Performance Comparison between SIFT and SURF Descriptors for Face Recognition using Wavelet Transforms

by

Musa M.AMEEN

Submitted to the Institute of Graduate Studies in Science and Engineering

in partial fulfillment of

the requirements for the degree of

Master of Science

in

Computer Engineering

Mevlana (Rumi) University

2016

PERFORMANCE COMPARISON BETWEEN SIFT AND SURF DESCRIPTORS FOR FACE RECOGNITION USING WAVELET TRANSFORMS

submitted by **Musa M.AMEEN** in partial fulfillment of the requirements for the degree of Master of Science in Computer Engineering at Mevlana (Rumi) University.

APPROVED BY:

Examining Committee Members:

Assist. Prof. Dr. Alaa ELEYAN . . . .
(Thesis Supervisor)

Assist. Prof. Dr. Armağan ÖZKAYA . . . .

Assoc. Prof. Dr. Halis ALTUN . . . .

Assist. Prof. Dr. Armağan ÖZKAYA . . . .
Head of Department, Computer Engineering

Prof. Dr. Ali SEBETCİ . . . .
Director, Institute of Graduate Studies in Science and Engineering

DATE OF APPROVAL (…..../…..../2016)

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Musa M.AMEEN

Signature:

ABSTRACT

PERFORMANCE COMPARISON BETWEEN SIFT AND SURF DESCRIPTORS FOR FACE RECOGNITION USING WAVELET TRANSFORMS

Musa M.AMEEN

M.Sc. Thesis, 2016

Thesis Supervisor: Assist. Prof. Dr. Alaa ELEYAN

Keywords: Face Recognition, Speeded-Up Robust Features, Scale-Invariant Feature Transform, Discrete Wavelet Transform, Gabor Wavelet Transform

Automatic face recognition is a major research area in computer vision which aims to recognize human face without human intervention. Significant developments in this field have shown that in many face recognition applications the automated techniques outperform human. Face recognition involves in many fields, such as machine vision, pattern recognition, bioinformatics, etc., and has become one of the hottest subjects. The key problem in face recognition is how to find a feature set to identify a face. Many algorithms about feature extraction have been proposed, which mainly include three aspects: face geometrical, facial and statistical features.

In this thesis, the conventional SIFT and SURF performances are tested in face recognition. They provide high performance. However, this performance can be improved further by transforming the input into different domain from the real time. Hence, we apply Discrete Wavelet Transform (DWT) or Gabor Wavelet Transform (GWT) at the input face images which provides us denser and clearer images compared to those by the conventional SIFT or SURF. Simulations show that the proposed approaches based on DWT or GWT using SIFT or SURF provides very high performance compared to the conventional algorithms.

ÖZET

Dalgacık Dönüşümleri Kullanarak Yüz Tanıma İçin SIFT Ve SURF Tanımlayıcılarının Arasındaki Performanslarının Karşılaştırılması

Musa M.AMEEN

Yüksek Lisans Tezi, 2016

Tez Danışmanı: Yrd. Doç. Dr. Alaa ELEYAN

Anahtar Kelimeler: Yüz tanıma, Hızlandırılmış Dayanıklı Özellikler, Sabit Ölçek Öznitelik Dönüşümü, Ayrık Dalgacık Dönüşümü, Gabor Dalgacık Dönüşümü

Otomatik yüz tanıma, insan yüzünü insan müdahalesi olmadan tanımayı amaçlayan bilgisayarla görmede başlıca bir araştırma alanıdır. Bu alandaki önemli gelişmeler, birçok yüz tanıma uygulamasında otomatik tekniklerin insanlardan daha üstün olduğunu göstermektedir. Yüz tanıma, makine görmesi, örüntü tanıma, biyoinformatik vb. gibi birçok alanı kapsamaktadır ve en gözde konulardan birisi olmuştur. Yüz tanımadaki asıl problem, bir yüzü tanımlamak için öznitelik setinin nasıl bulunacağıdır. Öznitelik çıkarımı hakkında birçok algoritma önerilmiştir ve bunlar esasen üç bakış açısını içerir: yüz geometriği, yüze ait ve istatistiksel öznitelikler.
Bu tezde, geleneksel SIFT ve SURF performansı yüz tanımada test edildi. Bunların yüksek performans gösterdiği görüldü. Fakat bu performans, veriyi gerçek zamandan farklı etki alanına dönüştürerek daha da geliştirilebilir. Bundan dolayı, Ayrık Dalgacık Dönüşümünü (ADD) ya da Gabor Dalgacık Dönüşümünü (GDD) yüz resimleri verilerine uyguladık ki bu bize geleneksel SIFT veya SURF'e göre daha yoğun ve daha net resimler sağladı. Simülasyonlar, SIFT veya SURF'u kullanan KDD veya GDD tabanlı önerilen yaklaşımların geleneksel algoritmalara kıyasla çok yüksek performans sağladıklarını gösterdi.

Dedicated to my Family, Friends, and Noora (dabdoob ☺)

ACKNOWLEDGEMENTS

Firstly, I give all love, thanks, honors and glories to our creator, ALLAH the sustainer, the cherisher for making everything achievable.

I would like to express my deepest gratitude to my supervisor Assist. Prof. Dr. Alaa ELEYAN, for his guidance, advice, criticism, encouragement and insight throughout the research. I cannot thank you enough for the unlimited support you have extended to me sir!
I dedicate this thesis to my parents, brothers and sisters, for their unconditional love and support. Despite of the distance, you were always helping, assisting and making me feel home again. I love you all.
Among the people, I would like to specially deliver my sincere gratitude to my friends; Muhammed Bazzaz, Muhammed Anwar, Berhem Jalal, Bayram Uzunçam, Selim Kaplan, Abubakir Ashir, Jawad Muhammed and Gaddafi Sani Shehu.
In addition, I wish to thank Ibrahim Bala for his grateful work in preparing database images.

At last, it is an honor for me to offer my gratitude to the Ishik University, Erbil for giving me this magnificent opportunity to undergo for my master’s degree program.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS/ABBREVIATIONS

Symbol	**Explanation**
2D	Two Dimensional
2D-DWT	Two Dimensional Discrete Wavelet Transform
3D	Three Dimensional
AAM	Active Appearance Models
DoG	Difference of Gaussian
DWT	Discrete Wavelet Transform
HH	High pass High pass subband
HL	High pass Low pass subband
ICA	Independent Component Analysis
kNN	k-Nearest Neighbor
LH	Low pass High pass subband
LL	Low pass Low pass subband
LBP	Local Binary Pattern
LH	Low pass High pass subband
LDA	Linear Discriminant Analysis
LoG	Laplacian of Gaussian
ORL	Olivetti Research Laboratory
PCA	Principal Component Analysis
PUT	Poznan University of Technology
SIFT	Scale-Invariance Feature Transform
SURF	Speeded-Up Robust Features

CHAPTER ONE

INTRODUCTION

1.1. Overview

During the past two decades, both consumer and business worlds have witnessed a swift growth in video and image processing to fulfill the needs of object detection for various applications such as data query and object retrievals. One of the most widely researched areas, thoroughly investigated for various applications is face recognition.

Conventionally, computer vision systems have been used in some particular tasks, like executing tedious checks on the visual tasks of assembly line. Nowadays development in this subject has moved in the direction of a global vision applications like face recognition, video tracking methods, biometrics, surveillance, robots, human-machine interaction and database indexing and many more applications that have face detection as the primary building block of their systems.

Face recognition is a significant research problem covering many areas and fields. The reason behind this is face recognition, furthermore to having various real-world applications like social media, identity verification, security and surveillance systems, is a critical human behavior that is important for active communications and interactions among people.

1.2. Related work

Face recognition is one of the most common biometric systems. Due to its higher acceptability rate, researchers have developed various algorithms for face recognition purpose. The process of recognition using these algorithms has been described as a difficult task because of the similarity nature or shapes of human faces [1]. Despite the difficulties encountered in designing these systems, several reasons contributed to the enormous attention in automatic digital image processing and video processing in a different types of applications, which include wide availability of powerful and low-cost desktop and embedded computing systems. Also, it has been described as one of the best applications of image processing and analysis [2]. Different statistical methods and algorithms such as Principal Component Analysis or Eigenface (PCA) [3], Local Binary Pattern (LBP) [4] and Independent Component Analysis (ICA) [5] algorithms has been developed for face recognition purposes. Due to continuous research, a significant improvement in recognition performance is obtained over years as in PCA and ICA comparison

[6], PCA and Linear Discriminant Analysis (LDA) [7], and improving eigenface recognition as in [8].

Turk and Pentland [9] initially developed face recognition using eigenface techniques. The approach provides for a form of computational pattern recognition for the face. The aim of using eigenfaces term is, because the primary components of the face is represented by mathematical algorithms using eigenvectors. The eigenface features are represented by weights later weights are used to compare with individual faces from a database for identification purposes.

Characteristic faces are more easily recognized than typical faces. Low frequency bands contain information that determines the sex of the specific subjects, while recognition of individuals depends on the high frequency features. The global description is determined by the low frequency, while the finer descriptions high frequency modules give to the finer information required for the identification procedure [10, 11, 12]. Since upper part of face contains more features, it's more important and useful than the lower part [13].

