysal my guide, during the research work. It is with great pleasure that I express my deep sense of gratitude to him for his valuable guidance, constant encouragement, motivation, support and patience throughout this work. I express my gratitude to Prof. Dr. Hasan Demirel chair of Department of Electrical and Electronic Engineering for his constant encouragement, and support during the completion of this work.

I owe a lot to my parents. Their moral support, encouragement and blessings always helped me in completion of this work.

I would like to express my sincere gratitude to Mr. Bashir Sadeghi for the continuous support of my MS study and related research, for his patience, motivation, and immense knowledge. I could not have imagined having a better friend and mentor for my MS study.

Finally, I would like to thank Mr. Iman Beheshty and Mr. Kehinde Ehineni my classmate, and all those who have helped directly or indirectly during the course of this work.

# TABLE OF CONTENTS

ABSTRACT .....	iii
ÖZ .....	iv
DEDICATION .....	vi
ACKNOWLEDGMENT .....	vii
LIST OF TABLES .....	xi
LIST OF FIGURES .....	xii
LIST OF SYMBOLS AND ABBREVIATIONS .....	xiv
1 INTRODUCTION .....	1
1.1 Introduction .....	1
1.2 Aim of this Thesis .....	3
1.3 Thesis Organization .....	3
2 FACE RECOGNITION .....	5
2.1 Introduction .....	5
3 DISCRETE WAVELET TRANSFORM .....	10
3.1 Overview .....	10
3.2 Necessity of Obtaining the Frequency Information .....	11
3.3 Multiresolution Analysis .....	13
3.4 Mathematical Backgrounds .....	13
3.4.1 Vector Space .....	13
3.4.2 Basis .....	13
3.4.3 Inner Product .....	17
3.4.4 Orthogonality and Orthonormality Property of Vectors .....	18
3.4.5 Direct Sum and Projection .....	19



3.5 Wavelet .....	21
3.5.1 Haar Wavelet .....	21
3.5.2 Theory of Wavelet .....	24
3.5.2.1 Continuous Wavelet Transform (CWT) .....	24
3.5.2.2 Discrete Wavelet Transform (DWT) .....	28
4 PROPOSED ILLUMINATION INVARIANT FACE RECOGNITION METHODS .....	35
4.1 Methodology .....	35
4.1.1 Benchmark Face Databases .....	35
4.1.2 Wavelet Transform .....	37
4.1.3 Z-score Normalization (ZN) .....	39
4.1.4 Luminance Quality (LQ) Metric.....	40
4.1.5 Nearest Neighbor (NN) Classifier .....	41
4.2 Proposed Methods .....	41
4.2.1 None Method .....	41
4.2.2 Histogram Equalization (HE) Method.....	42
4.2.3 Quality Base Histogram Equalization (QbHE) Method .....	42
4.2.4 Regional Histogram Equalization (RHE) Method.....	43
4.2.5 Regional Quality Base Histogram Equalization (RQbHE) Method .....	44
4.2.6 10 Fold Cross Validation Method.....	45
5 EXPERIMENTS AND RESULTS .....	48
5.1 LQ as an Appropriate Measure for Evaluating the Quality .....	48
5.2 Experiment and Discussion.....	49
6 CONCLUSION AND FUTURE WORK.....	59
6.1 Conclusions.....	59

6.2 Future Work .....	59
7 REFERENCES.....	61

## LIST OF TABLES

Table 2.1: Application of face recognition in some areas.[23] .....	9
Table 3.1: Table of wavelet sub-bands and corresponding applied filters.....	34
Table 4.1: Subsets and their LQ correspond to methods .....	47
Table 5.1: Identification accuracy rates base on different illumination normalization techniques for Extended Yale B data base for one level DWT.....	50
Table 5.2: Identification accuracy rates base on different illumination normalization techniques for Extended Yale B data base.....	52
Table 5.3: The accuracy of fix weighted of RHE method for LL and LH subbands.	53
Table 5.4: The accuracy of fix weighted of RHE method for LH2 and HL2 subbands in RQbHE method.....	54
Table 5.5: The accuracy of combination of LL2 and LH2 with fix weights for five different approaches. ....	55
Table 5.6: The accuracy of fix fusion method for LL and LH subbands in RQbHE approach. ....	55
Table 5.7: The accuracy of fix weighted for RQbHE method for LH and HL subbands.....	56
Table 5.8: Identification accuracy rates correspond to each approach achieved by 10 fold cross validation for Extended Yale B data base. ....	57
Table 5.9: GLQ and RLQ of each region for RHE and RQbHE methods when threshold sets to 0.88.....	58

## LIST OF FIGURES

Figure 2.1: Diagram of face detection and recognition in general.....	8
Figure 3.1: On the top a plot of the signal and on the bottom the frequency spectrum of an arbitrary Sinusoid function.[25].....	11
Figure 3.2: the interpretation of basis and duality in $\mathbb{R}^2$ (a, b) and $\mathbb{R}^3$ (c). ....	14
Figure 3.3: The geometric representation of orthogonal projection (a) and Oblique projection (b). $v = PV, Wy$ , $w = PW, Vy$ .....	20
Figure 3.4: The function $\varphi_0$ or scaling function (a) and the wavelet functions for $r = 0$ and $r = 1$ which named $\psi_1$ (b) and $\psi_2$ (c) respectively. ....	22
Figure 3.5: This plot is showing $\psi_{2,0x}$ and shifted versions of that geometrically in x-y plane.....	23
Figure 3.6: The relation between subspaces that are made by wavelet basis. ....	25
Figure 3.7: Geometrically representation of concept of direct sum.....	26
Figure 3.8: The relation between different wavelet functions .....	29
Figure 3.9: The FIR version of Haar scale function (a) and Haar wavelet function (b). .....	29
Figure 3.10: The wavelet analyze filter bank of a 1D signal in two levels.....	31
Figure 3.11: Different wavelet subbands in an image space.....	32
Figure 3.12: Block diagram of 2D wavelet transform in one level.....	33
Figure 3.13: Block diagram of 2D wavelet transform. ....	34
Figure 4.1: (a) Illumination subbands of Extended Yale B and (b) Example images of the ORL database. Some of Extended Yale B images used to calculate reference image (c) and reference face image (d) [1]......	36
Figure 4.2: Diagram of 2D wavelet hierarchical steps for $k=1$ .....	38

Figure 4.3: Decomposition of an image in one level (a) and two level (b).....	39
Figure 4.4: Block diagram of None method (right) and HE method (left) .....	42
Figure 4.5: Block diagram of QbHE method.....	43
Figure 4.6: Region segments that use in RHE and RQbHE techniques (a) and the number of pixels that included in each regions (b).....	44
Figure 4.7: Block diagram of RHE method .....	44
Figure 4.8: Block diagram of RQbHE method .....	45
Figure 4.9: selecting probe and train images in 10 fold cross validation method in 10 levels. ....	46
Figure 5.1: Combination of weighted subbands and save in a new vector. ....	51

## LIST OF SYMBOLS AND ABBREVIATIONS

DNA	Deoxyribonucleic acid
WT	Wavelet Transform
PCA	Principal Component Analysis
LDA	Linear Discriminant Analysis
IR	Infra-Red
DWT	Discrete Wavelet Transform
CCTV	Closed-circuit television
STFT	Short Time Fourier Transform
FT	Fourier Transform
FIR	Finite impulse response
WT	Wavelet Transform
DWT	Discrete Wavelet Transform
MRA	Multiresolution Analysis
2D	Two Dimension
LQ	Luminance Quality
ZN	Z-score Normalization
Std	Standard Deviation
HE	Histogram Equalization
NN	Nearest Neighbor classifier
K-NN	K-Nearest Neighbors
QbHE	Quality base Histogram Equalization
RHE	Regional Histogram Equalization
RQbHE	Regional Quality Base Histogram Equalization

PCA	Principal Component Analysis
$\mathbb{R}^n$	n-dimensional vector space over the field of real numbers
$\mathbb{C}^n$	n-dimensional vector space over the field of Complex numbers
$\mathbb{R}^2$	2-dimensional vector space over the field of real numbers
$\mathcal{L}_2(\mathcal{R})$	square integrable real space
$\varphi$	Basis
$\varphi_k$	k <sup>th</sup> basis component
$k, n, j, r$	An arbitrary integer number
$\mathbb{Z}$	Set of Integer numbers
$x$	An arbitrary Subspace
$\alpha_k$	Coefficient of kth basis
$\tilde{\varphi}$	Dual basis
$\tilde{\varphi}_k$	kth Dual basis
$e_k$	k <sup>th</sup> dimension axes
$\delta$	Direct delta function
N	The maximum number of linear basis of a space ( size of space)
$W \& V$	Subspace of vector space
$\mathbb{Z}$	Set of Integer numbers
$v, w \& y$	Vector
$V^\perp$	Orthogonal compliment of $V$
$W^\perp$	Orthogonal compliment of $W$
$\omega$	Angular frequency

$i$	Imaginary unit
$\varphi_{r,s}(x)$	Scaling function
$\psi_{r,s}(x)$	Wavelet function
$f(x)$	A function of a signal
$a_{r_0,s}$	Scaling coefficient
$b_{r,s}$	Wavelet coefficient
$h_\varphi(n)$	Haar scale function
$h_\psi(n)$	Haar wavelet function
$I(m, n)$	A digital image of size $m \times n$
$\varphi_{j_0,k}$	Kernel of transformation for scaling function
$\psi_{j_0,k}$	Kernel of transformation for wavelet function
$\varphi(m, n)$	LL subset of wavelet
$\psi^H(m, n)$	HL subset of wavelet
$\psi^V(m, n)$	LH subset of wavelet
$\psi^D(m, n)$	HH subset of wavelet
$W_\varphi$	Scale function
$W_\psi$	Wavelet function
H	Low-pass filter
G	High-pass filter



# Chapter 1

## INTRODUCTION

### 1.1 Introduction

Biometric recognition is referred to the biometric details of human body. These identification details of a person are based on his/her anatomical or behavioral characteristics features. There are many research activities that give an in-depth analysis in biometrics features like face, DNA, signature, fingerprint, handwriting geometry, voice print and eye verification. Among all these features, face recognition is one of the most common methods that used in the identification processes. Because of non-intrusive nature of image acquisition, face recognition is a very accurate identification and recognition technique. Developing face recognition system has received a considerable volume of consideration by the machine learning and computer vision researchers.

Generally, in facial recognition we must do several steps as follows. Firstly, we should do some pre-processing on images to make them ready for further phases. In this step we can resize the images, change the image's format, background matching and so on. Second phase is extracting the image's features and the final phase is recognition work.

Scientists attempted to design a system which is robust to against illumination variation, facial expressions, facial pose, and also noise. Continued investigation in

this area is derived by the demand for universal, efficient and trustworthy person identification methods in order to make the recognition process more convenient and reliable.

Changing in lighting condition has been one of the challenges in machine vision systems. Obtained face biometric samples that extracted from an image can lead to a reliable source of data if this information is robust against different lighting condition especially when the background is uncontrolled. Since vision cameras are commonly used in streets, airports, shops and many private and public places, the face recognition system plays an important role in centralized control rooms, security systems, crime and international terrorism.

Using information in images that were taken from a distance has a significant effect for person identification issues. Information in an image can be used as a tool for person identification by extracting essential and basic details. Wavelets transform (WT) as a multiresolution method is a great tool for excerpting and choosing these materials from selected images. In addition using wavelet has a great advantage in data size reduction. Data size reduction property of WT is very crucial for handheld and mobile devices which have limitation in data storage. Moreover, dealing with less data leads us to a faster process. After extracting facial details, data are normalized by using Z-score normalization before going to the next levels. For classification I go through two different and distinct methods. First, I followed the work that is done by Harin Sellahewa and Sabah A. Jassim in [1] and then, for extended experimental results, I used 10-fold cross validation to choose train and probe images of Extended Yale B database [2] in a different way for classifier. At the

end of the work I offered fusion method and compared methods together and discussed about the results.

## **1.2 Aim of this Thesis**

Besides the pose problem, illumination problem makes face recognition more complicated. Many researchers have been working in face recognition area to introduce an unaffected method in the presence of varying lighting condition in image processing and videos processing field. Meanwhile, the proposed method must be powerful and need low-cost computing system.

The work presented in this thesis targets discrete wavelet transform ability in feature extraction and usage of a quality based face recognition system to reach a better accuracy in face recognition.

The discrete wavelets transform and face recognition illumination problem is extensively studied and defined. Although many efforts have been spent on this problem, still it is not completely solved. I am interested in discrete wavelet transform applications and this face recognition problem because it is quite challenging to teach a machine to do efficient and reliable face recognition.

## **1.3 Thesis Organization**

This study is organized into six chapters; the first chapter is an introduction to the thesis and review. Chapter 2 gives a quick literature survey of face recognition that has been done in this field. Chapter 3 gives some preliminaries related to the wavelets and explain how features are extracted with WT and used for image processing.

Proposed approaches in face recognition are discussed in Chapter 4. Thesis followed with experimental results of the work that have done in Chapter 5. Finally, conclusions and further work ideas are drawn in Chapter 6.

## Chapter 2

# FACE RECOGNITION

### 2.1 Introduction

Facial recognition problem, as an important part of biometric recognition, has been studied extensively since 1960 and continued up to now by early works that done by Kelly [3], Bledsoe[4], and Kanade [5]. In the first stages, from 1960's till 1995, primary approaches for face recognition can be found in [6] and [7]. Generally, face recognition's researches, can be grouped into two distinct classes, first Feature-Based approaches and second Appearance-Based which is also known as holistic approach. Moreover, there is one more class which is achieved by combining other two methods together; hence named Hybrid method. Facial features like eyes, brows, mouth, nose, etc. and their geometric relationships are fundamental features in Feature-Based approaches. Edge detection and signature detections are some example of earlier works of biometric recognition which just involved basic methods of very simple image processing techniques. For instance, in feature base matching methods that proposed in [3], recognition is done by calculating the size (width) of the head and distance between face elements such as the distances between eyes and from eyes to the mouth. Different features investigated in [5] rely on distances and also angles between chin, nostrils, mouth extrema and eyes. References [3] and [5] rely on finding the exact locations of facial features in a face, unlike [8] which used Hidden Markov Model (HMM) methods. In this method, bands of pixels that cover the facial features (forehead, nose, mouth, etc.) are utilized. Studies are illustrating

that the feature based approaches are most robust against rotation, pose and scale and illumination variation. Since feature base methods use the facial features, they are extremely relying on the accuracy of facial feature selection procedure.

In Appearance-based approach, instead of analyzing face detail, just some features of the face are extracted as facial features. Two of the most commonly used appearance-based approaches - which are based on statistical methods - are Principal Component Analysis (PCA) [9-10] and Linear Discriminant Analysis (LDA). A valuable comparative analysis of PCA and LDA can be found in [11]. Computation of these statistical approaches depends on the dimension of the original data and the number of images that choose as train samples. Therefore, by growing the size of face database, a larger memory demand to handle the system data and also the process takes significantly longer time to be done for train. Disadvantages of appearance-based approaches are generally two main problems: firstly, the features can be extracted from the background of face. Secondly, the accuracy can be significantly affected by deviating from the average face of a gallery set because of lighting, orientation and scale [11].