Bruce in [14] made an experiment on Margaret Thatcher's face by imposing low spatial frequency of its face on Tony Blair's face using its high spatial frequency components. When, the image is seen in near distance Tony Blair's face appears, but when distance becomes far then Tony Blair's face disappears and Margaret Thatcher's face appears. By this experiment, it can be demonstrated that faces distinctive textures to recognize that particular face contained within a specific range of spatial frequencies.

The advances in face recognition have reached an advance stage in recent years. The algorithms developed by researchers now detect even age range of the enrolled individual besides the faces identification. Age prediction can be done by using a set of gallery images to test and train the model that is in turn predicts age of face images. In [10] age prediction and face recognition are analyzed and identified using Eigenface. They projected a new face image into this face space and then comparing its position in the face space with those of recognized faces for age prediction. The system preserves the distinctiveness of the person while applying a realistic recognition properties on adult facial images dividing it into histogram of 11 bins between 15 to 60 years old with 5 years old range for each bin.

1.3. Face Recognition

Facial recognition can be categorized as a visual pattern recognition job. The 3D human face, which can be easily changed and its subject to change in expression, light, pose etc. must be recognized. This recognition process can be applied on different types of input images like; a 2D face image, Stereo 2D images and 3D laser scans.

Video sequence is a still image with a time dimension. The identification of face images in a video sequence is easier, precise and has advantage over still images because face images of a person cannot be changed from two frames taken in a video sequence. While changing face images in a still image is easier, it's difficult to identify a person.

There are four core steps in a face recognition systems which are; face detection, face preprocessing, feature extraction, and feature matching, as shown in Figure 1.1. The steps are described in the following sections.

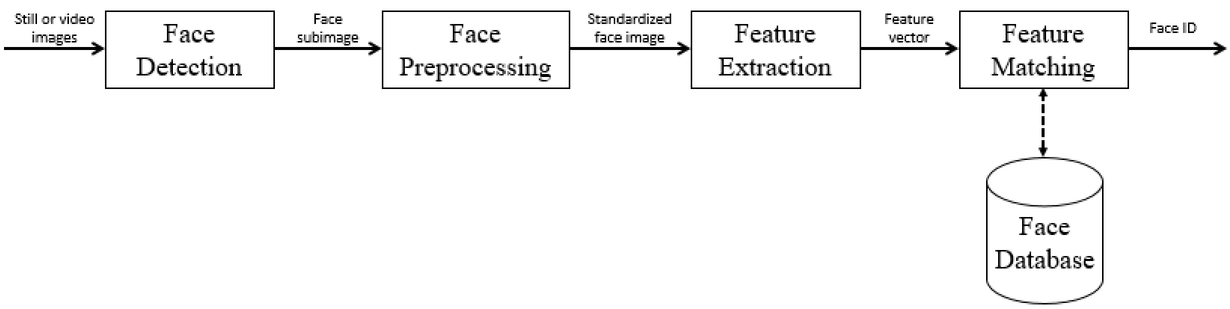


Figure 1.1. The four general steps in facial recognition.

1.3.1. Face Detection

One of the main points in face detection is localization of face images. When we apply this on a video input, the main advantage is in tracking the face images in between multiple frames; this has the advantages of reducing computational time and preserving the identity of an individual face image between frames.

1.3.2. Face Preprocessing

Face preprocessing step aims to stabilize the output of face detection step, so that a better feature extraction can be gained. Face preprocessing includes different processes depending on the application, like; alignment (translation, rotation, scaling), illumination normalization, and illumination correlation.

1.3.3. Feature Extraction

Feature extraction's goal is to extract a solid and stable set of discriminating specific features and keypoints of the facial images. Algorithms used in feature extraction include: SURF [19], SIFT [18], PCA [3], LDA [7]…

1.3.4. Feature Matching

The last and actual recognition process feature matching. The feature vector that's obtained from the feature extraction step is matched to classes (individuals) of face images already stored in a database. There are different types of matching algorithms and they vary from the k-Nearest Neighbor to advanced algorithms like Neural Networks.

1.4. Problem Definition

An overall statement of the problem can be expressed as follows: given an image of an individual face image, identify subject using an enrolled database of face images.

Face recognition is one of the difficult problems because of the overall similar shape of faces combined with many differences between images of the same face. Recognition of face images taken from a wild environment is a very complex and difficult process: illumination condition may change with a big difference; face shapes and expressions are also differ from time to time; face may appear at diverse directions. Furthermore, depending on the type of application, handling and extracting face aging should be required.

Although existing applications and algorithms perform well under controlled environments, the difficulties with the performance and recognition rate still remained unsolved. The proposed approach, tries to increase performance rate of recognition over the conventional algorithms.

Since the algorithms and methods that are used and available in the face recognition applications might rely on the task and job of the system, there are two general classes of face recognition that can be identified [15]:

1. Finding an individual within a large database of face images.
2. Identifying specific people in real time like tracking systems.

In this thesis, we mainly focus on the first task. Our goal is to provide a better performance in recognizing the correct faces from all subjects in the database.

1.5. Organization of the Thesis

The thesis is organized as follows; Chapter 1 serve as general introduction and it covers the face recognition. The chapter also states the contributions of the thesis work.

In Chapter 2, the feature extraction algorithms that are being used in proposed approach, were given in detail. The working mechanism and properties of these feature descriptors are discussed.

Chapter 3 presents the detailed methodology followed in carrying out the work and the explanation of the proposed approaches. The feature extraction algorithms used and transformation algorithms were also discussed.

Chapter 4 gives the simulations results and comments on every result obtained. This chapter also covers the general discussion of results.

Chapter 5 gives the work conclusions and recommendations for future work.

CHAPTER TWO

Feature Extraction Algorithms

2.1. Overview

The basics of feature extraction is the dimensionality reduction, by choosing some dominant or distinct features that can best represents the face image with less distortion to the original image. Appropriate algorithms are used to extract the salient features from the relevant patterns. The face representation is done in two ways: the first way is the appearance (holistic) texture features and is applied to the whole face image; the second way is the component based which utilizes the linear relationships between the facial features such as eyes, mouth, and nose [16]. The unpopular component (feature) based approaches utilize some special facial points, and characterize them by applying a bank of filters which extract the typical texture around them [17]. The holistic approaches attract more attention than the component based methods. In this thesis work, two of the popular holistic or appearance based methods are studied briefly to extract features from face images.

2.2. Scale-Invariant Feature Transform

SIFT (Scale-Invariant Feature Transform) developed in 2004 by D. Lowe from University of British Columbia [18]. SIFT is able to detect and extract distinctive features from any face image to perform stable and robust matching between face images of the same subject (person) with various facial expressions, face poses, and the features extracted form face images are scale, illumination and rotation invariance.

Figure 2.1 shows four important stages involved for detecting keypoints in the SIFT algorithm.

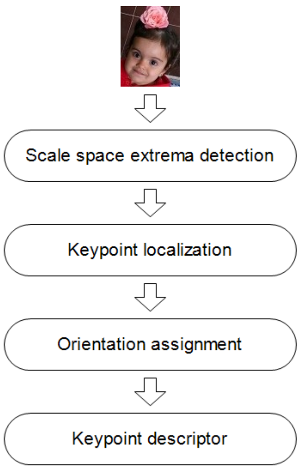


Figure 2.1. SIFT features extraction process.

2.2.1. Scale-space Extrema Detection

It is noticeable that, with different scales we can't detect keypoints using the same window. With small corner it is working fine. But larger windows are needed to detect larger corners. For this reason, we used scale-space kernels. In it, Laplacian of Gaussian (LoG) of the fade image is found with different σ values. LoG has been used as a blob detector to find blobs in different scales because of variation in σ. Briefly, σ used as a scaling parameter. For example, for small corners Gaussian kernel that has low σ outputs high value while for larger corners Gaussian kernel that has high σ fits well. So, local maxima can be found across the scale and space which provides us a set of (x, y, σ) values that proves, there is a potential keypoint at (x, y) at σ scale [18].

But this LoG is slightly costly, so SIFT algorithm applies Difference of Gaussians (DoG) on LoG which is an approximation of LoG. DoG is defined as the difference of Gaussian blurring of a face image with two different σ, let it be σ and $k\sigma$. This procedure is applied for various scales (octaves) of the face image that's found in Gaussian Pyramid as shown in Figure 2.2.