To pass over the illumination variation problem, the effectiveness of visible and infra-red (IR) imagery for recognizing faces was compared in [12]. The performance for IR and visible imagery was found to be similar. Another approach based on Discrete Wavelet Transform (DWT) was successfully used in [13-15]. Many types of systems have been successfully applied to the task of face recognition, but they all have some advantages and disadvantages. Appropriate schemes should be chosen starting from the specific requirements of a given task.

Beside alterations in lighting situations, facial expressions, poor camera instrument quality and pose cause identification faults. There is a propensity to offer “standard reference” images with respect to these variations, increasing to extents of image feature [17-18]. Statistical techniques such as PCA need train phase. However, DWT as a multiresolution technique, used as a tool without train phase, to extract a multiresolution feature representation of a given face image [19-20]. Sellahewa and Jassim [21-22] have shown that the low-frequency approximation subbands of wavelet transform is an appropriate face descriptor for recognition when illumination is controlled; however it is greatly pretentious by varying illumination. Contrariwise, other subbands (which obtained by high-pass filters and represent horizontal features and/or vertical features) are robust in contradiction of varying illumination conditions. However, they are influenced by pose and facial expressions.

When we have a gallery or a data set of facial images of people and we want to recognize a given image as input image by using facial recognition algorithm, face recognition should be done. The recognition algorithm matches any of images in the input set to a person from the gallery. Face recognition also known as facial recognition. The most commonly use of facial recognition is in video surveillance to match the identity of people in surveillance footage to an existing database. (See Table 2.1)

Face recognition is a part of larger area called object recognition. As it discussed before, face recognition is programing a machine or a computer, based on composition mathematic relations that formed by a computer language in order to get the machine the ability of recognizing face of different people. In this issue, the

brightness level of the image, the direction of the face and generally, variation in the angle between camera, subject (person) and light source is one of the reasons of failure in face recognition systems. This problem is marked especially when the background is uncontrolled. Table 2.1 shows some of the applications of face recognition briefly.

In automated systems, identifying of a face in a photo or a film that is taken with a camera is a challenging problem. For humans, this task has been done pretty well without any affords but, for programing a machine, it is quite different. In machine vision at first, we must detect the location of the face in a photo then recognize the subject that detected. Face detection and face recognition is widely used in many fields from centralized control room to handhelds and mobile phones. Face recognition is highly affected by pose and illumination. Figure 2.1 shows some steps of face detection and face recognition tersely.

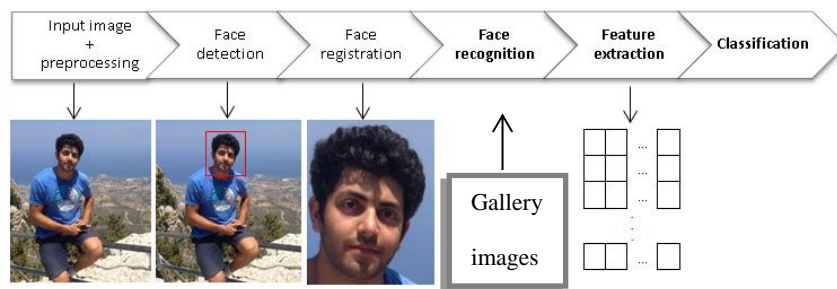


Figure 2.1: Diagram of face detection and recognition in general

In normal vision cameras (ordinary digital cameras) the quality of images are highly affected by using the techniques to solve the lighting condition problem. That's why Infrared (IR) option added to many cameras later and various techniques applied to eliminate this weakness. Regardless of pose, it is obviously that recognition of an image with an appropriate lighting quality has less error while face recognition of



dark images is pretty hard and challenging. Therefore, the accuracy of an automated face recognition system is depending on varying illumination in captured images. Solving this problem is the core of subject of various researches in recent studies. This thesis focuses to solve this obstacle as discussed before.

Many techniques are used to increase the accuracy of the computer vision's results by decreasing the negative affect of varying lighting condition between enrollment and test images during the recognition progress.

Table 2.1: Application of face recognition in some areas [23]

Area	Applications
Access Control	Facility Access, Vehicular Access
Biometrics	Drivers' Licenses, Entitled Programs, Immigration, National ID, Passports, Voter Registration
Information Security	Computer Logon Application Security, Database Security, File Encryption Intranet Security, Internet Access, Medical Records Secure Trading Terminals
Law Enforcement and Surveillance	Advanced Video Surveillance, CCTV Control

## Chapter 3

### DISCRETE WAVELET TRANSFORM

#### 3.1 Overview

In practice the majority of the signals are TIME-DOMAIN signals. It means that we can do the measuring as a function of time. By plotting time-domain signals, a time-amplitude representation of the signal is achieved. In other words, the plot of these signals has axes of time (independent variable) and a dependent variable that usually called amplitude axes. Related to the applications, this representation is not always suitable representation of the signal for most cases. Usually, majority of distinguished information is concealed in the frequency content of the signal. The information in the frequency spectrum of a signal tells what frequencies exist in our signal. [18]

If changes in a variable are fast, we can say that it is a high frequency variable, whereas if the changes do not occur rapidly, i.e., it changes smoothly, we call that low frequency variable. The FT of a signal in time domain gives us the frequency-amplitude representation of that signal. The plot of FT of a signal represent us how much of each frequency exists in our signal. In Figure 3.1 you can see a time-domain signal and its frequency spectrum.

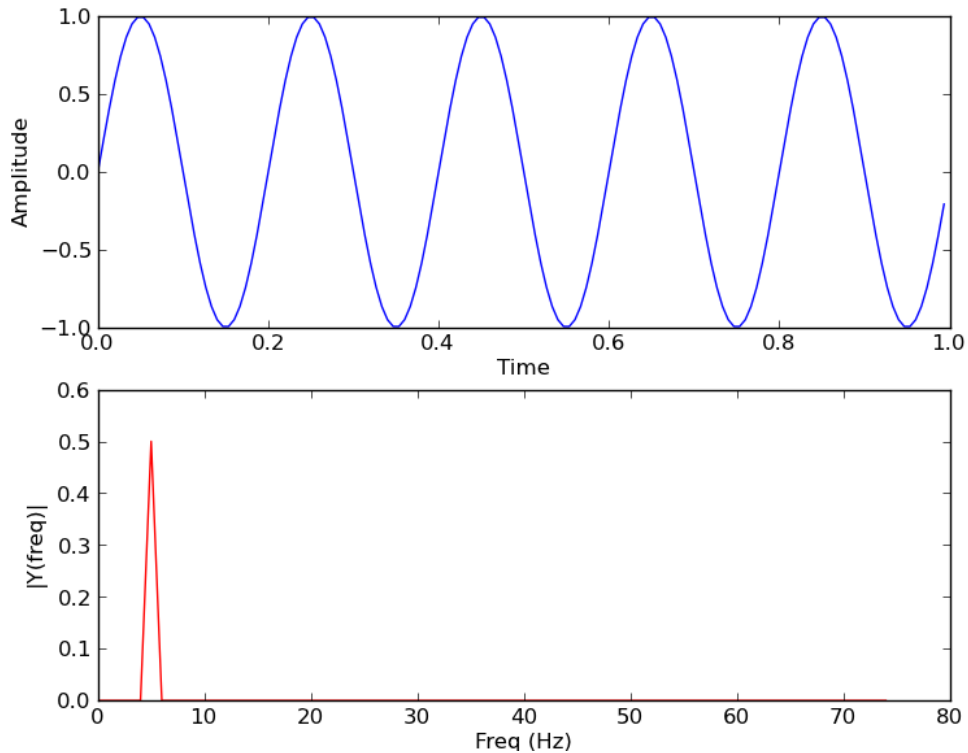


Figure 3.1: On the top a plot of the signal and on the bottom the frequency spectrum of an arbitrary Sinusoid function [25]

### 3.2 Necessity of Obtaining the Frequency Information

Most of the time information that obtained from the time-domain representation is not adequate to do further processing. Thus, by frequency domain representation we can see those hidden information.

Fourier transform just gives us frequency components of the signal, nothing more. It means that there is no access to the time information in the Fourier transformed signal. On the other hand, there is no frequency information obtainable in the time-domain representation of a signal. Frequency information is not required when the signal is so-called stationary.

If there are no changes in frequency content of a signal, this signal is called stationary. In other words, when we are working with stationary signals we do not see any frequency changes when the time changes. In this case, it is not necessary to have the detail of the frequency component's exact time.

The time localization of the spectral components of a signal is expressed by the time-frequency representation of the signal. To reach this aim we can use Short Time Fourier Transform (STFT) and the WT was developed as an alternative to the STFT.

Briefly, we pass the time-domain signal from individual distinct high pass and low pass filters. Which the output of filters, both high frequency and low frequency are fragments of the signal. This procedure is imitated, every time some portion of the signal corresponding to some frequencies being removed from the signal. In other word, it split the signal into two main parts, low frequency and high frequency. This operation is called decomposition.

Although with information given in FT, we are not able to find out what spectral component is exists at any exact time, we can investigate what spectral components exist at any given interval of time. It makes a new problem that called resolution. To solve difficulty, scientists have transferred to Wavelet transform (WT). STFT gives an unchanged resolution at all times, while WT offer a range of resolutions as follows: lower frequencies which are better determined in frequency, and higher frequencies that are well resolved in time. In the other word, a clear high frequency component can be located in time with less relative error in compare with a low frequency component. In contrast, low frequency components are able to locate

better in frequency in compare with a high frequency component. Wavelet Transform (WT) is appropriate to analyze non-stationary signals.

### 3.3 Multiresolution Analysis

The Wavelet transform of a signal is obtained by passing the time-domain signal through various high-passes and low-passes filters repeatedly. It means that we can analyze any signal at different frequency with different resolution. This representation of signal is called Multiresolution Analysis (MRA). Before going through wavelet transform details, we need to describe the main idea of wavelet analysis theory.

### 3.4 Mathematical Backgrounds

#### 3.4.1 Vector Space

A vector space is defined over a set. This set can be real or complex denoted by  $\mathbb{R}$  and  $\mathbb{C}$  respectively. By definition, any linear combination of elements in a vector space must be another element of it.

#### 3.4.2 Basis

Now consider linear expansions of signals (or functions). Let consider  $S$  be a space that is finite-dimension (for instance  $\mathbb{R}^n$  or  $\mathbb{C}^n$ ) or infinite-dimension (for instance  $\mathcal{L}_2(\mathcal{R})$ ) and  $x$  is a subspace of  $S$ . Based on linear theorem, we will be able to find a set like  $\{\varphi_k\}_{k \in \mathbb{Z}}$  to write  $x$  as a summation of linear combination. Since  $x \in S$ ,  $x$  can be expanded by equation:

$$x = \sum_k \alpha_k \varphi_k \quad (3.1)$$

where  $\{\varphi_k\}$  is spanning the complete space  $S$ . In signal processing topic,  $\{\tilde{\varphi}_k\}_{k \in \mathbb{Z}}$  namely dual basis is defined in order to compute expansion coefficients represent in equation (3.1):

$$\alpha_k = \sum_n \tilde{\varphi}_k[n] x[n] \quad (3.2)$$

There is some different type of dual basis defined such as orthogonal, biorthogonal and over complete (frame). In Figure 3.2 some possible sets of vectors for the expansion of the plan ( $\mathbb{R}^2$ ) is shown.

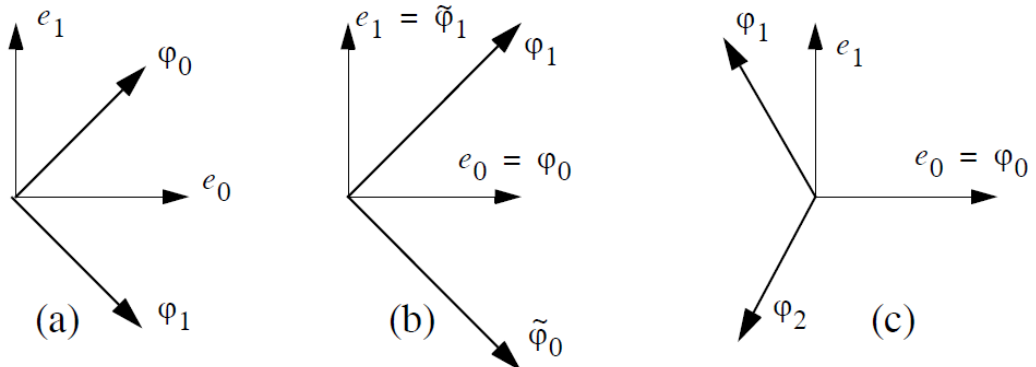


Figure 3.2: the interpretation of basis and duality in  $\mathbb{R}^2$  (a, b) and  $\mathbb{R}^3$ (c)

In Figure 3.2 (a)  $e_0$  and  $e_1$  are orthogonal to each other and  $\varphi_0$  is orthogonal to  $\varphi_1$ . Since both  $e_0, e_1$  and  $\varphi_0, \varphi_1$  can span  $\mathbb{R}^2$ , we call them orthogonal basis for  $\mathbb{R}^2$ . Moreover in Figure 3.2 (b)  $e_0$  and  $e_1$  are orthogonal but  $\varphi_0$  and  $\varphi_1$  are not orthogonal to each other thus, to compute expansion coefficients (3.2) we need to define  $\tilde{\varphi}_0$  as a dual for  $\varphi_1$ . Accordingly  $\tilde{\varphi}_1$  is a dual for  $\varphi_0$ . In part (c) of Figure 3.2, there are three orthogonal basis  $e_0, e_1$  and  $\varphi_2$ , which can span the  $\mathbb{R}^3$  and  $\varphi_1$  is a frame.

Since  $x$  and  $\tilde{\varphi}_k$  are discrete-time functions (sequence), the summation is appeared in definition of  $\alpha_k$ .  $\alpha_k$  will express by integral when  $x$  and  $\tilde{\varphi}_k$  both are continuous-time functions:

$$\alpha_k = \int \tilde{\varphi}_k(t) x(t) dt \quad (3.3)$$

The result of (3.1) and (3.2) can be written by expending inner product as follow  $\langle \tilde{\varphi}_k, x \rangle$ . For simplicity of calculation, we define an especial case which that the set of  $\{\varphi_k\}$  (that known as basis) is orthonormal and complete, since then its dual is the same, that is,  $\varphi_k = \tilde{\varphi}_k$ . Base on inner product and duality property then we have:

$$\langle \varphi_k, \varphi_j \rangle = \delta[k - j] = \delta_{kj} \quad (3.4)$$

In this equation  $\delta_{k,j}$  is used as direct delta function.

$$\delta_{k,j} = \begin{cases} 1 & \text{for } k = j \\ 0 & \text{otherwise} \end{cases} \quad (3.5)$$

For basis, biorthogonal case is happening when the set is complete and bases are linear independent but the orthonormality property does not take place. In this case the basis and its dual satisfy

$$\langle \varphi_k, \tilde{\varphi}_j \rangle = \delta_{kj} \quad (3.6)$$

In one exceptional type of basis, that called *frame*, the set is complete but because of redundancy the linear independency is no longer satisfy (so we do not have a basis) (Figure 3.2 (c)).