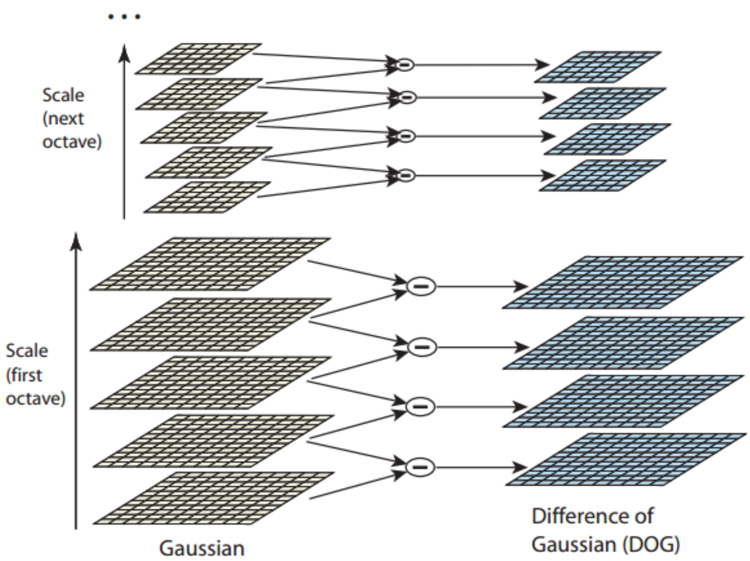


Figure 2.2. Gaussian pyramid [18].

As soon as DoG is found, images are examined to find local minimum and maximum over scale and space. For example, 1 pixel in a face image is compared with 26 pixels which is 8 neighbors, 9 pixels in previous scale and 9 pixels in next scale. It is defined as a possible keypoint, if it's a local minimum or local maximum among pixels. Simply it means that keypoint that's found is the best represented in that scale as shown in Figure 2.3.

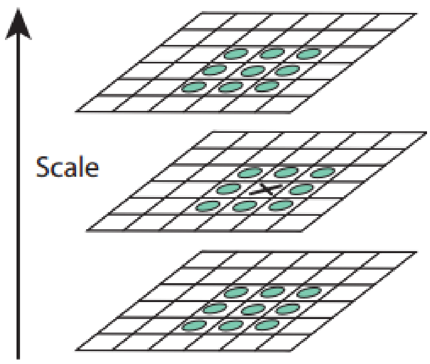


Figure 2.3. Scale space for keypoint [18].

2.2.2. Keypoint Localization

Whenever possible keypoints' positions are found, to get more accurate results from them, they have to be refined. For refining keypoints, Taylor series used as an development of scale space to get more precise position of extrema and if the intensity at this extrema is below the threshold level. Difference of Gaussians has more edge responses, so the edges need to be vanished, for this, a notion is used that is same as Harris corner detector. The principal curvature is computed using a 2x2 Hessian matrix (H) [18]. Harris corner detector states that for any edge, one eigenvalue is bigger than the other eigenvalue. So here they defined a simple function, that is, if the ratio is greater than a threshold value, that keypoint is rejected.

This way, any keypoints and edge keypoints with low-contrast were removed, so the remaining is robust and strong keypoints.

2.2.3. Orientation Assignment

Image rotation invariance is achieved by assigning orientation to each keypoint. Depending on the scale, the direction and gradient magnitude of the neighborhood pixels around the keypoint position are taken and calculated. An orientation histogram with 36 bins is formed to cover 360 degrees. It is weight is calculated by Gaussian-weighted circular window and gradient magnitude with σ equal to 1.5 times the scale of keypoint. The calculations of orientation is done by taking the highest peak in the histogram and any other peak above 80% is also considered and taken. It generates keypoints with different directions, but same scale and position. The contribution goes to the stability of matching.

2.2.4. Keypoint Descriptor

A neighborhood of 16x16 scale pixels around the keypoints are taken. It is divided into 16 sub-regions of 4x4 size. 8 bin orientation histogram is created for each sub-block. So a total of 128 bin values are obtained. And, the values are denoted as a vector to create keypoint descriptor as shown in Figure 2.4.

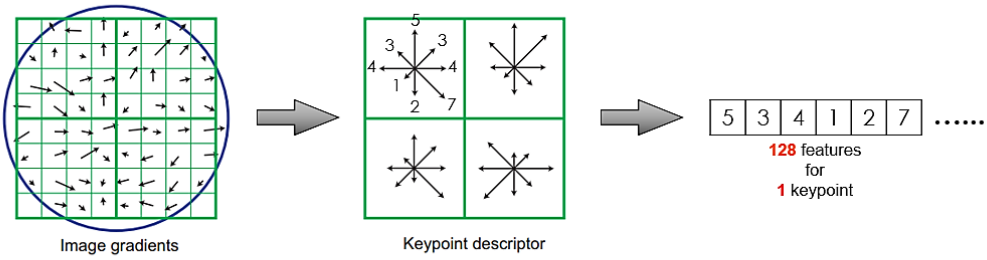


Figure 2.4. SIFT keypoint descriptor [18].

Along with this, some procedures are applied to gain robustness alongside lighting changes, rotation etc. Figure 2.5 shows distinctive keypoints detected in face image.

Figure 2.5. SIFT keypoints detected in a face image.

2.3. Speeded-Up Robust Features

SURF (Speeded-Up Robust Features) was created by Bay, Tuytelaars, and Van Gool in 2006 from ETH Zurich [19]. SURF algorithm is a robust keypoint detector of local features in a face image. It is a developed version of Shift-Invariant Feature Transform (SIFT) and Hessian blob detectors integer approximation to the determinant is calculated with integral images.

In SIFT, Lowe [18] approximated LoG using DoG for scale-space step. SURF goes a slightly more than Laplacian of Gaussian using Box Filter. Figure 2.6 shows approximation demonstration. This approximations biggest advantage is that, it simply uses integral images to calculate the convolution with box filters. Also, for different scales it can be done in parallel. The determinant of Hessian matrix is major component of SURF for both position and scale.

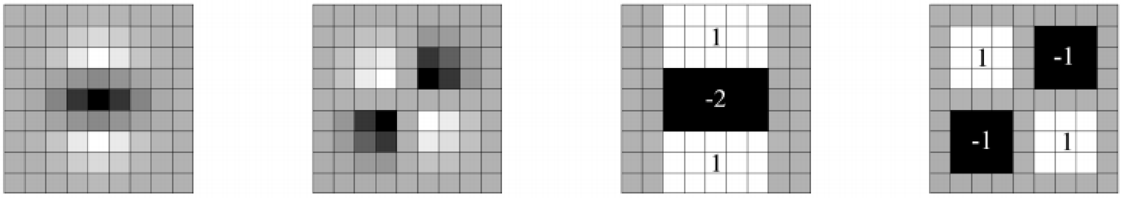


Figure 2.6. The box filters of approximations of Gaussian second order partial derivative [19].

Orientation assignment achieved using wavelet responses in vertical and horizontal direction for a neighborhood of size 6 multiplied by scale in which keypoint is detected. Suitable Gaussian weights are also performed on it. The calculation of the sum of all responses within a sliding orientation window of 60° estimates the main orientation. Exciting part is that, simply integral images can be used to find out wavelet response at any scale. Rotation invariance is not requires in many applications, so finding this orientation is not needed, by this speed of process increases. SURF delivers an extra method called Upright-SURF or U-SURF which increases speed and is strong up to $\pm 15^{\circ}$. Figure 2.7 shows distinctive SURF keypoints of a face image.

Figure 2.7. SURF keypoints detected in a face image.

Around each keypoint a neighborhood of size 20sX20s is taken where s is scale at which keypoint is detected. It is then distributed into 4x4 subregions as shown in Figure 2.8. Vertical and horizontal wavelet responses are taken for each subregion and $v = (\sum dx\,,\ \sum dy\,,\ \sum|dx|\,,\ \sum|dy|)$ like, a vector is formed. As it's represented as a vector, it gives a total of 64 dimensions of SURF feature descriptor. So with the SURF keypoint feature vector dimension becomes low, it causes to increase the speed of computation and also increases speed of matching, providing better uniqueness of features.

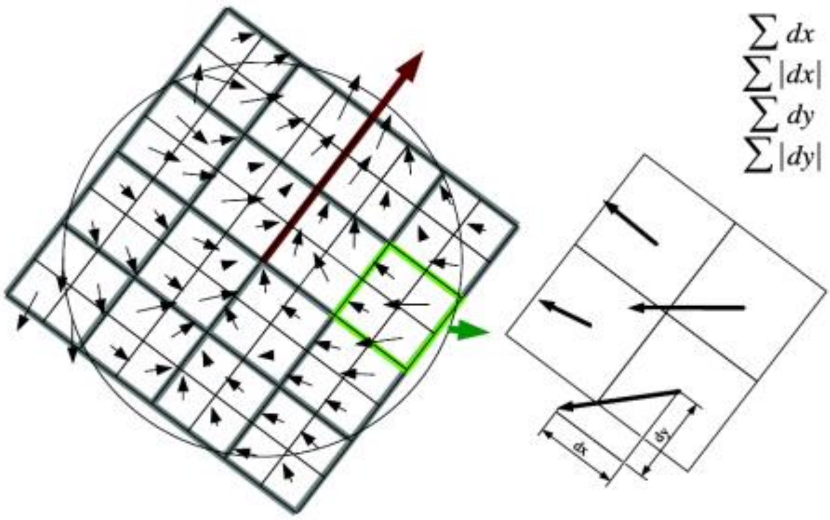


Figure 2.8. The demonstration of descriptor building [19].

When we use sign of Laplacian (trace of Hessian Matrix), it gives a significant improvement for underlying distinctive keypoints. Since it is already computed during detection, there is no additional computation cost. The bright blobs on dark backgrounds are distinguished using sign of the Laplacian from the reverse situation. In the feature matching step, features with the same contrast rate will be compared as shown in Figure 2.9. By this slight information it will make matching process faster, without affecting or decreasing the performance of descriptor.

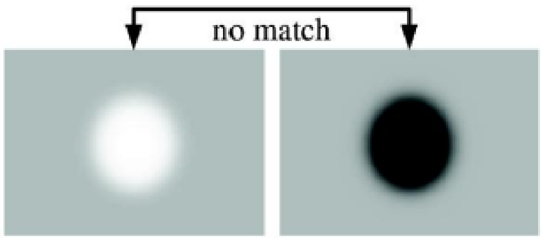


Figure 2.9. The fast index for matching [19].

In short, to improve the speed of the process SURF adds a lot of features in every step. SURF is good at handling images with blurring and rotation.

In SIFT and SURF keypoints of two images are matched by identifying their nearest neighbors (k-NN) which is shown in Figure 2.10. But sometimes, the second closest-match may be very close to the first. Due to noise or some other reasons. In that case, ratio of closest-distance to second-closest distance is taken into the consideration.

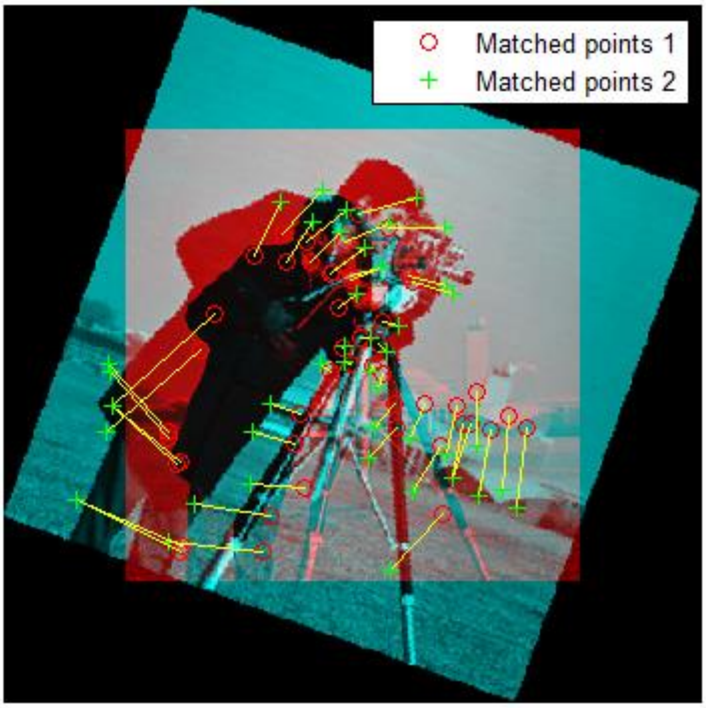


Figure 2.10. Example of point matching result.

2.4. Wavelet Transform

2.4.1. 2D-Discrete Wavelet Transform

The two dimensional discrete wavelet transform (2D-DWT) is basically a one dimensional analysis of a two dimensional signal [20]. It basically works on one dimension at a time, in a separable fashion it examines the rows and columns of an input image. At first it convolves the high pass and low pass kernels (filters) among the rows of an input face image. This creates two new images, where one image is a set of detailed row coefficients and the other a set of coarse row coefficients. In the next, step for the columns of each new image analysis kernels are convolved, which creates four different images called subbands or subimages. H defines rows and columns convolved with a high pass filter. Similarly, L defines rows and columns convolved with a low pass filter. For example, HL subband defines a subimage or subband that was produced using a high pass filter on the rows and a low pass filter on the columns. Figure 2.11 describes the whole procedure.

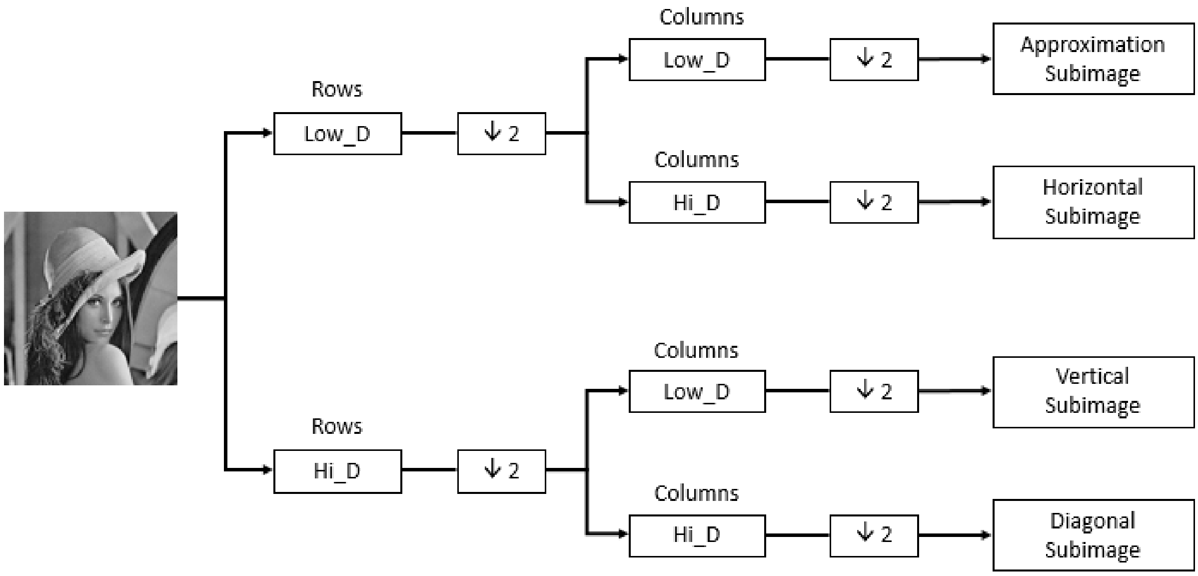


Figure 2.11. 2D-DWT, the high and low pass filters operate separately on the rows and columns to create four different subimages.

Each subimage or subband provides different information about the image as shown in Figure 2.12. The approximation (LL) subimage is an approximation of the image and removes all high frequency textures. The horizontal (LH) subimage eliminates high frequency textures along the rows and emphasizes high frequency textures along the columns so output is an image with emphasized vertical edges. The vertical (HL) subimage emphasizes horizontal edges, while the diagonal (HH) subimage emphasizes diagonal edges.

Figure 2.12. 2D-DWT transform on face image.

2.4.2. Gabor Wavelet Transform

Gabor functions developed by Dennis Gabor as a tool for detecting signals in noise. Gabor [21] showed that there exists a "quantum principle" for information; in order no signal can conquer less than certain minimal area in it, the conjoint time-frequency domain must be quantized for 1D signals. Gabor decomposition is described by its orientation and scaling sensitivity with directional microscope. Since curves have some low-level salient features in an image, such curves form low level feature map of the image intensity.

A Gabor wavelet filter is a Gaussian kernel function moderated by a sinusoidal plane wave (1).

$$\psi_g(\mathrm{x,y}) = \frac{f2}{\eta\gamma\pi}\exp(\beta^2 y'^2 - \alpha^2 x'^2)\exp(2\pi\, j\, f\, x')$$

$$x' = x\cos\theta + y\sin\theta, \tag{1}$$

$$y' = y\cos\theta - x\sin\theta,$$

where f defines central frequency of the sinusoidal plane wave, the anticlockwise rotation of the Gaussian defined by θ and the envelope wave defined by α, which is the sharpness of the Gaussian along the major axis parallel to the wave and the sharpness of the Gaussian minor axis perpendicular to the wave defined by β. To keep the ratio between frequency and sharpness constant $\gamma = f/\alpha$ and $\eta = f/\beta$ is defined [22]. (2) defines the 2D Gabor wavelet which has Fourier transform.