Vector space generated by span of basis vectors. Thus we must discuss about the concept of span. If  $S \subset E$ , the span of  $S$  is a subspace of  $E$  which contain the whole

possible linear combination of vectors in  $S$ . For finite-dimension the span of  $S$  is define as below:

$$span(S) = \left\{ \sum_{k=0}^{N-1} \alpha_k \varphi_k \mid \alpha_k \in \mathbb{R} \text{ or } \mathbb{C}, \varphi_k \in S \right\} \quad (3.7)$$

Generally it can be said that if  $S = \{\varphi_k\}_{k=0}^{N-1}$  is a linear independent set, any vector in  $s$  can be represented uniquely by linear combination of its basis. For instant assume  $S_1 = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$ ,  $S_2 = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$  and  $S_3 = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\}$ , We can conclude that  $S_1$  is not a basis for  $\mathbb{R}^2$  but  $span\{S_2\} = span\{S_1\} = \mathbb{R}^2$ , because the linear combination of vectors of  $S_1$  are gathered in  $span\{S_1\}$ . For  $S_1$  and  $S_2$  this relation is confirmed:

$$S_1 \subseteq S_2 \Leftrightarrow span\{S_1\} \subseteq span\{S_2\} \quad (3.8)$$

By adding a linear combination of basis to a set we are increasing the redundancy and we will not see any changes in span of that set. The linear independency is guaranty that the representation is unique but in presence of redundancy linear representation is not unique.

The linear independency for  $\{\phi_k\}_{k=0}^{N-1}$  is defined as

$$\sum_{k=0}^{N-1} \alpha_k \varphi_k = 0 \Leftrightarrow \alpha_k = 0 \text{ for all } k. \quad (3.9)$$

A subset  $S = \{\varphi_k\}_{k=0}^{N-1}$  that defined over a vector space like  $E$  is called a basis for  $E$  if and only if  $E = span(\varphi_0, \varphi_1, \dots, \varphi_{N-1})$  and  $\varphi_0, \varphi_1, \dots, \varphi_{N-1}$  are linear independent. In this case the vector space  $E$  known as  $N$ -dimension space. If  $E$  includes an



infinite linear independent set of vectors in its basis,  $E$  called infinite-dimensional space.

### 3.4.3 Inner Product

If  $V$  and  $W$  are subspace of vector space  $S$  then the inner product is a function that assign to each order pair of vectors (3.10). Using inner product gives us optionality to define various mathematical topics such as Norm. By means of inner product many important theorem is define e.g., Cauchy-Schwarz inequality, Triangle inequality, Parallelogram law.

*if  $V, W \subset S$  and  $v \in V$  and  $w \in W$*

$$\Rightarrow \begin{cases} \text{for continuous } (\mathbb{R}): \langle v, w \rangle = \int_t v^*(t) w(t) dt \\ \text{for discrete } (\mathbb{Z}): \langle v, w \rangle = \sum_{k=0}^{N-1} v^*[k]w[k] \end{cases} \quad (3.10)$$

By applying the inner product, the norm of a vector can define simply:

$$\|v\| = \langle v, v \rangle^{1/2} \quad (3.11)$$

And the distance between two vectors is simply defined by the norm of differences  $\|v - w\| = \langle v - w, v - w \rangle^{1/2}$ . This norm is known as Euclidean or square norm. There are other norms that defined for different customs. For example City Block norm and City Block distance are defined by:

$$\|v\| = \sum_{i=0}^{N-1} |v_i| \quad (3.12)$$

and

$$\|v - w\| = \sum_{i=0}^{N-1} |v_i - w_i| \quad (3.13)$$

respectively. In some texts, City Block distance is called Manhattan distance. Be informed that the energy of a signal is obtained by Euclidean norm.

$$\|v\|_2^2 = \sum \langle v, v \rangle^2 \quad (3.14)$$

### 3.4.4 Orthogonality and Orthonormality Property of Vectors

A vector  $v$  is said to be orthogonal to a set of vectors  $W = \{w_k\}$  if the inner product of that vector and any vector in  $W$  is equal to zero:

$$\langle v, w_k \rangle = 0 \Rightarrow v \perp W \quad (3.15)$$

As a rule, two subspace  $W_1$  and  $W_2$  are called orthogonal ( $W_1 \perp W_2$ ) if any vector in  $W_1$  is orthogonal to any vector in  $W_2$ . Moreover, orthogonality is defined for any arbitrary set in similar way. If  $V = \{v_k\}_{k=0}^{N-1}$ ,  $\langle v_j, v_k \rangle = 0$  ( $v_j \perp v_k$ ) when  $j \neq k$ ,  $V$  named an orthogonal set. For some reasons such as simplicity in mathematical calculations we attempt to normalize a vector to achieved unit norm. By doing normalization we obtain an orthonormal system which is satisfies

$$\langle v_k, v_j \rangle = \delta_{kj} \quad (3.16)$$

Because of convenience in calculating and also large number of mathematics options that available, we are trying to work with the basis that chosen orthonormal. One of the great advantages of orthonormal basis is shows below:

$$\begin{aligned} \text{if } x[k] = \sum_{k=0}^{N-1} \alpha(k) \varphi(k) &\xrightarrow{\text{if } \varphi \text{ is orthonormal basis}} a_k = \langle x(k), \varphi(k) \rangle \\ a_k &= \sum_{k=0}^{N-1} x^*(k) \varphi(k) \end{aligned} \quad (3.17)$$

As an implicitly assume, a Hilbert space is a space that contain a countable number of orthonormal basis. In other word, Hilbert space satisfies complete inner product specification. For continuous-time signals this property is define by substituting integral instead of summation. In dealing with continuous time signals since the equation (3.3) is satisfied (in presence of orthonormal basis) equation (3.17) can use to present the coefficients( $\alpha$ ).

$$\langle \alpha_k, \alpha_j \rangle = \delta_{kj} \Rightarrow x(t) = \int_t \alpha(t) \varphi(t) dt \Rightarrow \alpha(t) = \langle x(t), \varphi(t) \rangle, \quad (3.18)$$

### 3.4.5 Direct Sum and Projection

If  $S$  is a Hilbert space, both  $W$  and  $V$  are subspace of  $S$  ( $W, V \subset S$ ),  $S = W \cup V$  and  $W \cap V = \emptyset$  therefore, we can represent  $S$  by direct sum of two subspaces  $W$  and  $V$ :

$$S = W \oplus V \quad (3.19)$$

In this case decomposition of  $S$  is unique. This decomposition is known as Oblique case of Oblique projection. In special case,  $W$  and  $V$  are orthogonal so it can determined that  $W$  and  $V$  are orthogonal compliment:

$$if W \perp V \Rightarrow W = V^\perp (V = W^\perp) \Rightarrow S = V \oplus V^\perp$$

To illustrate these concepts, consider the following example. If  $y$  is a vector and  $V$  and  $W$  are linear vector spaces, it can be said that  $v$  is projection of  $y$  along  $W$  and similarly  $w$  is the projection of  $y$  along  $V$ . Refer to Figure 3.3 to see the concept of projection geometrically .It obvious that  $V$  and  $W$  which plotted in Figure 3.3 can construct  $\mathbb{R}^2$ . In the other word,  $W \oplus V = \mathbb{R}^2$ .

In discrete Fourier transforms

$$F(e^{i\omega}) = \sum_{k=-\infty}^{+\infty} f[k]e^{-i\omega k} \quad (3.20)$$

$\phi = \{e^{ik\omega}\}$  for all  $k \in \mathbb{Z}$  and  $s = \{\phi_k\}_{k=0}^{N-1}$  is basis. But this representation has its own weakness for instance  $v = \text{span}\{s\}$  is always a subspace since, the basis of  $v$ ,  $(\phi)$  cannot be zero in following statement.

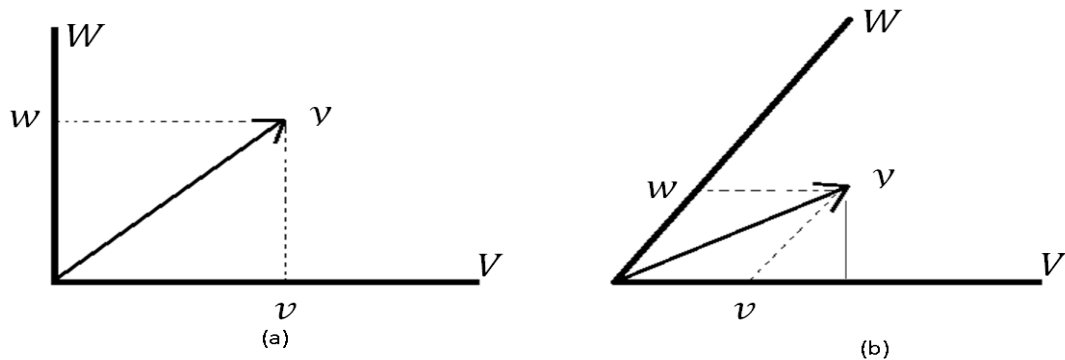


Figure 3.3: The geometric representation of orthogonal projection (a) and Oblique projection (b);  $v = P_{V,W}y$  ,  $w = P_{W,V}y$

Every vector in a vector space can be written as a linear combination of the basis vectors in that vector space.

In Fourier based methods we assumed that the signals are linear and stationary and this is one of the most significant motives to avoid using Fourier transform. The key difference between Fourier bases methods and Wavelets is the nature of basis functions that made the fundamental of both theorems. In Fourier, the basis is a function of sine and cosine. Meanwhile, in wavelet transform, basis can be chosen by mother wavelet function adequate to the problem.

However the main challenging issue on Wavelet is the proper selection of mother wavelet function and the accuracy of results is purely based on mother wavelet function that chosen.

### 3.5 Wavelet

#### 3.5.1 Haar Wavelet

Up to now we had pointed out some of crucial deficiencies and it had said that  $L^2(A)$  is a linear space of finite energy signals with duration in  $A$ . Fourier transform decompose our function into sine and cosine. In the other word, it could give use a basis in  $L^2([0,1])$  which consisting of sin waves. Alfred Haar in the year 1910 discovered different basis for a subspace of  $L^2([0,1])$ . After 100 years from the time that Alfred Haar discovery has borne, the signal processing become very much akin to Haar wavelet (which also called mother wavelet). Haar wavelet is famous because of its simplicity in calculation and also speed of computation. These two specifications make it suitable for a large area of application in digital signal processing. By using Haar wavelet, two disparate type of information (coefficients) is obtained. 1- Course approximation and 2- fine detail of function. One of the prominent properties of Haar wavelet function is reversibility. The forward transform of scaling function is obtained easily by add two adjacent samples value and divide by two. As well, the wavelet coefficient can obtain by subtracting two adjacent samples value and divide by number 2. The reverse transform can calculate by simple adding and subtracting.

Suppose that  $x$  is a continuous signal and  $\varphi_0$ ,  $\psi_0$  and  $\psi_1$  are defined as

$$\varphi_0(x) = \begin{cases} 1 & \text{for } 0 \leq x < 1 \\ 0 & \text{elsewhere} \end{cases} \quad (3.21)$$

$$\psi_0(x) = \begin{cases} 1 & \text{for } 0 \leq x < \frac{1}{2} \\ -1 & \text{for } \frac{1}{2} \leq x < 1 \\ 0 & \text{otherwise} \end{cases} \quad (3.22)$$

and

$$\psi_1(x) = \sqrt{2}\psi_0(2x) = \begin{cases} \sqrt{2} & \text{for } 0 \leq x < \frac{1}{4} \\ -\sqrt{2} & \text{for } \frac{1}{4} \leq x < \frac{1}{2} \\ 0 & \text{otherwise} \end{cases} \quad (3.23)$$

respectively.

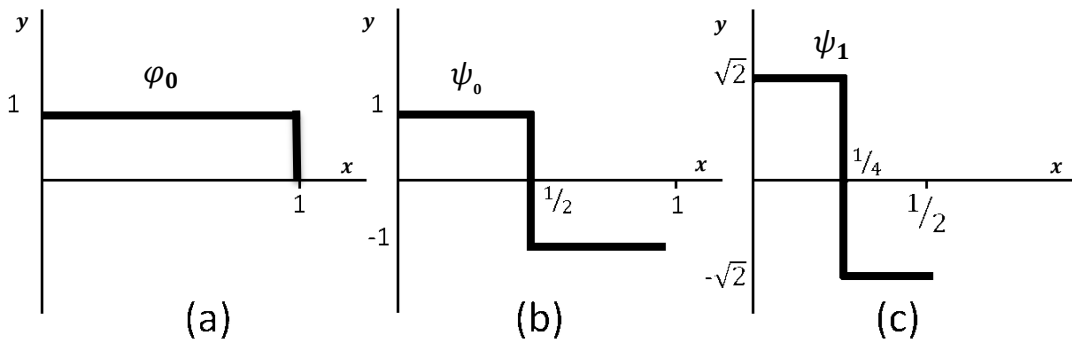


Figure 3.4: The function  $\varphi_0$  or scaling function (a) and the wavelet functions for  $r = 0$  and  $r = 1$  which named  $\psi_1$ (b) and  $\psi_2$  (c) respectively

As equation (3.23) illustrated,  $\psi_s(x)$  when  $s = 1$ ,  $\psi_1(x)$  is obtained by squeezing  $\psi_0(x)$  along  $x$  axis and stretch it along  $y$  axis by the value of  $\psi_0(x)$  coefficient and ratio of  $x$ . In Figure 3.4 the plot of  $\varphi_0(x)$ ,  $\psi_0(x)$  and  $\psi_1(x)$  are shown. Follow this process; function  $\psi_{r,s}(x)$  for  $r = 0$  and for  $s = 0, 1, 2$  and 3 are defined as below: (We assumed  $\psi_{0,0}(x) = \psi_0(x)$ )

$$\psi_{2,0}(x) = 2\psi_0(4x) = \begin{cases} 2 & \text{for } 0 \leq x < \frac{1}{8} \\ -2 & \text{for } \frac{1}{8} \leq x < \frac{1}{4} \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad \psi_{2,1}(x) = 2\psi_0(4x - 1),$$

$\psi_{2,2}(x) = 2\psi_0(4x - 2)$ ,  $\psi_{2,3}(x) = 2\psi_0(4x - 3)$  which are shifted version of  $\psi_2(x)$  that shown in Figure 3.5. We can also define shifted version of  $\psi_{2,0}(x)$ .

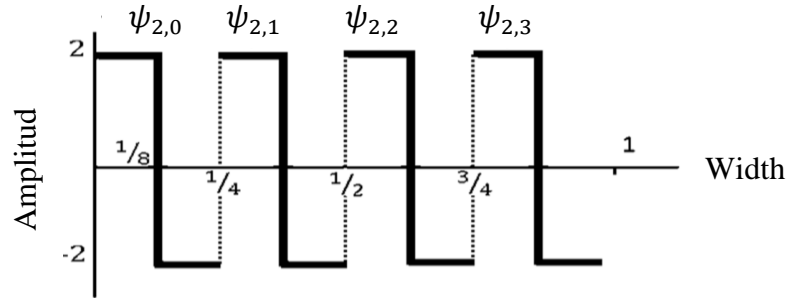


Figure 3.5: This plot is showing  $\psi_{2,0}(x)$  and shifted versions of that geometrically in x-y plane

Moreover, Figure 3.5 is showing that what is happening in an interval and also it's certain scale. This structure can continued to build more functions such as  $\psi_{14,s}(x)$ ,  $\psi_{30,s}(x)$  and  $\psi_{62,s}(x)$  which have smaller length scale and eventually we get functions that are too small so we can neglect them because they have poor resolution.

$$\psi_{14,s}(x) = 2\sqrt{2} \psi_0(8x - s) \quad |s = 0, 1, \dots, 7$$

$$\psi_{30,s}(x) = 4\psi_0(16x - s) \quad |s = 0, 1, \dots, 15$$

$$\psi_{62,s}(x) = 4\sqrt{2}\psi_0(32x - s) \quad |s = 0, 1, 2, \dots, 31$$

When we have a continuous signal like  $f(x)$  and want to know the coefficients respect to  $\psi_n(x)$ , it just needs to do  $V_n = \int_0^1 \psi_n(x)f(x)dx$ . If we do so,  $V_0, V_1, V_2, \dots$  will obtain which shows the resolution of that signal [16]. Finally, since the value of resolution and shift parameter is set to zero in  $\psi_{0,0}(x)$ , let assume  $\psi_{0,0}(x) = \psi(x)$ .

## 3.5.2 Theory of Wavelet

### 3.5.2.1 Continuous Wavelet Transform (CWT)

Let define a new class of functions  $\varphi_{r,s}(x)$  as below:

$$\varphi_{r,s}(x) = a^{r/2}\varphi(a^r x - s) \quad (3.24)$$

Which  $r$  and  $s$  are integers,  $a$  is a positive value greater than 1 and  $x$  is a variable in continuous space. With  $\varphi_{r,s}(x)$  and for any possible value of  $r$  and  $s$  we are able to produce the entire square integrable real space ( $span(\{\varphi_{r,s}(x)\}) = L^2(\mathbb{R})$ ).

Let assume  $a = 2$ , by substituting this assumption into equation (3.24) we will have the first idea of Haar wavelet:

$$\varphi_{r,s}(x) = 2^{r/2}\varphi(2^r x - s) \quad (3.25)$$

This set of functions is called scaling function. Now let choose  $r = r_0$ . Which  $r_0$  is a specific value of  $r$ . In this case we can say that  $\{\varphi_{r_0,s}(x)\}$  is just depends on changing the value of  $s$  since we assume a constant quantity for  $r$ . Moreover, since there is no shifting and resolution parameter in  $\varphi_{0,0}(x)$ , we usually assume  $\varphi_{0,0}(x) = \varphi(x)$ .

Now let analyze this problem geometrically. As we discussed,  $r$  is an integer value and  $r_0$  and  $r_1$  have this relation  $r_1 = r_0 + 1$  and  $r_2 = r_1 + 1 = r_0 + 2$ . Due to equation (3.24), by substituting  $r_0$  with  $r_1 = r_0 + 1$  we will obtain  $V_0 = span(\{\varphi_{r_0,s}(x)\})$  and  $V_1 = span(\{\varphi_{r_1,s}(x)\})$ . The comparison between  $V_0$  and  $V_1$  tells us the amplitude is increase by a factor of  $\sqrt{2}$  meanwhile the width is reduced by a factor of 2. Based on the spaces that subspaces  $V_0$  and  $V_1$  are contained, we can conclude that  $V_0 \subset V_1$ , because the resolution of  $V_1$  is higher than  $V_0$ .



The subsets  $V_2, V_3, \dots$  can obtain by the same way and since the  $r$  can takes any integer value we will have:

$$V_{-\infty} \subset \dots \subset V_{-1} \subset V_0 \subset V_1 \subset V_2 \subset \dots \subset V_{\infty} \quad (3.26)$$

These relations in (3.26) are well illustrates in Figure 3.6.

Based on linearity, any arbitrary function that lying within  $V_1$  can approximate by linear combination of  $V_1$  basis ( $\{\varphi_{r,s}(x)\}$ ). Also by Figure 3.6 it is obvious that any function that lying within  $V_0$  can represent by  $V_1$  basis (since  $V_0 \subset V_1$ ). Now let assume that

$$V_1 = V_0 \oplus W_0 \quad (3.27)$$

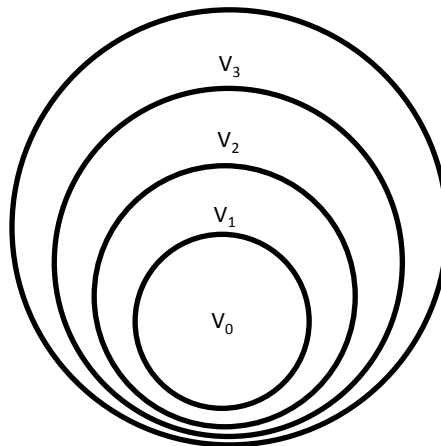


Figure 3.6: The relation between subspaces that are made by wavelet basis

In the other word, subspace  $W_0$  contain the difference between  $V_0$  and  $V_1$  by applying this change of presenting  $V_2$  we will have

$$V_2 = V_1 \oplus W_1 = V_0 \oplus W_0 \oplus W_1 \quad (3.28)$$

Assume the set of basis that presented in equation (3.25) and let  $r = 1$ . Generally, we can compose any function that spanned by set of  $\{\varphi_{1,s}(x)\}$  from summation of different shifted versions of next higher space functions that have specific weights.

$$\varphi(x) = \sum_n h(n)\sqrt{2}\varphi(2x - n)$$

In this expression the shifted basis is  $\varphi(2x - n)$  which  $n$  is shifting parameter and  $h(n)$  are the coefficients with respect to each basis.

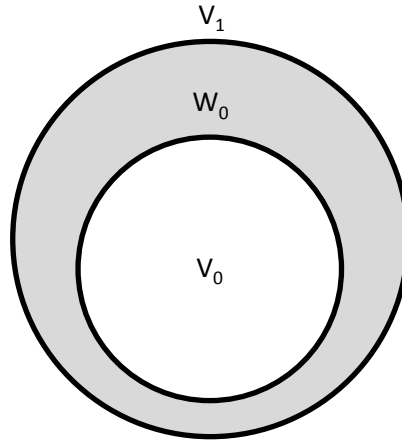


Figure 3.7: Geometrically representation of concept of direct sum

By developing this algorithm up to  $V_n | n \geq 0$ ,  $V_n$  can represent by direct sum of  $V_0$  and all  $W_k | 0 \leq k < n - 1$

$$V_n = V_0 \oplus W_0 \oplus W_1 \oplus \dots \oplus W_{n-1} \quad (3.29)$$

By applying  $\varphi_{0,0}(x)$  or (3.21) that had been shown in Figure 3.4 (a) over a signal, technically do the low-pass filtering. But whenever we consider a function that cover the differences in the subspace which has covered by the two low-pass filters we must use high-pass filters. Thus, the type of filters that can span the difference spaces

that covered by two low filter is high-pass filter. The original forms of this class of functions which is called *Wavelet function* are introduced as bellow:

$$\psi_{r,s}(x) = 2^{r/2}\psi(2^r x - s) \quad (3.30)$$

This presentation is very similar to (3.25) however they are completely distinct and span different spaces. When the shifted versions of (3.25) and (3.30) is considered, it is imperative for them to satisfy orthogonality with respect to each other. So, due to (3.28), to represent  $V_2$  we need one scale function  $V_0$  and two wavelet functions  $W_0$  and  $W_1$ . As well, to find  $W_2$ , we need  $V_2$  and  $V1$  cause  $W_2$  is the difference between them. The relation between wavelet function and scaling function is  $\psi(x) = \sum_n g(n)\sqrt{2}\varphi(2x - n)$  and it tells us the wavelet function can construct by using a series summation of shifted versions of scaling function of the next higher subspace. For example, reassume equation (3.21) as scale function (known as Haar scale function) and in equation (3.30) let  $s = 0$  and calculate the wavelet function for each values of  $r = 0, r = 1$  and  $r = 2$ . The results are shown in Figure 3.8 (a), (b) and (c) respectively. By adding up the scaling function subspace and wavelet function subspace with higher resolution order that had been offered in this example, in fact, we are able to analyze any function in  $L^2(\mathbb{R})$ . This concept is well defined by *wavelet series*:

$$f(x) = \sum_s a_{r_0,s}\varphi_{r_0,s}(x) + \sum_{r=r_0}^{\infty} \sum_s b_{r,s}\psi_{r,s}(x) , r \geq r_0 \quad (3.31)$$

That  $a_{r_0,s}$  is corresponding coefficients which associated with  $\varphi_{r_0,s}$ . Also, this  $\varphi_{r_0,s}$  is using in set of  $\psi(x)$  as well. Moreover,  $\sum_s a_{r_0,s}\varphi_{r_0,s}(x)$  can cover  $V_{r_0}$  subspace. In the wavelet series,  $b_{r,s}$  are coefficients associated with  $\{\psi(x)\}$ .

The basic objective reason to applying the wavelet transform to an image is achieving space frequency localization with the image. In the other word, wavelet transform tells us at what position, what frequency component exists. If we fix  $r$  to a constant value, the scaling and wavelet functions become orthogonal since the scale parameter  $s$  is change by integer numbers. As the orthogonality property is satisfying, we can obtain coefficients in (3.31) as below:

$$a_{r_0,s} = \int f(x)\varphi_{r_0,s}(x) dx \quad (3.32)$$

$$b_{r,s} = \int f(x)\psi_{r,s}(x) dx \quad (3.33)$$

For discrete Haar wavelet the filters that used are as follow.

$$h_\varphi(n) = \left\{ \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right\} \quad (3.34)$$

$$g_\psi(n) = \left\{ \frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}} \right\} \quad (3.35)$$

Which  $h_\varphi(n)$  is known as Haar scale function and  $h_\psi(n)$  is known as Haar wavelet function (see Figure 3.9)

### 3.5.2.2 Discrete Wavelet Transform (DWT)

However we defined  $x$  as a continuous signal, in point of fact, when we work with computers we deal with digital signals which obtained from continuous sequences by applying kind of sampling. For instant, any image can be sampled as the form of  $I(m,n)$  which  $m = 1,2, \dots, M$ ,  $n = 1,2, \dots, N$  and  $M \times N$  is the size of image.

Consequently, the DWT for a one dimension signal like  $I(n), n = 1,2, \dots, N$  can obtain by replacing integral with summation

$$W_\varphi(j_0, k) = \frac{1}{\sqrt{N}} \sum_n I(n) \varphi_{j_0, k}(n) \quad (3.36)$$

$$W_\psi(j, k) = \frac{1}{\sqrt{N}} \sum_n I(n) \psi_{j, k}(n) \quad (3.37)$$

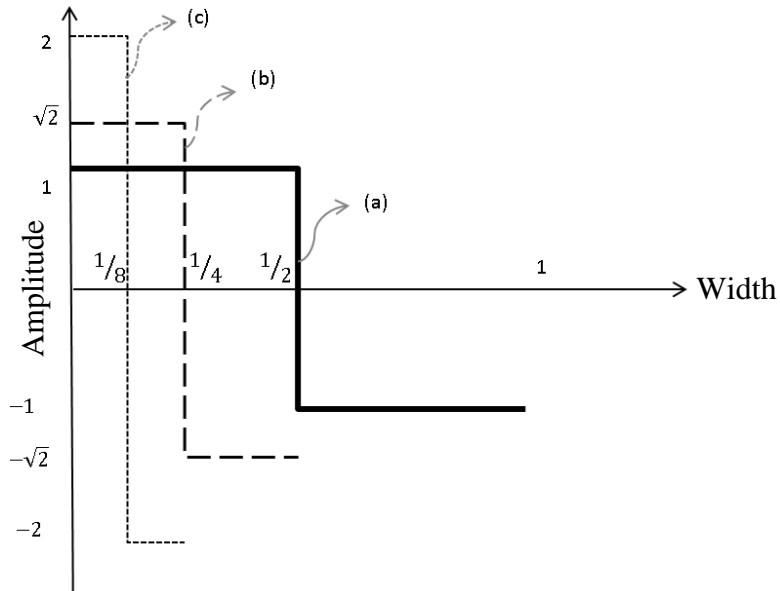


Figure 3.8: The relation between different wavelet functions

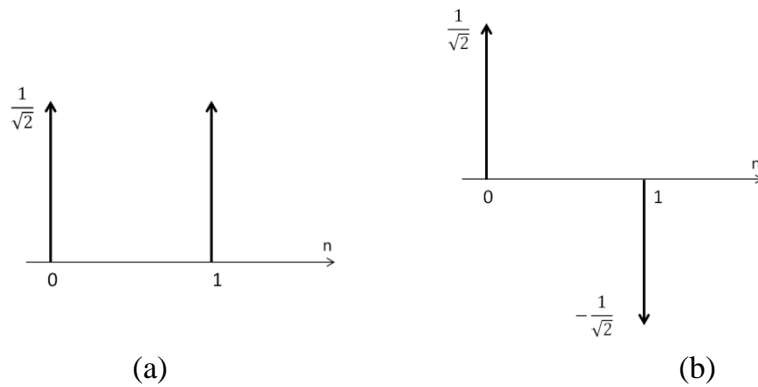


Figure 3.9: The FIR version of Haar scale function (a) and Haar wavelet function (b).

Which  $W_\varphi$  and  $W_\psi$  are the scaling function and the wavelet function quantity respectively that are similar to  $a_{r_0, s}$  and  $b_{r, s}$ ,  $k$  is an integer number and  $j \geq j_0$ . In compare with (3.32) and (3.33), here we use  $j_0$  instead of  $r_0$  and also use  $k$  as scaling

parameter. Since we convert the signal  $S(n)$  to a new domain ( $W_\psi(j_0, k)$  and/or  $W_\varphi(j_0, k)$ ), we need to normalize the phrases by  $\frac{1}{\sqrt{N}}$  to be sure that the energy of signal remains unchanged after shifting the domain. The wavelet series for discrete signals makes us able to produce the original signal from scaling function and wavelet function as follow:

$$f(x) = \frac{1}{\sqrt{N}} \sum_k W_\varphi(j_0, k) \varphi_{j_0, k}(x) + \sum_{j=j_0}^{\infty} \sum_k W_\psi(j_0, k) \psi_{j_0, k}(x) , r \geq r_0 \quad (3.38)$$

Furthermore,  $\varphi_{j_0, k}$  and  $\psi_{j_0, k}$  are the kernels of this transformation. For convenience in calculation, normally we set  $j_0 = 0$  and  $N = 2^j$ , ( $j = 0, 1, \dots, N - 1$ ). Due to equation (3.30), we will have:

$$\psi_{j, k}(n) = 2^{j/2} \psi(2^j n - k) \quad (3.39)$$

By substituting this equation into equation (3.37) we will have  $W_\psi(j_0, k) = \frac{1}{\sqrt{N}} \sum_n I(n) 2^{j/2} \psi(2^j n - k)$ . This equation leads us to two new useful relations:

$$W_\psi(j, k) = \sum_n g(n - 2k) W_\varphi(j + 1, n) \quad (3.40)$$

$$W_\varphi(j, k) = \sum_n h(n - 2k) W_\varphi(j + 1, n) \quad (3.41)$$

If  $W_\varphi(j + 1, n)$  is available, we just need to convolve it with  $h$  and  $g$  to obtain the scale function and wavelet function of lower resolution respectively. Usually,  $h$  is use to express low-pass filter and  $g$  shows high-pass filter. Wavelet essentially permits us to realize a bank of analysis filters which means we can obtain the original signal back if we apply a corresponding bank of synthesis filters. Figure 3.10(a) shows analyze of a one dimension (1D) signal from  $S(n)$  to three bands and its frequency response. As it illustrates in Figure 3.10 (b), filter banks do not have the same bandwidth. In this figure,  $S_1(n)$  which is the ultimate low-pass filter output has

the one quarter,  $S_2(n)$  has one fourth and  $S_3(n)$  as the result of passing the original signal through the high-pass filter, has half of the original bandwidth. Also, the frequency response of  $S_1(n)$ ,  $S_2(n)$  and  $S_3(n)$  are shown correspondingly in (1), (2) and (3) fragments of Figure 3.10 (b).