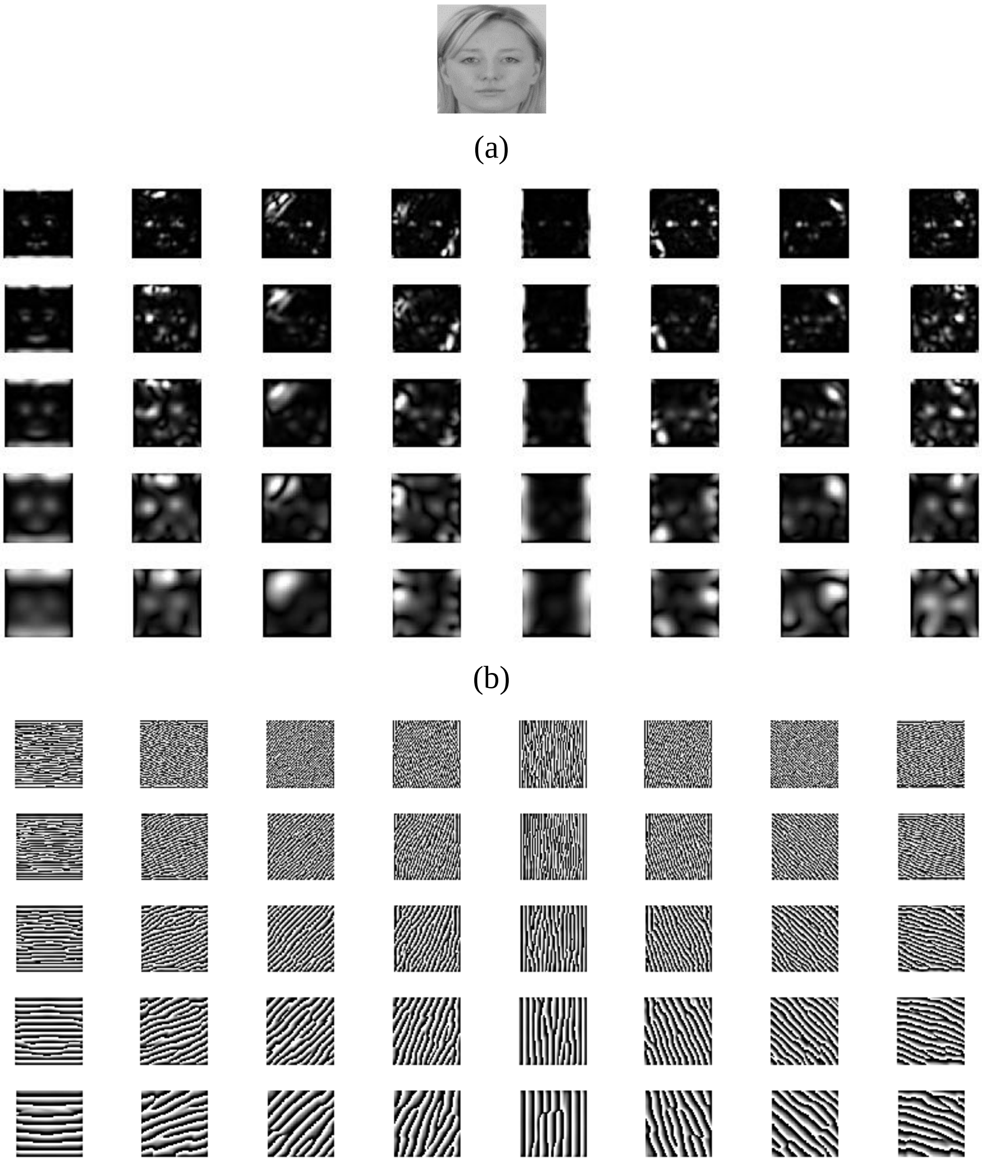
$$\psi_g\ (u,v) = \exp\left(-\pi^2\ \left(\frac{(u'-f)^2}{\alpha^2} + \frac{v'^2}{\beta^2}\right)\right)$$

$$u' = u\ cos\theta + v\ sin\theta, \tag{2}$$

$$v' = v\cos\theta - u\sin\theta.$$

Since GWT is mainly developed for vision applications and systems, one of the most significant applications it can be used for is face recognition. GWT usage in vision area and applications was first utilized by Daugman in the 1980s. Lately, a face recognition system based on Gabor wavelet was developed by B. S. Manjunath [23] et al. Then, research in using GWT in the face recognition applications is continued with attaching dynamic elastic graph matching [24] and link architecture [16] to the system.

Gabor wavelet transform can define an input image by allowing the description of both the spatial relations and spatial frequency structure. Convolving an input image with Gabor filters using 5 scales and 8 orientations captures the whole frequency spectrum, and the response is always complex. In Figure 2.13, the magnitude and phase responses of the Gabor filter on an input image are shown.



(a)

(b)

(c)

Figure 2.13. (a) The original image, (b) the magnitude and (c) the phase of the Gabor kernels at one scale and eight orientations.

CHAPTER THREE

METHODOLOGY

3.1. Overview

Face recognition is a pattern recognition task that is implemented specially on face images. Face recognition is a challenging and difficult task and has been one of the popular applications of image processing and analysis. Despite its wide spread usage, most of face recognition methods suffer from challenges like rotation, illumination, facial expression, aging and pose. The core task of this thesis work is to investigate the means by which the recognition performance can be enhanced and speeded up. Therefore, image transformation approach is used as a pre-processing stage before the feature extraction stage.

In this thesis, to extract salient features from face image, a feature based algorithm is used. The motivation behind this is the demonstration of the face image in a very compact way. This fact mainly gains attention and importance when we want to make the system as accurate as possible. Feature based techniques are based on detecting distinctive keypoints on a face image and defining its feature vector in a well-organized way. However, using these algorithms alone does not result in a good recognition performance, so choosing suitable approach is extremely critical to increase the performance and recognition rate of the system. So we perform transformations on images before extracting features from them.

3.2. The Proposed Approach

The proposed approach contains details of the stages taken in carrying out the simulations. All the images are transformed using DWT or GWT. We proposed two approaches using SIFT and SURF.

In first approach, SURF or SIFT was used as a feature extraction algorithm, but before extracting features, input face images are transformed using DWT. DWT generates four different subimages namely: approximate, vertical, horizontal and diagonal.

Figure 3.1 shows 1-scale transformation of input face images. Keypoint detection and description are performed on the output subimages using SURF defined as (DWT-SURF). All keypoint features that are extracted from SURF will be stored. Then, each corresponding feature of keypoints will be compared using kNN to get a score (that defines the number of matched keypoints). Then, summation of scores are stored. At last decision will be made based on the highest score, which will define if a subject belongs to a particular class or not.

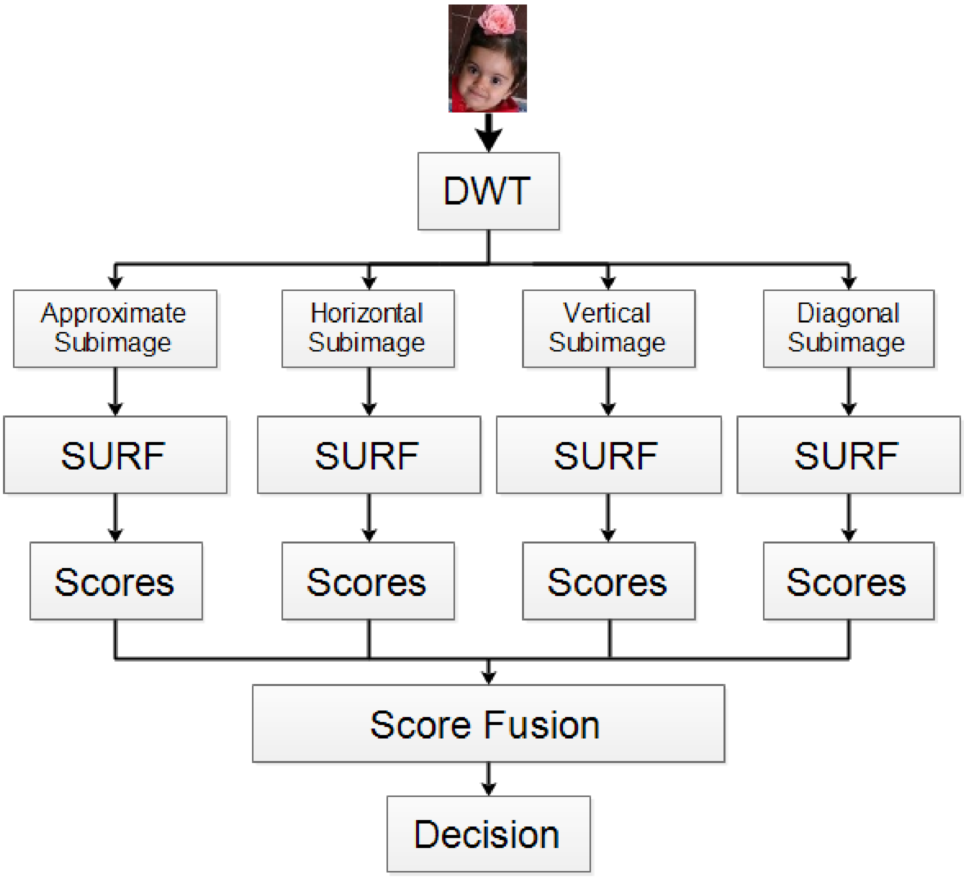


Figure 3.1. The block diagram of proposed approach for DWT-SURF.

In 2-scales transformation, after applying 1-scale transformation, DWT was applied as a second scale on approximate subimage, which produces four subimages. Scores of all eight subimages will be fused and decision will be made based on results. Figure 3.2 describes steps of 2-scales transformation using DWT-SURF.