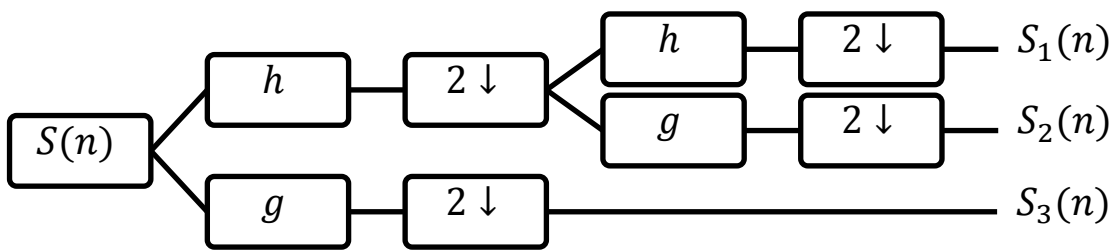
The four different possible functions are:

$$\varphi(m, n) = \varphi(m)\varphi(n) \quad (3.42)$$

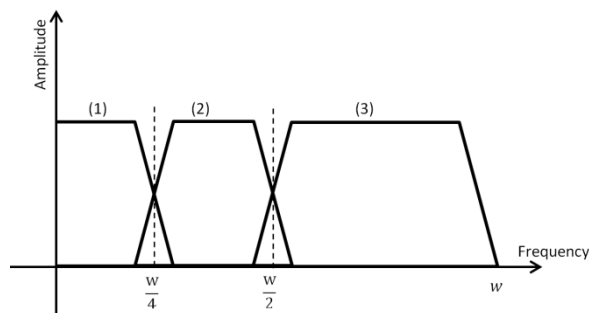
$$\psi^H(m, n) = \psi(m)\varphi(n) \quad (3.43)$$

$$\psi^V(m, n) = \varphi(m)\psi(n) \quad (3.44)$$

$$\psi^D(m, n) = \psi(m)\psi(n) \quad (3.45)$$



(a)



(b)

Figure 3.10: The wavelet analyze filter bank of a 1D signal in two levels.

In (3.42),  $\varphi(m, n)$  represents the approximation of image which is obtained from passing both rows and columns of image through the low-pass filter. In (3.43) the

supper script  $H$  shows that rows are high-pass filtered (wavelet function) but columns are low-pass filtered (scale function). Equation (3.44) represents  $\psi^V(m, n)$  which is calculated by taking low-pass filter from rows and high-pass filter form columns. Finally the last subband is diagonal subband  $\psi^D(m, n)$ . Here supper script  $D$  is referring to diagonal.

Whenever we applying 2D DWT to an image, since any level is made by the product of two filters, we have to do double decimation by the factor of two. Sharply, decimation by factor of two in horizontal direction and decimation by another factor of two in vertical direction and in overall, we do the decimation by factor of four. This is the reason that each subbands contained one quarter of image space as it shows clearly in Figure 3.11. in this figure Lena benchmark image (Figure 3.11 (a)), discrete wavelet subbands of that in one level (Figure 3.11 (b)) and the responses of different wavelet subbands in the image space (Figure 3.11 (c)) are shown.

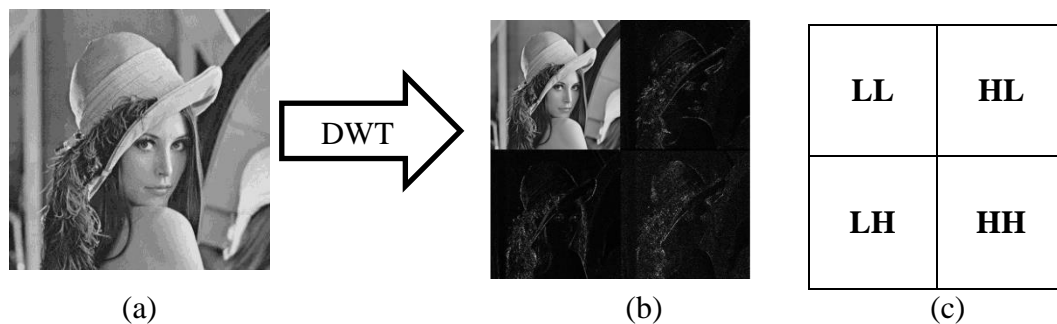


Figure 3.11: Different wavelet subbands in an image space

We discussed about how wavelet transform is applicable on the images. To understand it better assume that  $W_\phi(j + 1, m, n)$  is an image with size  $m \times n$  in the first scale  $j + 1$ . By applying 2D DWT to the image, at the first stage we must pass the image through the high-pass and low-pass filter distinctly, which indicate as



$h_\varphi(-n)$  and  $h_\psi(-n)$  correspondingly. Firstly step, the scale function and wavelet function must apply to columns; this is the reason of using  $(-n)$ . Likewise, rows will analyze at the second stage. The negative sign is for shifted part of  $h_\psi$  and  $h_\varphi$  which is introduced in (3.40) and (3.41). Decimation by the factor of two means selecting the alternate samples. In the other word, we remove redundant samples that will not carry any information. The bandwidth of signal essentially gets half in each scale and wavelet subbands. Figure 3.12 shows the diagram of filters and decimation blocks.

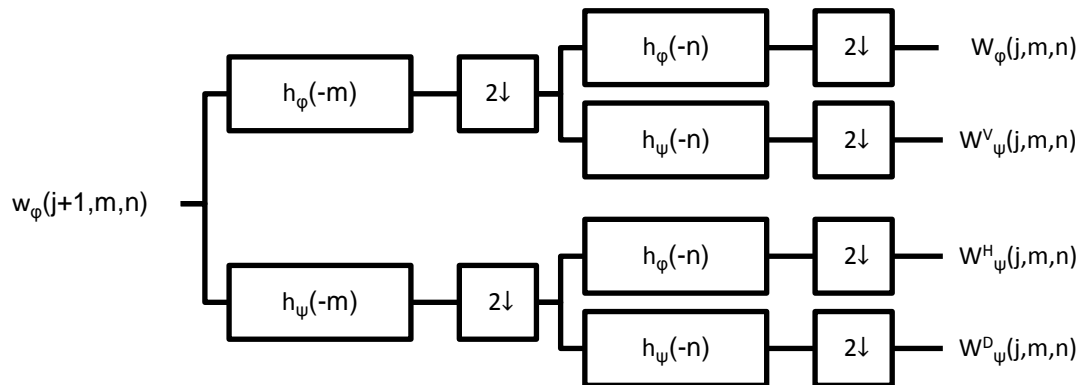


Figure 3.12: Block diagram of 2D wavelet transform in one level

Phrase  $W_\psi^D(j, m, n)$  that shows the diagonal edges, is the result of extracting the high-pass features along rows and columns so, it is corresponding to HH segment that showed in Figure 3.11.  $W_\varphi(j, m, n)$ ,  $W_\psi^H(j, m, n)$  and  $W_\psi^V(j, m, n)$  in this block diagram are referred to LL, HL and LH respectively. HL and LH are show the extract of horizontal and vertical edges individually (see Table 3.1). Interestingly, from Figure 3.12 it can conclude that by giving an image with a scale of  $j + 1$  to the wavelet transform, the output of transform will be at scale of  $j$ . Decrease the scale by factor of one means effectively lose the resolution by a factor of 2 for both rows and columns. Moreover, each of these four sub bands can be analyzed further but, since

images are very rich in low frequency content, we mostly do the further analysis on LL sub-bands.

Accordingly, the squeeze version of original image is offered by LL sub-band. Highest scale is referring to original image which has the maximum resolution as well. By going further in decomposition, the coarseness of sub-bands will decrease. In the other word, by partitioning we move from finer domain to coarser domain of analysis. This is exactly the objective of doing wavelet which is frequency localization.

Table 3.1: Table of wavelet sub-bands and corresponding applied filters

Direction	Horizontal	Vertical	Subbands
Filter type	(Row)	(Column)	representatives
Low	X	X	LL
High			$W_\varphi(j, m, n)$
Low		X	HL
High	X		$W_\psi^H(j, m, n)$
Low	X		LH
High		X	$W_\psi^V(j, m, n)$
Low			HH
High	X	X	$W_\psi^D(j, m, n)$

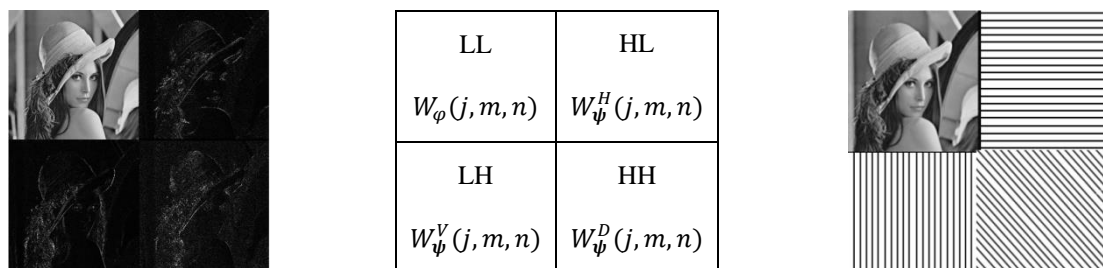


Figure 3.13: Block diagram of 2D wavelet transform.

## Chapter 4

# PROPOSED ILLUMINATION INVARIANT FACE RECOGNITION METHODS

### 4.1 Methodology

#### 4.1.1 Benchmark Face Databases

In order to evaluate the proposed face recognition system, our experiments are performed on following benchmark face databases.

**Extended Yale Database B (Extended Yale B)** [23]: Extended Yale B is containing 2414 images of 38 different subjects. For each subject there are five different illumination subsets which were divided base on the angle between light source and optical axis of the camera. Images in subset 1 are captured in good lighting condition. The majority of images in subset 2 have similar illumination condition in compare with set 1 and the light source and subject just have 20 to 25 degree difference. In Subset 3 because of larger angle between light source and subject, some images have shadow and it cause some dark area over faces. Because of these dark shadows, the lighting condition in set 3 is not as good as set 1 and 2. In subset 4 and subset 5, the angle between light source and subject increase up to 105 degree. The worst lightening condition is referred to subset 5 (Figure 4.1 (a)). All of the images in this database are captured in frontal pose and have a same size of  $168 \times 192$ . The number of images for each subject is varying between 59 and 64 images. For this experiment, I resize the images to size of  $128 \times 128$ . In each subsets there is one image which captured in frontal pose and direct illumination which named by

P00A+000E+00. These 38 images are used to build the gallery images which are used in classifier. Each of gallery images set in one row matrix with a size of [1, 16384] and then put them together. For example gallery group is a matrix with 38 rows and 16384 columns.

**AT&T (ORL) face database [26]:** The ORL database is used to find the luminance quality and threshold. In whole, this database has 400 images which contains of 40 different subjects. Each of 40 subjects, has a collection of 10 images were captured at different pose, facial expression and time (Figure 4.1 (b)).



Figure 4.1: (a) Illumination subbands of Extended Yale B and (b) Example images of the ORL database. Some of Extended Yale B images used to calculate reference image (c) and reference face image (d) [1]

Also, images in this data base captured against a dark homogeneous background and cropped in the size of  $92 \times 112$ . I resampled these images to a fix size of  $128 \times 128$  before using them. Since the images in this data base are captured in good lighting condition and there is no illumination variation, this database is an appropriate choice for introducing an average of face images this average image is called reference image and I use it to calculate Luminance Quality (LQ) (see Figure 4.1(c) and (d)).

#### **4.1.2 Wavelet Transform**

The objective of using WT in this work is achieving space frequency localization in images and to know at what position, what frequency component exists. By knowing this information we are able to select the most appropriate subbands of wavelet for face recognition. Since the images that used in this work have different lighting condition, in theorem,  $LL_k$  subband of wavelet transform cannot gives us the best face features in compare with  $LH_k$ ,  $HL_k$  and  $HH_k$ . Although by using some low-pass and high-pass filters we can reach us to same point but using wavelet has it own advantages. Using wavelet can compress the information then we can represent images with less data and it lead us to faster analyses. Moreover, once DWT applied to an image, the information of image in four different subbands can achieved.

Haar Wavelet, that known as simplest and the first wavelet is applied to test and train images in order to extract features that needed over recognition work. The Haar wavelet gives us an image into four different segments. These subbands represent different features base on utilized filters.

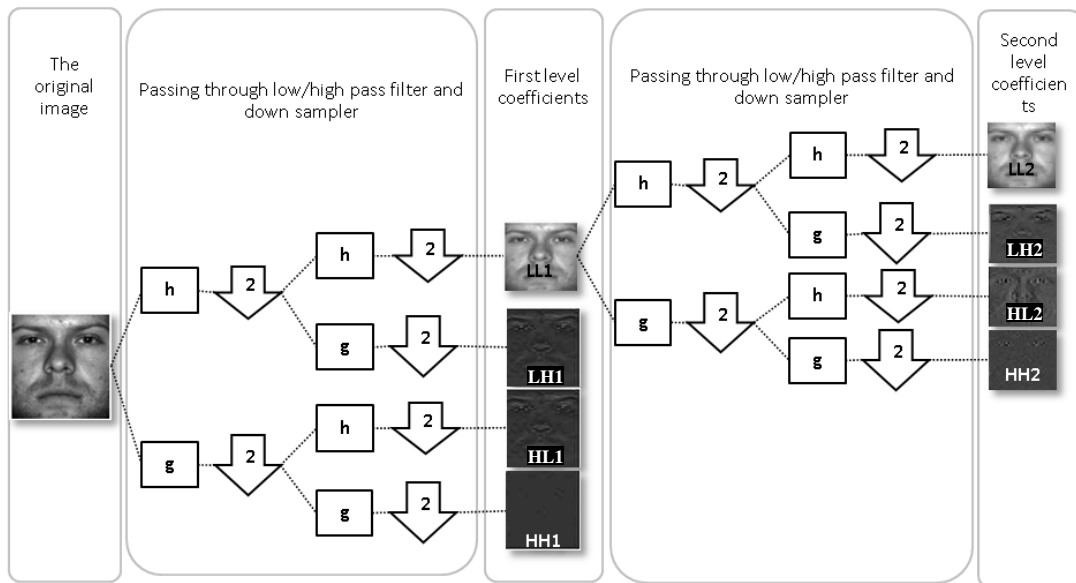


Figure 4.2: Diagram of 2D wavelet hierarchical steps for  $k=1$

I benefit the multiresolution property of WT to decompose images into low and high frequencies. These hierarchically decomposition of images, at the resolution of  $k$ , gives us  $3k + 1$  subbands. In Figure 4.2 the result of applying wavelet on an image is showed. In this figure,  $h$  and  $g$  are represent low pass and high pass filter respectively.

These sub-bands are known as  $LL_k, LH_k, HL_k, HH_k, \dots, LL_1, LH_1, HL_1, HH_1$ . In this arrange of sub-bands,  $LL_k$  is achieved by passing the signal (image) through a low-pass filter. Since the luminance is a trait of DC component of an image and the  $LL_k$  subband considered as a  $k_{th}$  level approximation of image, it is mostly affected to illumination variation in compare with other subbands.

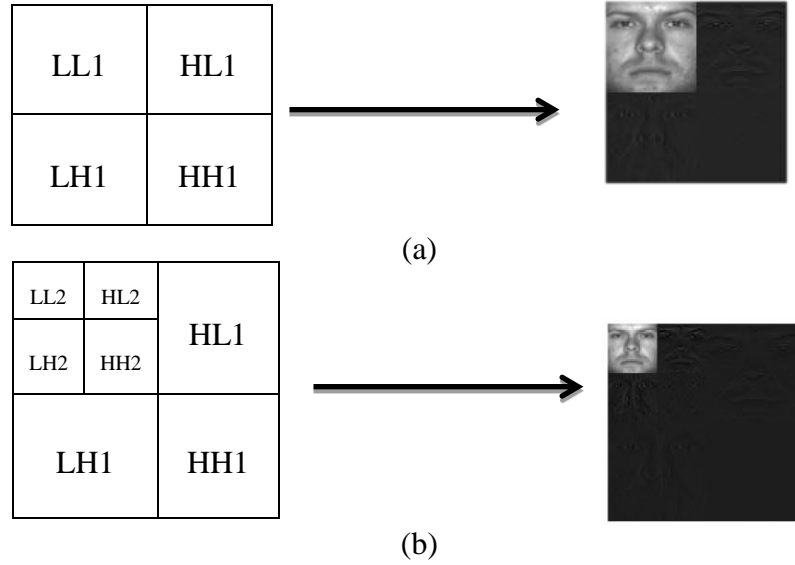


Figure 4.3: Decomposition of an image in one level (a) and two level (b)

#### 4.1.3 Z-score Normalization (ZN)

After extracting the facial feature, in order to improve recognition accuracy, there is need to apply a kind of preprocessing on data before using them in classifier. Typically, sub-bands coefficients are normalized. Here, face features normalized by Z-score normalization method. Assume that the wavelet coefficient that I want to normalize is  $x = \{x_i | i = 1, 2, \dots, N\}$  then

$$ZN = \frac{x - \bar{x}}{std(x)} \quad (4.1)$$

where  $\bar{x} = \frac{1}{N} \sum_{i=1}^N (x_i)$  is the average value of feature and  $std(x) = \frac{norm(x-\bar{x})}{\sqrt{N-1}}$  is the standard deviation.

ZN is based on calculating mean and standard deviation. By normalizing the wavelet coefficients with ZN, the recognition algorithm becomes robust against illumination in compare with lack of ZN. Therefore, it leads us better accuracy.

#### 4.1.4 Luminance Quality (LQ) Metric

As discussed before, the aim of this thesis is remove the harmful effect of different lighting condition in captured images when we want to do face recognition. This issue is done by applying a method which must be generalizable to vast area of application in facial recognition. The first idea against illumination variation problem is handle a normalization method to normalize illumination. This normalization must apply to images before extracting the facial features. In this work I utilized histogram equalization (HE) to normalize the illumination in preprocessing stage. Although using HE is a common use method to improve face recognition accuracy, this improvement is depends on the level of illumination discrepancy between test and trained images. I offered three distinct techniques to exert HE in preprocessing section.

In some techniques of this work, I did the HE if the luminance quality (LQ) of the image was less than a predefined measure. In these methods, before feature extraction I achieved the luminance quality. Applying equation (4.2) by assuming train images and test images are  $x = \{x_i | i = 1, 2, \dots, N\}$  and  $y = \{y_i | i = 1, 2, \dots, N\}$  respectively we have:

$$LQ = \frac{2\bar{x}\bar{y}}{(\bar{x})^2 + (\bar{y})^2} \quad (4.2)$$

Where

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i \quad (4.3)$$

and



$$\bar{y} = \frac{1}{N} \sum_{i=1}^N y_i. \quad (4.4)$$

The value range of LQ is [0,1] and it shows the distance of illumination between  $x$  and  $y$ . In (4.2) the LQ has the maximum value (LQ=1) if and only if  $\bar{x} = \bar{y}$ .

#### 4.1.5 Nearest Neighbor (NN) Classifier

After normalizing features, I compare test (probe) and train images by using the nearest neighbor as the standard determination technique. The NN classifier is a special case of K-Nearest Neighbors (K-NN) Classifier. In the other word, the K Nearest Neighborhood classifier when K=1 gives us nearest neighbor. The nearest neighbor classifier applied to take a decision between probe (test) and train groups. Train group contains the P00A+000E+00 image from each 38 subjects and remain images which are 2376 items in total, are used as test to calculate the efficiency of identification system. The CityBlock (Manhattan) distance calculated the distance score between train images and probe images [1].

## 4.2 Proposed Methods

With reference to [1] five new techniques were investigated and compared with other previews techniques in face recognition in presence of varying illumination. As an additional job, I changed the classification inputs (probe and train groups) by using 10 fold cross validation method. In [1] train images selected from well-lit images but by applying 10 fold cross validation not only the number of train images are increased, but also train groups include images of all five subsets.

### 4.2.1 None Method

Firstly wavelet transform applied to images without any luminance normalization technique. This approach is maned None. (Figure 4.4 (right)). As it is shown in Figure 4.4 (right), we took DWT of probe and train images and after that the wavelet

coefficients normalized with Z-score normalization (ZN) algorithm. The LL, HL and LH subbands of probe and train images gave to the classifier.

#### 4.2.2 Histogram Equalization (HE) Method

In this method all images normalized before extracting features without any exception. This approach was called HE. (Figure 4.4 (left)). After HE, features (wavelet subbands) of each image achieved by DWT and then normalized by ZN. These processes applied for both probe and train images in a same way. At last, data was given to NN classifier to complete the recognition progression.

#### 4.2.3 Quality Base Histogram Equalization (QbHE) Method

Thirdly, Quality Base Histogram Equalization (QbHE) approach was exerted (Figure 4.5). In this approach the LQ calculated for each image and compared with a predefine threshold [1]. If the image's LQ was less than the threshold, the image normalized with HE method before feature extraction. Otherwise the original image was given to feature extracting stage.

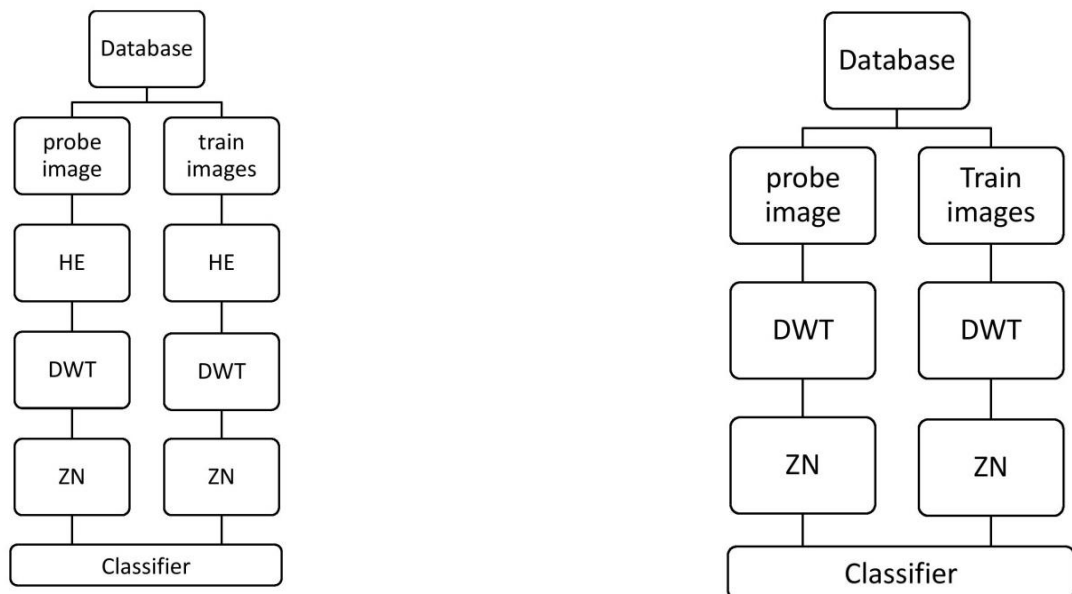


Figure 4.4: Block diagram of None method (right) and HE method (left)

After this level, features extracted by DWT and then ZN applied to normalize the features. For the last part, NN classifier applied to decide about the accuracy of method.

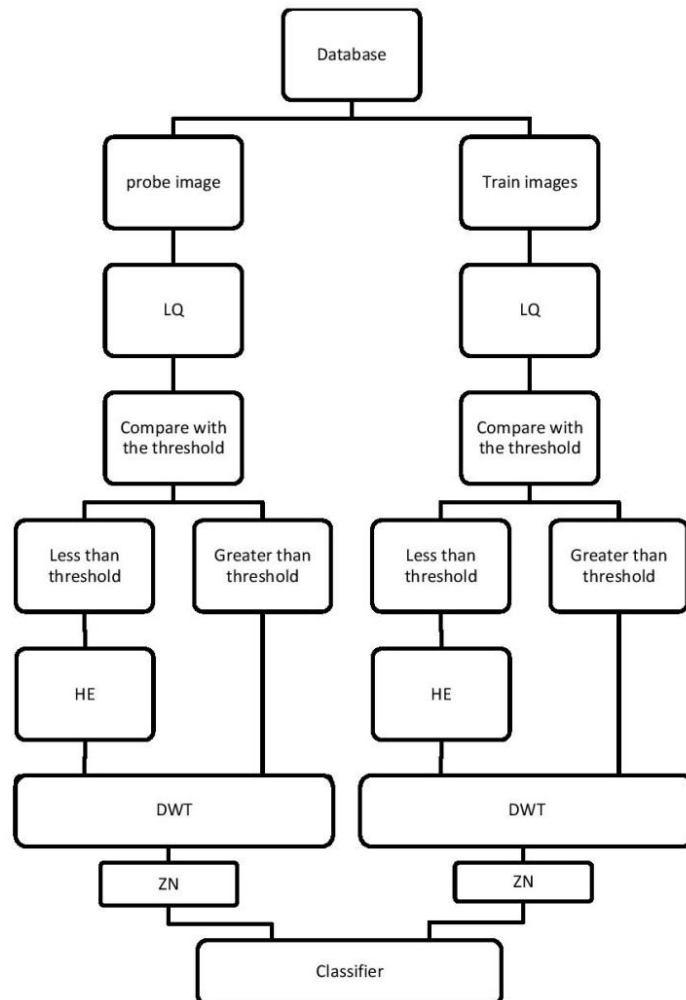


Figure 4.5: Block diagram of QbHE method

#### 4.2.4 Regional Histogram Equalization (RHE) Method

In this approach which named Regional Histogram Equalization (RHE), any images divided into four equal fragments by applying a 2 by 2 mask (Figure 4.6). For each region which had a size of  $64 \times 64$  pixels, HE applied separately. After normalizing each region of images distinctly, WT applied on any single images to extract features and NN classifier used to investigate the efficiency of method (Figure 4.7).

#### 4.2.5 Regional Quality Base Histogram Equalization (RQbHE) Method

At last, I tested the Regional Quality Base Histogram Equalization approach (RQbHE). This method had done by dividing images into four different regions similar to RHE technique and then compared the LQ of each single region with the threshold value that defined before in QbHE approach. HE only used for normalizing regions that have a LQ ratio less than the predefined threshold.

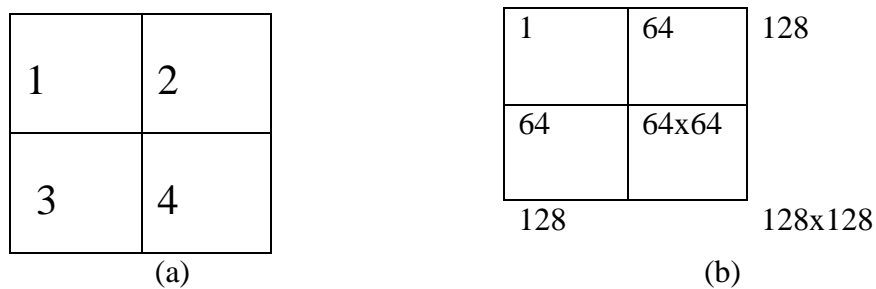


Figure 4.6: Region segments that use in RHE and RQbHE techniques (a) and the number of pixels that included in each regions (b)

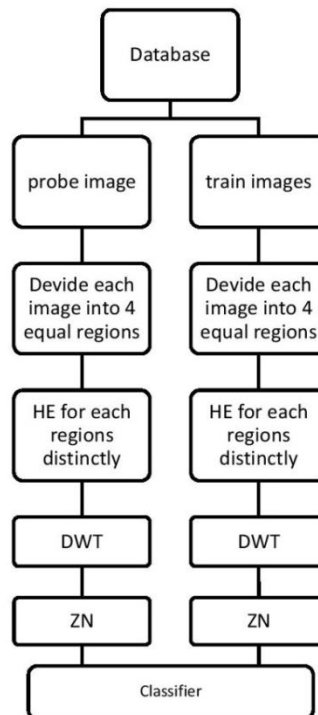


Figure 4.7: Block diagram of RHE method

A sample of each subset and the LQ measure of each image is shown in Table 4.1. Moreover, Table 4.1 shows an example for each method. As it discussed before, subset 1 contains images with a good illumination condition. In the other hand subset 5 images mostly captured under insufficient light. (Figure 4.8)

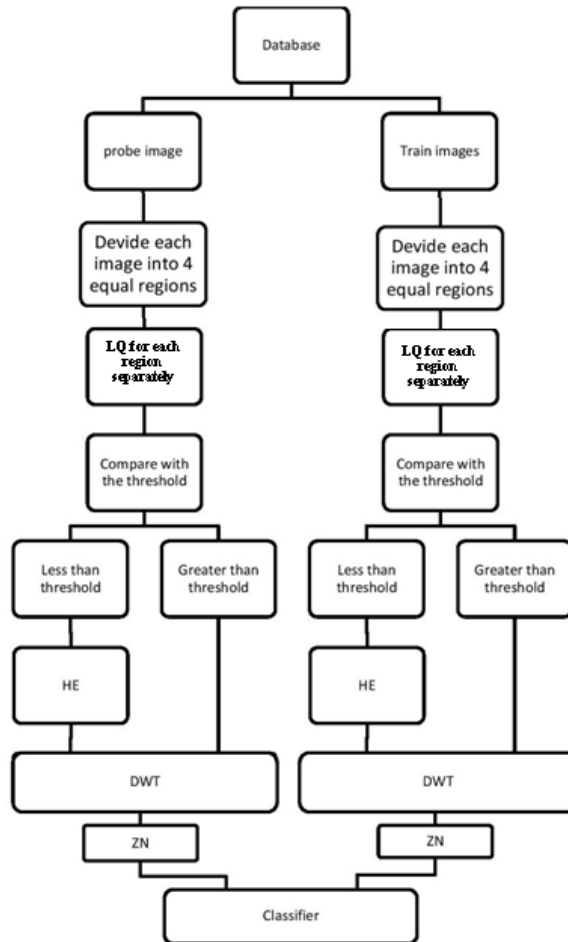


Figure 4.8: Block diagram of RQbHE method

#### 4.2.6 10 Fold Cross Validation Method

In this method, 10-fold cross validation technique applied to choose train and test images for classifier. In the previous techniques, the number of probe images was 2376 and train images were 38. But in new technique I have 10% of 2414 images as

test (probe) and the 90% rest assumed as train images and each time the accuracy calculated (Figure 4.9).