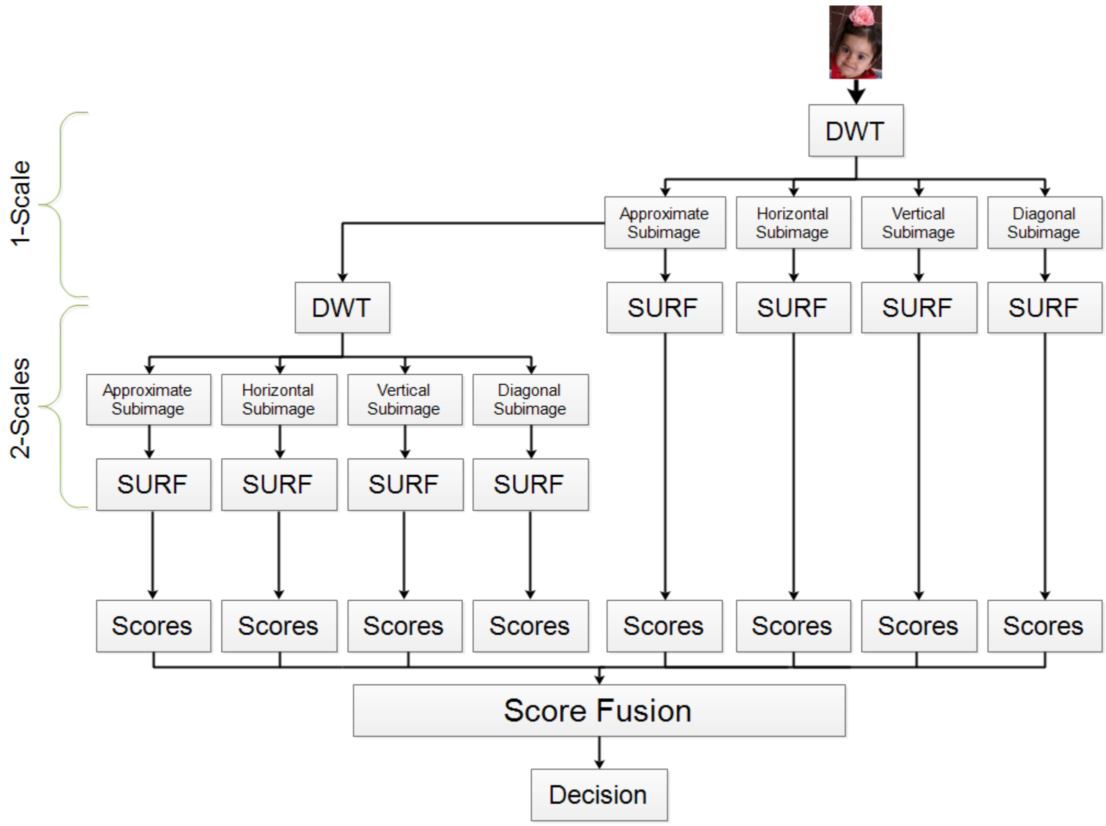


Figure 3.2. The block diagram of 2-scales of DWT-SURF.

The same scenario has been applied but instead of SURF, SIFT was used to extract features from face images. As described in Figures 3.3 and 3.4 shows the same procedure with SIFT algorithm.

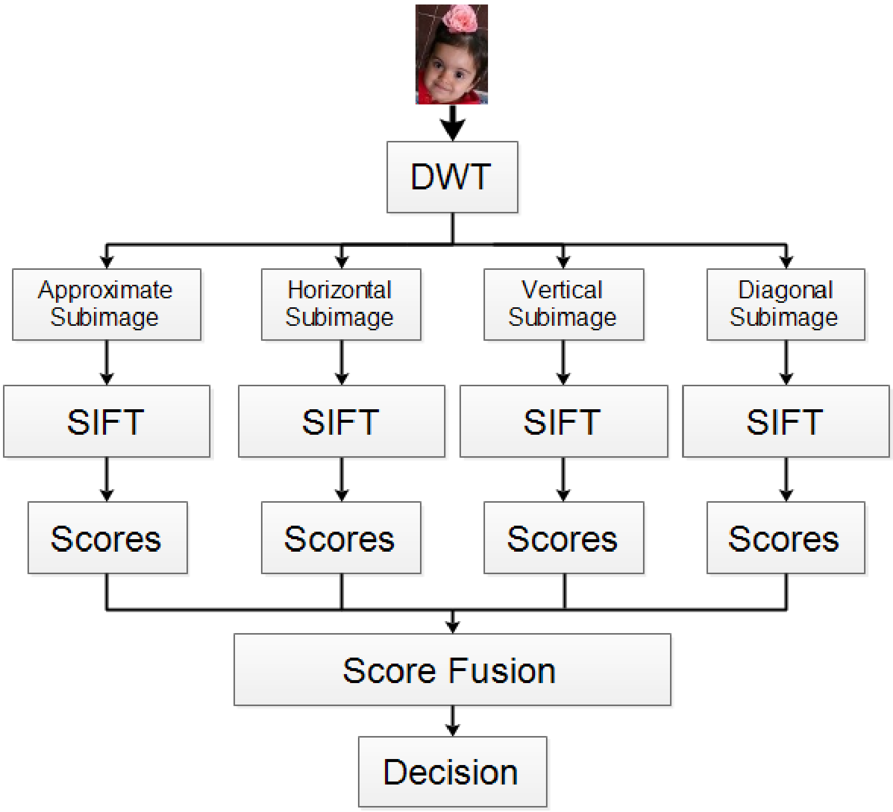


Figure 3.3. The block diagram of 1-scale of DWT-SIFT.

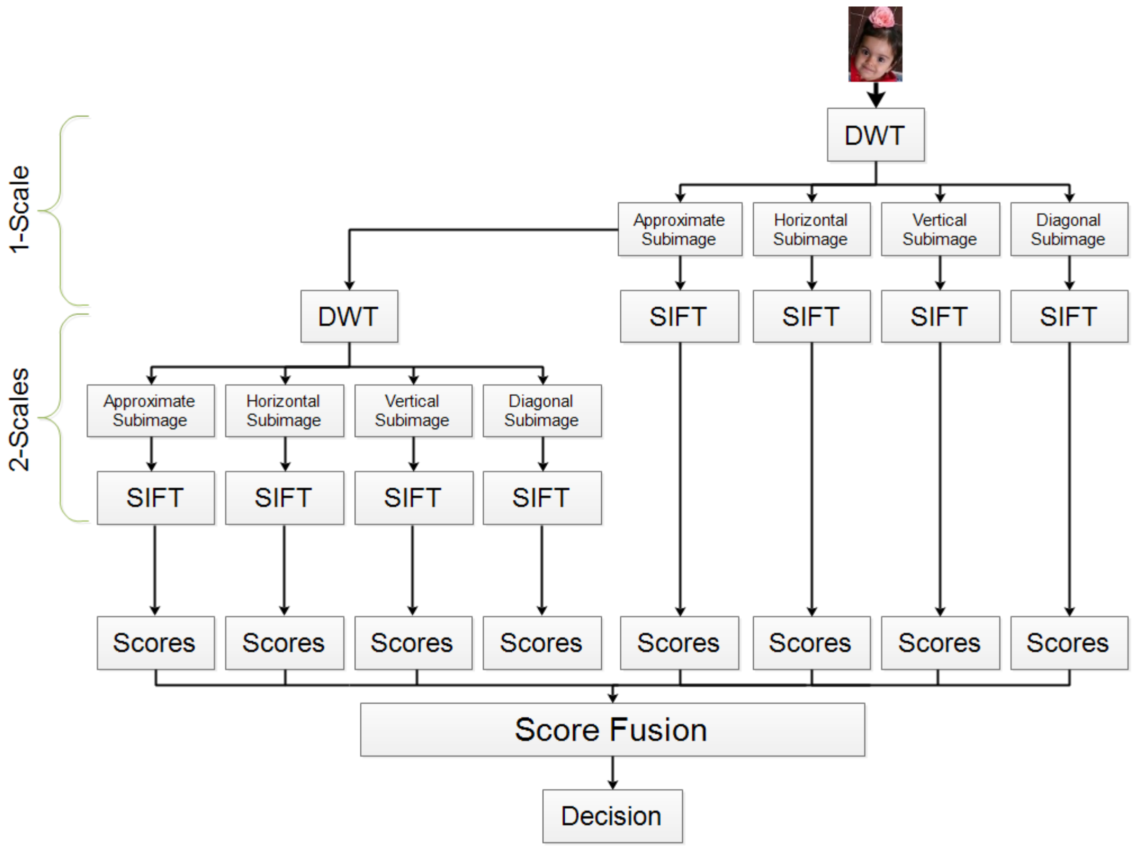


Figure 3.4. The block diagram of 2-scales of DWT-SIFT.

In second approach, SURF or SIFT was used as a feature extraction algorithm, but before extracting features input face images were transformed using GWT. GWT outputs eight different subimages in each scale.

Figure 3.5 shows 1-scale transformation of input images, and features are extracted from output subimages using SURF or SIFT defined as (DWT-SURF, DWT-SIFT). All keypoint features that are extracted from SURF or SIFT will be stored. Then, each corresponding feature of keypoints will be compared using kNN to get a score (that defines the number of matched keypoints). Then, summation of scores are stored. At last decision will be made based on the highest score, which will define if a subject belongs to a particular class or not.

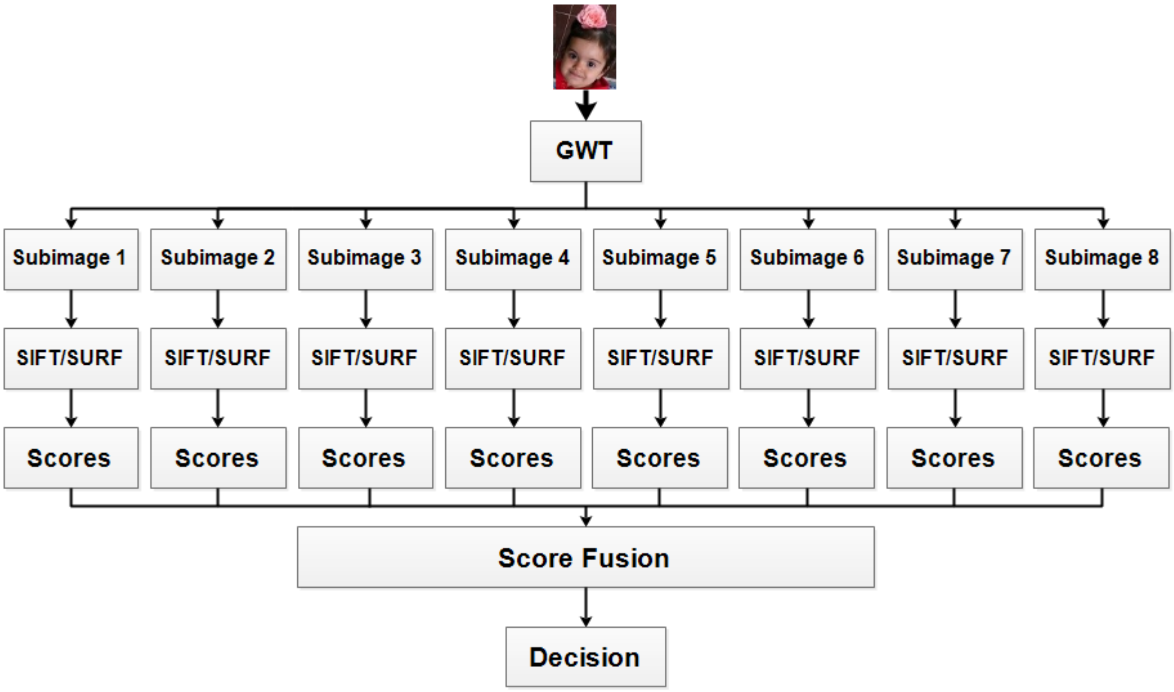


Figure 3.5. The block diagram of 1-scale of GWT-SURF and GWT-SIFT.

Figures 3.6 and 3.7 describe the block diagram of 2-scales transformation using GWT with SIFT and SURF. When input image is given to GWT, it will output eight subimages in complex vectors in each scale which will be a total of 16 subimages. In each subimage features will be extracted using SURF (as shown in Figure 3.7) or SIFT (as shown in Figure 3.6). Scores of each extracted feature will be fused, then based on fused scores a decision will be made.

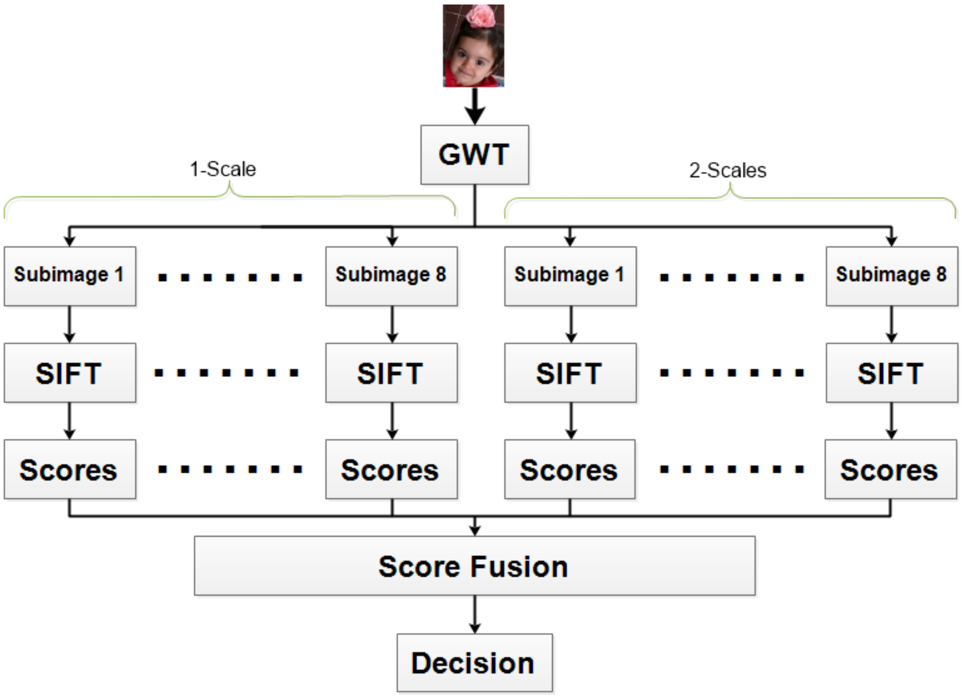


Figure 3.6. The block diagram of 2-scales of GWT-SIFT.

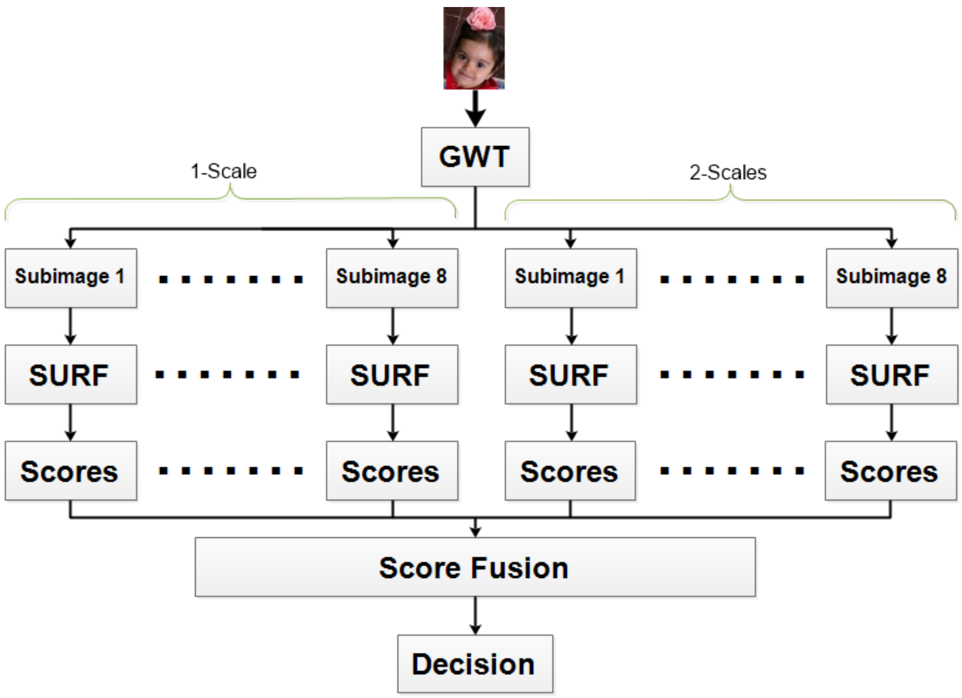


Figure 3.7. The block diagram of 2-scales of GWT-SURF.

CHAPTER FOUR

SIMULATION RESULTS AND DISCUSSIONS

4.1. Overview

In this chapter, the simulation results of the base line recognition performance of the ORL and PUT face database images were presented accordingly. All the simulations were conducted using the MATLAB R2013a software package. A short discussion and observation is drawn in each category of the simulation carried-out. Finally, general discussions of results were presented at the end.

4.2. Simulation Setup

The proposed approach is tested on two different face databases: ORL [26] and PUT [25] face databases. For most of the experiments in each dataset, 5 randomly chosen face images is considered as the gallery (train) set and the remaining face images are considered as the probe (test) set. Subjects in both databases here more than one face image and each subject has different image conditions like (illumination, pose, expression…etc.). Face images that are in gallery set do not exist in the probe set. Each of the images in probe set is matched against the images in the gallery set, and the results and scores are fused and the decision will be made.

Both of the stated databases have different properties to test and asses our proposed approach. ORL face images are taken at different time period and have different head poses. PUT face images were recorded in partly controlled light conditions over a constant background and they have different head pose variations and in some images subjects wear glasses. This provides us an opportunity to compare the recognition performance of our proposed approach with the others using a uniform and standardized databases.

Conventional algorithms and our proposed approach were tested with different number of subjects. We run program 10 times, each time with different randomly chosen subjects. After each experiment, scores will be fused and compared. In all of the experiments, database face images were separated into two classes; gallery (train) and probe (test) set. In most of the experiments, gallery set contains 5 images per subject and the rest is in probe set.