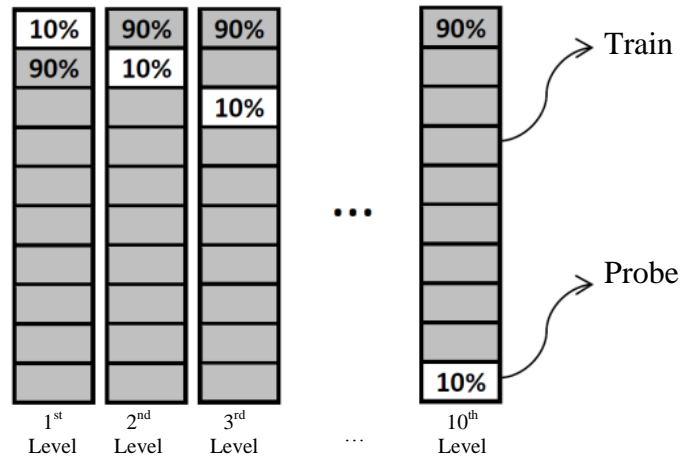







Figure 4.9: selecting probe and train images in 10 fold cross validation method in 10 levels.

At the end, the average of accuracies assume as final accuracy of approach. Each time, after calculating accuracy, I choose another 10 percent. By respect to this manner 10 distinct groups of probe and train images obtained. In this method, since 2414 subjects are sorted form subject 1 to subject 38, before 10 fold cross validation I mixed the place of rows to make sure about fairness in classification. The total number of image is 2414 and I select 241 images for first nine 10 percent and 244 for last 10 percent. Obviously, the number of train images is 2173 for the first nine 90% and for the last 90% is 2170.

Table 4.1: Subsets and their LQ correspond to methods

LQ	Method	Subsets					
		1	2	3	4	5	
	None						
		0.9987	0.8880	0.5647	0.4197	0.1481	
	HE						
		0.9865	0.9866	0.9863	0.9861	0.9850	
	QbHE						
		0.9987	0.9687	0.9863	0.9861	0.9850	
	RHE						
		0.9867	0.9865	0.9866	0.9860	0.9845	
	RQbHE						
		0.9987	0.9687	0.9866	0.9860	0.9845	

## Chapter 5

### EXPERIMENTS AND RESULTS

#### 5.1 LQ as an Appropriate Measure for Evaluating the Quality

Sabah A.Jassim and Harin Sellahewa in [1] by defining Global Luminance Quality (GLQ) showed that 60% of images in subset 1 (60 percent out of 225 images) have an equal or a higher GLQ score of 0.95. This rate is only 19% in the images of second subset. Additionally, 91% of subsets 1 images have a luminance quality mark of 0.90 or higher. This condition is same for 73% of the images of subsets 2. This measure proves that the GLQ quantity properly express the illumination excellence of images in first two groups. Since GLQ in subset one and two is very close to the predefined reference image, distinguishing that first subset images are closer to the predefine reference image in compare with images from subset 2. In addition in [1] is discussed that 96% of associates of subset 3 have a GLQ mark of less than 0.90, although, 33% of face images which made subset 3 have a global luminance quality value from 0.80 to 0.90. This specifies an obvious difference in lighting of images in subset 3. Approximately, only the LQ rate of 18% of whole 456 images in subset 4 is 60 or higher. While the maximum value of image's GLQ rate of this subset is merely 0.76. Almost 56% of the images in fifth subband recorded GLQ grade equal or greater than 0.70, whereas it's maximum GLQ score is individual 0.59. These numbers indicate that the images which grouped in subset 4 and 5 are poor is illumination quality aspect as a result of the great changes in illumination along horizontal and/or vertical directions (see Figure 4.1). By analyzing these data it



demonstrates that, as an illumination quality, the luminance quality (LQ) index nicely estimates the illumination quality for face image. [1]

## **5.2 Experiment and Discussion**

Recently, [22] shows us that by normalizing so-called well-lit images we cannot reach to better identification accuracy. However, normalizing images with bad lighting condition could lead to increase LQ factor markedly. These conclusions can be easily understood by Table 4.1. For more information Table 5.9 is given for threshold 0.8. In this table, a sample image refers to subset 1 remained unchanged during RQbHE. Meanwhile, the illumination of second subset image is normalized for regions 2 and 4 which had luminance quality equal 0.7773 and 0.77858 respectively. Additionally, to understand better the difference between adaptive technique like RQbHE and non-adaptive techniques, Table 5.9 gives the accuracy of different illumination normalization techniques respect to each subsets. The accuracy of each approach for first level of DWT is given by Table 5.1. Table 5.2 examined the accuracy of LL1 obtained from approximation subband of first level of 2D WT besides LL2, HL2, HH2 and LH2 which are features obtained by second level of 2D WT. Since the images are rich in low frequencies, LL1 and LL2 subbands are highly affected by bad condition of lighting. Unlikely, HL2 and LH2 are more robust against illumination variation. Then HH2 cannot assume as an appropriate descriptor of face base on Table 5.1 and Table 5.2. By comparing accuracy rates given in Table 5.1 and Table 5.2 generally it can conclude that LL1 given better accuracy in compare with LL2 nevertheless LH2 and HL2 has more accurate results in recognition approaches in compare with LH1 and HL1. In overall, HH2 has a higher accuracy in compare with HH1 but its result is not comparable with LL, HL and LH subbands. LL1 subband contains LL2, LH2, HL2 and HH2 subbands and also HH,

HL and LH are more robust against illumination variation. Accordingly, I chose LL2, LH2 and HL2 subbands as selected features in this work. Table 5.2 clearly implies these concepts. Furthermore, it shows that LH2 subband of RQbHE technique has the best accuracy in recognition in presence of luminance variation.

Table 5.1: Identification accuracy rates base on different illumination normalization techniques for Extended Yale B data base for one level DWT

Wavelet subbands	Method	Accuracy (%)					
		Set1	Set2	Set3	Set4	Set5	All sets
LL1	None	98.67	84.65	34.86	6.58	4.2	<b>35.82</b>
	HE	98.22	82.68	32	10.31	16.95	<b>39.31</b>
	QbHE	98.67	84.65	35.05	8.33	18.63	<b>40.53</b>
	RHE	100	100	72.38	25	20.45	<b>55.6</b>
	RQbHE	98.22	86.84	56.95	30.7	27.59	<b>52.74</b>
LH1	None	84.89	100.00	83.05	65.79	24.51	<b>65.57</b>
	HE	84.00	100.00	80.95	81.58	79.55	<b>84.60</b>
	QbHE	84.89	100.00	83.43	75.88	73.39	<b>82.28</b>
	RHE	83.11	100.00	81.33	81.14	76.05	<b>83.46</b>
	RQbHE	84.89	100.00	82.48	78.51	69.05	<b>81.27</b>
HL1	None	81.78	98.03	73.90	31.14	4.62	<b>50.25</b>
	HE	79.11	97.81	70.48	46.27	28.43	<b>59.26</b>
	QbHE	81.78	98.03	74.10	38.82	25.35	<b>58.00</b>
	RHE	77.33	97.59	70.67	49.34	36.41	<b>62.08</b>
	RQbHE	81.78	97.81	71.43	39.25	22.83	<b>56.69</b>
HH1	None	55/11	87/72	44/57	31/58	8/82	<b>40/61</b>
	HE	56/89	86/18	36/95	17/32	10/92	<b>36/70</b>
	QbHE	55/11	87/72	43/62	28/73	11/90	<b>40/78</b>
	RHE	58/67	87/50	37/71	19/08	11/62	<b>37/84</b>
	RQbHE	55/11	87/72	35/81	12/06	9/52	<b>35/14</b>

In the next stage, to achieve better results and base on the highest accuracy rates in Table 5.1 and Table 5.2; two wavelet coefficients which achieved by RQbHE and RHE approaches combined together distinctly with considering certain weights. This

work has been done by combining subbands with a constant factor. Assume the combination of two subbands saved in  $f_{Subband}$ :

$$f_{Subband} = \alpha S1 + \beta S2 \quad (5.1)$$

Which  $\alpha$  and  $\beta$  are real positive coefficients and  $\alpha + \beta = 1$ .  $S1$  and  $S2$  are wavelet subbands which can be LL, LH or HL. Assume that  $S1$  and  $S2$  have  $n$  features therefore they are a row matrix of size  $1 \times n$ . In order to combine subbands, at first the correspond coefficient multiplied to them, and then put weighted coefficients together in a new matrix like  $f_{Subband}$ . This new matrix has one row and  $2n$  columns (Figure 5.1).

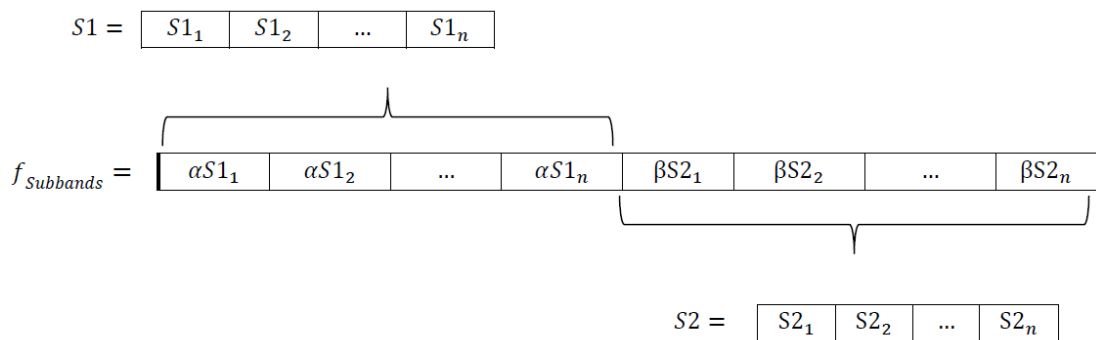


Figure 5.1: Combination of weighted subbands and save in a new vector

The accuracy based on different fix weighs of LL2 with LH2 subbands and LH2 with HL2 subbands are given in Table 5.6 and Table 5.7 respectively.

In Table 5.6 and Table 5.7, the identification accuracy is shown by corresponding subbands fusions. Followed by accuracies for RQbHE method it can observe that, fusion quality based histogram equalization method is much accurate than before. The higher accuracy is referring to fusion method with LH2 and HL2 subbands since these two subbands are robust against illumination variation.

Table 5.2: Identification accuracy rates base on different illumination normalization techniques for Extended Yale B data base

Wavelet subbands	method	Accuracy (%)					
		Set1	Set2	Set3	Set4	Set5	All sets
LL1	None	98.67	84.65	34.86	6.58	4.2	<b>35.82</b>
	HE	98.22	82.68	32	10.31	16.95	<b>39.31</b>
	QbHE	98.67	84.65	35.05	8.33	18.63	<b>40.53</b>
	RHE	100	100	72.38	25	20.45	<b>55.6</b>
	RQbHE	98.22	86.84	56.95	30.7	27.59	<b>52.74</b>
LL2	None	98.67	79.82	32.76	6.36	3.78	<b>34.26</b>
	HE	97.33	77.41	30.29	9.43	15.27	<b>37.16</b>
	QbHE	98.67	79.82	32.95	7.68	16.11	<b>38.26</b>
	RHE	100	100	68.19	21.93	19.33	<b>53.75</b>
	RQbHE	97.78	82.02	53.52	27.41	25.63	<b>49.79</b>
LH2	None	89.33	100	85.14	66.89	31.79	<b>68.86</b>
	HE	88.44	100	81.9	82.24	88.52	<b>88.05</b>
	QbHE	89.33	100	84.76	76.1	83.47	<b>86.07</b>
	RHE	87.56	100	83.43	84.43	89.64	<b>89.06</b>
	RQbHE	89.33	100	85.14	84.21	89.08	<b>89.39</b>
HL2	None	90.67	99.12	83.24	32.68	6.3	<b>54.17</b>
	HE	88	99.12	83.43	63.6	45.66	<b>71.72</b>
	QbHE	90.67	99.12	84	45.39	40.76	<b>67.13</b>
	RHE	88	99.12	84.19	62.94	43	<b>70.96</b>
	RQbHE	90.67	99.12	83.81	56.58	37.96	<b>68.39</b>
HH2	None	67.56	95.18	57.14	28.29	5.88	<b>44.49</b>
	HE	67.56	94.74	65.14	51.97	44.12	<b>62.21</b>
	QbHE	67.56	95.18	58.29	33.11	27.87	<b>52.27</b>
	RHE	69.78	94.96	62.10	46.05	38.94	<b>59.09</b>
	RQbHE	67.56	95.18	56.38	37.72	29.13	<b>53.11</b>

The accuracy rate in table 5.6 and Table 5.7 overall, illustrate that the multi-stream approach is more precise in identification in compare with individual sub-band representation. However, the results highly depend on selecting the efficient weights as well as choosing subbands. At first glance, by neglecting or giving a small weight

to LL2 subband the identification accuracy improve. However, LL2 subband contribution in fusion weighted method is more effective for images in subset 1.

Moreover, to find out the best results and choose the most accurate approach in this work the same process for unchanged fusion manner, applied to RHE method as well. Since RHE and RQbHE methods have the largest accuracy rate between other methods, comparing the results of these two approaches can lead us to better conclusion.

Table 5.3: The accuracy of fix weighted of RHE method for LL and LH subbands

Wavelet Subbands		Accuracy (%)					
LL+LH		Set 1	Set 2	Set 3	Set 4	Set 5	All sets
Factors							
1	0	100.00	100.00	68.19	21.93	19.33	<b>53.75</b>
0.9	0.1	100.00	100.00	74.10	28.73	23.11	<b>57.49</b>
0.8	0.2	100.00	100.00	78.29	37.28	29.13	<b>61.87</b>
0.7	0.3	100.00	100.00	81.33	47.81	40.34	<b>67.93</b>
0.6	0.4	99.56	100.00	83.81	57.89	55.32	<b>74.87</b>
0.5	0.5	96.89	100.00	85.71	66.67	67.65	<b>80.43</b>
0.4	0.6	96.00	100.00	86.86	74.56	76.89	<b>84.89</b>
0.3	0.7	95.56	100.00	86.10	78.73	82.77	<b>87.25</b>
0.2	0.8	93.78	100.00	85.90	82.02	87.39	<b>89.06</b>
0.1	0.9	89.78	100.00	84.57	84.65	89.08	<b>89.39</b>
0	1	87.56	100.00	83.43	84.43	89.64	<b>89.06</b>

The accuracy of RHE method's wavelet subbands are given in Table 5.3 and Table 5.4. Based on these two tables, by using LL+LH and LH+HL the maximum accuracy is 87.25% and 92% respectively. These accuracy rates were predictable since HL and LH are more robust in compare with LL.

After applying 10 fold cross validation, since the population of subset's images raised and train group did not only contains P00A+000E+00 images, the accuracy of system enhance significantly in compare with original selecting the probe and train groups.

Table 5.4: The accuracy of fix weighted of RHE method for LH2 and HL2 subbands in RQbHE method

Wavelet Subbands		Accuracy (%)					
LH+HL		Set 1	Set 2	Set 3	Set 4	Set 5	All sets
Factors							
1	0	87.56	100.00	83.43	84.43	89.64	<b>89.06</b>
0.9	0.1	90.22	100.00	85.90	85.53	90.06	<b>90.19</b>
0.8	0.2	92.89	100.00	87.43	87.28	90.20	<b>91.16</b>
0.7	0.3	93.33	100.00	89.71	87.94	90.48	<b>91.92</b>
0.6	0.4	94.22	100.00	90.10	88.16	90.06	<b>92.00</b>
0.5	0.5	94.67	100.00	90.29	87.94	87.25	<b>91.20</b>
0.4	0.6	96.44	100.00	90.67	87.72	83.89	<b>90.40</b>
0.3	0.7	94.22	100.00	90.48	85.09	77.87	<b>87.84</b>
0.2	0.8	92.89	100.00	89.33	79.82	67.79	<b>83.42</b>
0.1	0.9	89.33	99.78	86.67	72.37	55.88	<b>77.44</b>
0	1	88.00	99.12	84.19	62.94	43.00	<b>70.96</b>

Table 5.5: The accuracy of combination of LL2 and LH2 with fix weights for five different approaches

Wavelet Subbands		Accuracy (%)				
LH+HL		Methods				
Factors		None	HE	QbHE	RHE	RQbHE
1	0	34.26	37.16	38.26	53.75	49.79
0.9	0.1	38.22	42.59	44.07	57.49	56.48
0.8	0.2	42.51	48.74	49.79	61.87	62.21
0.7	0.3	47.01	55.35	56.27	67.93	69.53
0.6	0.4	51.64	62.84	63.43	74.87	76.01
0.5	0.5	55.60	69.82	70.08	80.43	81.61
0.4	0.6	59.72	77.15	76.77	84.89	86.11
0.3	0.7	63.80	82.32	81.78	87.25	88.05
0.2	0.8	67.17	86.11	85.40	89.06	89.10
0.1	0.9	68.81	87.67	86.45	89.39	89.56
0	1	68.86	88.05	86.07	89.06	89.39

Table 5.6: The accuracy of fix fusion method for LL and LH subbands in RQbHE approach

Wavelet Subbands		Accuracy (%)					
LL+LH		Set 1	Set 2	Set 3	Set 4	Set 5	All sets
Factors							
1	0	97.78	82.02	53.52	27.41	25.63	<b>49.79</b>
0.9	0.1	98.67	92.76	61.71	33.99	30.53	<b>56.48</b>
0.8	0.2	99.11	96.93	69.33	42.11	35.99	<b>62.21</b>
0.7	0.3	99.11	99.78	76.00	51.97	47.34	<b>69.53</b>
0.6	0.4	98.22	100.00	80.76	62.28	58.96	<b>76.01</b>
0.5	0.5	96.89	100.00	83.81	71.05	70.17	<b>81.61</b>
0.4	0.6	95.11	100.00	85.90	78.07	79.69	<b>86.11</b>
0.3	0.7	94.22	100.00	85.90	81.36	84.31	<b>88.05</b>
0.2	0.8	92.00	100.00	85.71	83.55	87.25	<b>89.10</b>
0.1	0.9	91.11	100.00	85.71	84.87	88.24	<b>89.56</b>
0	1	89.33	100.00	85.14	84.21	89.08	<b>89.39</b>

Table 5.7: The accuracy of fix weighted for RQbHE method for LH and HL subbands

Wavelet Subbands		Accuracy (%)					
LH+HL		Set 1	Set 2	Set 3	Set 4	Set 5	All sets
Factors							
1	0	89.33	100.00	85.14	84.21	89.08	<b>89.39</b>
0.9	0.1	91.56	100.00	86.10	85.96	89.36	<b>90.24</b>
0.8	0.2	92.44	100.00	87.62	85.96	89.92	<b>90.82</b>
0.7	0.3	94.22	100.00	89.14	86.40	89.22	<b>91.20</b>
0.6	0.4	94.67	100.00	90.86	87.72	87.39	<b>91.33</b>
0.5	0.5	95.56	100.00	91.81	87.94	85.01	<b>90.95</b>
0.4	0.6	96.89	100.00	92.38	86.18	78.99	<b>89.06</b>
0.3	0.7	97.78	100.00	91.81	81.36	71.29	<b>85.77</b>
0.2	0.8	94.67	100.00	89.33	75.88	59.80	<b>80.43</b>
0.1	0.9	92.89	100.00	86.29	66.01	47.90	<b>74.12</b>
0	1	90.67	99.12	83.81	56.58	37.96	<b>68.39</b>

The earlier method shows the effect of adaptive face recognition however; in new method I could investigate the best approximation to choose the better accuracy for recognition system. The results in Table 5.8 illustrated the accuracy rate of all approaches with benefit of 10-fold valid classification. In this table the whole accuracy rates are significantly greater than last results. This increase is just about changing in probe and test images that NN classifier used. In previews tables, results were achieved by using 2376 images for probe and 38 images for train groups and the train group is include the well-lit photos. Results in this technique demonstrate the effectiveness of recognition system to illumination. Then, results in Table 5.8 show the accuracy of methods in illumination normalization and feature extraction technique. High accuracy rates in this table illustrate that DWT extract important features of photos that makes the classifier more reliable even for None method.





Comparable to data in Table 5.2 here the minimum error is refers to RHE and RQbHE approaches. Since results in Table 5.8 for LH and HL subbands are very close, combing them with LL subband cannot present useful information.


Table 5.8: Identification accuracy rates correspond to each approach achieved by 10 fold cross validation for Extended Yale B data base

		<b>Accuracy (%)</b>					
<b>Wavelet subbands</b>	<b>Method</b>	<b>Set1</b>	<b>Set2</b>	<b>Set3</b>	<b>Set4</b>	<b>Set5</b>	<b>All sets</b>
<b>LL2</b>	None	100	100	86.10	52.63	41.46	<b>70.71</b>
	HE	100	100	89.71	81.80	92.02	<b>91.96</b>
	QbHE	99.62	99.78	85.71	56.36	83.61	<b>83.72</b>
	RHE	100	100	97.14	90.57	98.32	<b>97.10</b>
	RQbHE	88.65	100	96.93	76.38	78.51	<b>94.68</b>
<b>LH2</b>	None	100	100	100	99.56	88.24	<b>96.44</b>
	HE	100	100	100	99.78	99.16	<b>99.71</b>
	QbHE	100	100	100	99.78	99.16	<b>99.71</b>
	RHE	100	100	99.62	100	99.16	<b>99.67</b>
	RQbHE	99.67	100	100	99.81	99.78	<b>99.16</b>
<b>HL2</b>	None	100	100	99.43	93.42	80.25	<b>92.79</b>
	HE	100	100	99.81	98.25	99.16	<b>99.38</b>
	QbHE	100	100	100	98.03	99.02	<b>99.34</b>
	RHE	100	100	99.62	98.90	99.16	<b>99.46</b>
	RQbHE	99.38	100	100	99.62	98.25	<b>99.30</b>

Table 5.9: GLQ and RLQ of each region for RHE and RQbHE methods when threshold sets to 0.8

Method	Subsets									
	1		2		3		4		5	
None										
GLQ	0.9987		0.8880		0.5647		0.4197		0.1481	
RLQ	0.9978	0.9995	0.9964	0.7773	0.7202	0.4013	0.3005	0.5467	0.0794	0.13162
	0.99897	0.9982	0.9999	0.77858	0.7116	0.3576	0.2588	0.5459	0.1099	0.2686

RHE										
GLQ	0.9867		0.9865		0.9866		0.9860		0.9845	
RLQ	0.9865	0.9865	0.9866	0.9866	0.9866	0.9864	0.9844	0.9867	0.9804	0.9840
	0.9867	0.9870	0.9865	0.9865	0.9867	0.9865	0.9866	0.9863	0.9865	0.9869

RQbHE										
GLQ	0.9987		0.9687		0.9866		0.9860		0.9845	
RLQ	0.9978	0.9995	0.9964	0.9866	0.9866	0.9864	0.9844	0.9867	0.9804	0.98404
	0.99897	0.9982	0.9999	0.9865	0.9867	0.9865	0.9866	0.9863	0.9865	0.98695

## **Chapter 6**

### **CONCLUSION AND FUTURE WORK**

#### **6.1 Conclusions**

In this thesis human face recognition is investigated in the presence of variation in lighting condition. A quantitative quality value is used which is named the luminance quality (LQ) index. LQ is used for the whole pixels and for four regions of images. Based on [1] the probe and train images are selected from Extended Yale database B which confirmed the effectiveness of illumination normalization procedures. By changing the group of train and probe images the effectiveness of feature extraction method and illumination normalization procedure are investigated.

Moreover, Wavelet transform gave us different features of images that presented specific information of images. Since the images captured in extreme variation in lighting conditions, combining different features that obtained from wavelet transform led us to a more accurate recognition system.

Finally, adaptive fusion approach, done by choosing appropriate masses according to LQ index of probe images, significantly increased the accuracy of the face recognition system.

#### **6.2 Future Work**

The future work will contain other applications of wavelet transform in face detection and face recognition. 3D wavelet transform can use in 3D face recognition

and even face detection. Moreover, it can utilize the features of human brain 3D MRI images in an automated system to detect the position and even recognize the type of tumors. Beside choosing better databases and different classification approaches according to probe and train groups, wavelet base adaptive automated systems enable us to reach a wider range of features of data that lead to more accurate and reliable systems.

## REFERENCES

- [1] Sellaheewa, H., & Jasim, S. A. (2010). Image-Quality-Base adaptive face recognition, *IEEE Trans. Instrument and measurement*, 9(6), 805–813
  
- [2] The Extended Yale Face Database. (2015, September 3). Retrieved from <http://cvc.yale.edu/projects/yalefaces/yalefaces.html>.
  
- [3] Georghiades, A. S., Belhumeur, P. N., & Kriegman, D. J. (2001). From Few to Many: Illumination Cone Models for Face Recognition under Variable Lighting and Pose. *IEEE Trans. Pattern Anal. Mach. Intelligence*, 23(6), 643-660.
  
- [4] Kelly, M. D. (1970). Visual identification of people by computer. *Technical Report AI-130*, Stanford AI Project, Stanford, CA.
  
- [5] Bledsoe, W. W. (1964). The model method in facial recognition. *Technical Report PRI: 15*, Panoramic Research Inc., Palo Alto, CA.
  
- [6] Kanade, T. (1977). Computer recognition of human faces. Basel and Stuttgart, Birkhauser Verlag.
  
- [7] Phillips, P. J., Rauss, P., & Der, S. (1996, October). FERET (face recognition technology) Recognition Algorithm Development and Test Report. *Technical*

report arl-tr 995, U.S. Army Research Laboratory.

- [8] Zhao W., Chellappa, R., Phillips, J., & Rosenfeld, A. (2003). Face recognition in still and video images: *A literature survey*. *ACM Computing Survey*, 35, 399–458.
  
- [9] Samaria, F., & Young, S. (1994). Face HMM based architecture for face identification. *Image and Computer Vision*, (1994).
  
- [10] Kirby, M., & Sirovich, L. (1990, January). Application of the karhunen-loeve procedure for the characterization of human faces. *IEEE Transactions on Pattern Analysis Machine Intelligence*, 12(1), (pp. 103–108).
  
- [11] Turk, M. A., & Pentland, A. P. (1991). Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, 3(1), (pp. 71–86).
  
- [12] Delac, K., Grgic, M., & Grgic, S. (2005). Independent comparative study of PCA, ICA, and LDA on the FERET data set. *Wiley Periodicals*, 15, (pp. 252–260).
  
- [13] Wilder, J., Phillips, P. J., Jiang, C. H., & Wiener S. (1996). Comparison of visible and infra-red imagery for face recognition. *International Conference on Automatic Face and Gesture Recognition*, (pp. 182–187).

- [14] Nicholl, P., & Amira, A. (2008, January). DWT/PCA face recognition using automatic coefficient Selection. *IEEE International Workshop on Electronic Design, Test and Applications (DELTA '08)*, (pp. 390–393).
- [15] Shen, L., Ji, Z., Bai, L., & Xu, C. (2007). DWT based HMM for face recognition. *Journal of Electronics*, 24(6), 835–837.
- [16] Wright, J., Yang, A. Y., Ganesh, A., Sastry S. S., & Ma Y. (2009). Robust face recognition via sparse representation. *Transactions on Pattern Analysis and Machine Intelligence*, 31(2), (pp. 210–227).
- [17] Sadun, L. (June, 2008). *Applied Linear Algebra: The Decoupling Principle.*, Second Edition, Department of Mathematics, University of Texas , Austin, TX.
- [18] Kryszczuk, K., Richiardi, J., Prodanov, P., & Drygajlo, A., (2005, September). Error handling in multimodal biometric systems using reliability measures In *Proceedings. 13th EUSIPCO*.
- [19] Grother, P., & Tabassi, E. (2007, April). Performance of biometric quality measures, *IEEE Transaction. Pattern Anal*, 29(4), (pp. 531–543).
- [20] Lai, J. H., Yuen, P. C., & Feng, G. C. (2001, January). Face recognition using holistic Fourier invariant features, *Pattern Recognition*, 34(1), (pp. 95–109).

- [21] Jassim, S. A., & Sellahewa, H. (2006, September). Multi-stream face recognition on dedicated mobile devices for crime-fighting, In Proceedings. SPIE Opt. Photon. Counterterrorism Crime Fighting II, 6402.
- [22] Sellahewa, H. (2006). Wavelet-based automatic face recognition for constrained devices, Ph.D. dissertation, University of Buckingham, Buckingham, U.K.
- [23] Sellahewa, H., & Jassim, S. A. (2008, September). Illumination and expression invariant face recognition: Toward sample quality-based adaptive fusion. *IEEE International Conference in Biometrics*, (pp. 1–6.)
- [24] Georghiadis, A. S., Belhumeur, P. N., & Kriegman D. J. (2001, Jun). From few to many: Illumination cone models for face recognition under variable lighting and pose, *IEEE Transaction. Pattern Anal. Mach. Intel*, 23(6), (pp. 463–474).
- [25] Lee, K. C., Ho, J., & Kriegman, D. J. (2005, May). Acquiring linear subspaces for face recognition under variable lighting, *IEEE Trans. Pattern Anal, Intell*, 27(5), (pp. 684–698).
- [26] Frequency Spectrum. (2015, September 3). Retrieved from <http://glowingpython.blogspot.com.tr/2011/08/how-to-plot-frequency-spectrum-with.html>
- [27] AT & T: The database of faces (formerly: The ORL database of faces). (2015, January 3). Retrieved from



<http://www.cl.cam.ac.uk/Research/DTG/attarchive/facedatabase.html>.

- [28] Tzu, H., Henry, L. (2004). Wavelet Analysis for Image Processing. Graduate Institute of Communication Engineering, National Taiwan University, Taipei, ROC.