Proposed approaches were applied with various extraction algorithms like (SURF [19], SIFT [18]) and different transform algorithms like (DWT [20], GWT [21] …). With DWT, different types of filters were applied like (db1, db2, db3, db4, db5, haar…). GWT transformation outputs

subimages in complex, SIFT or SURF didn't work properly on real or imaginary parts separately, so magnitude and phase parts of subimages were tested and used separately.
In the following sections, two face databases are explained in detail and performance results of using these databases in our proposed approach are given, with comparisons with some popular conventional face recognition algorithms.

(a) (b)

Figure 4.1. Different face poses of two subjects from PUT face images [25], a) gallery faces, b) probe faces.

Different types of conventional algorithms were used to test images, then results were compared with our proposed approaches using the same face database images.

4.3. Databases Used

4.3.1. The ORL Face Database

The Olivetti Research Laboratory (ORL) [26] face database is tested so as to assess our proposed approach in the existence of head poses and variations in time since images were taken between April 1992 and April 1994. There are 40 different subjects (persons), 10 images per subject, a total of 400 face images. For most of the subjects, the face images were recorded at light variance, time variance, face details (glasses / no glasses), face expressions (open / closed eyes, smiling / not smiling) and head poses (rotation and tilting up to 20°). Most of the face images were recorded against a dark regular background. Figure 4.2 shows the complete set of 40 subjects' face images.

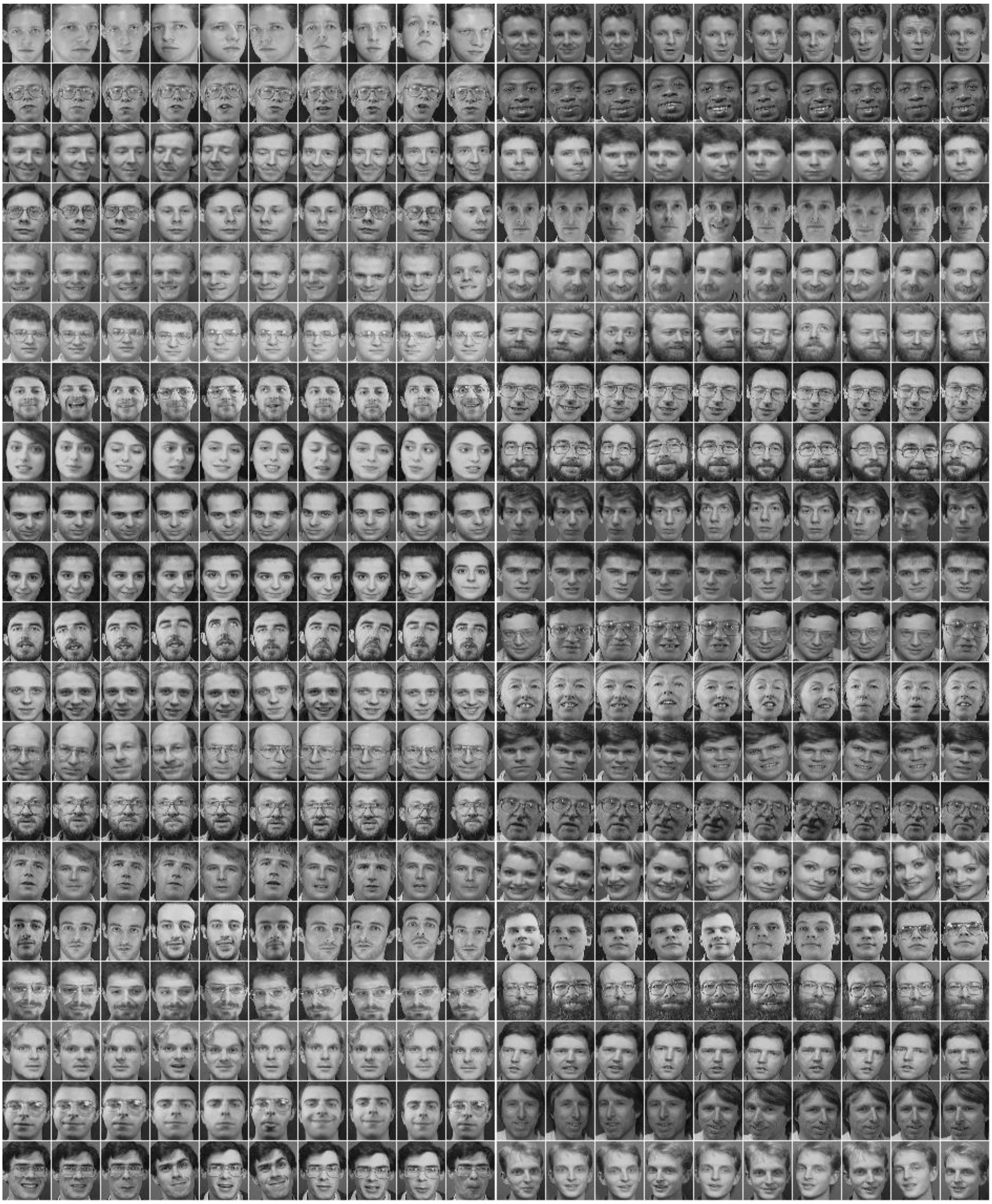


Figure 4.2. Complete set of face images of 40 subjects, 10 images per person.

4.3.2. The PUT Face Database

The Poznan University of Technology (PUT) [25] face database is used so as to test partially controlled illumination conditions over a homogenous background and they have different head pose variations and in some images subjects wear glasses.. There are 100 distinct subjects, 10 images per each subject, a total of 1000 face images. For most of the subjects, the images were recorded with different face expressions, illumination and head poses. The database supplies additional information including: rectangles containing face, nose, eyes, and mouth. Figure 4.3 shows a sample set of images from the PUT database.

Figure 4.3. Sample set of face images for PUT database.

4.4. Results for the ORL Face Database

4.4.1. Results for ORL face database with DWT

- **Results for ORL face database using SURF**

With SURF only, with 5 to 40 subjects were tested, the average recognition rate was 90.09% and the average of correctly classified different number of subjects are shown in Figure 4.4. Performance of all algorithms were decreasing when the number of subjects was increased.

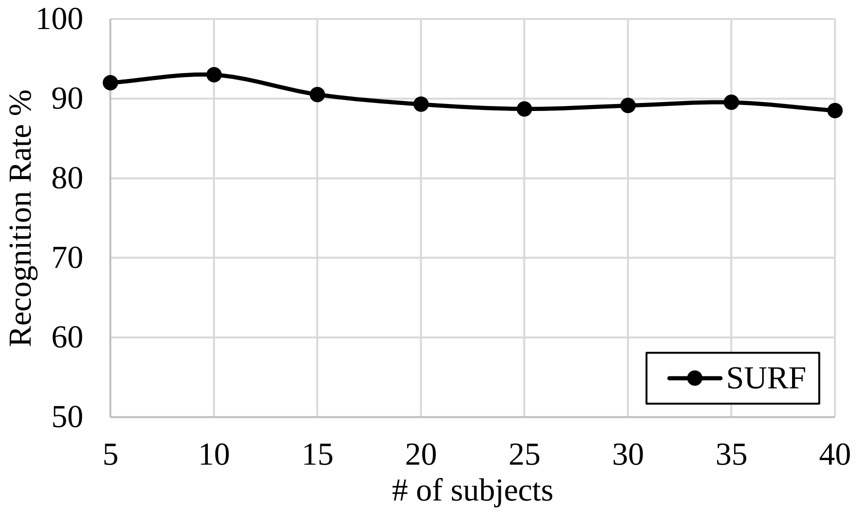


Figure 4.4. Performance of SURF using ORL face database.

At first, 1-scale transformation was applied on face images, using different transformation filters. Results are tabulated in Table 4.1. Performance of our proposed approach was not good because after transformation, SURF couldn’t extract enough features to describe face images. For cA (Approximate), there was 0 SURF points for all images in test and train sets. So we had to use cH (Horizontal), cV (Vertical) and cD (Diagonal) subimages of DWT output.

Table 4.1. Recognition performance of ORL database after applying 1-scale DWT-SURF.

	Wavelet Filters					
# of Subjects	**db1**	**db2**	**db3**	**db4**	**db5**	**haar**
5	82.00	77.60	78.80	78.40	85.20	82.00
10	79.40	81.20	79.80	80.40	82.00	79.40
15	79.20	80.80	79.60	77.87	81.33	79.20
20	79.20	80.40	77.80	74.50	79.60	79.20
25	78.32	77.92	77.76	75.76	78.56	78.32
30	73.00	73.67	75.47	72.20	73.73	73.00
35	72.57	72.51	75.77	72.23	74.57	72.57
40	70.90	69.65	75.05	71.20	73.40	70.90

After applying 1-scale transformation, 2-scales transformation was applied on images with the same filters that we used in 1-scale transformation. Performance of 2-scales transformation is tabulated in Table 4.2.

Table 4.2. Recognition performance of ORL database after applying 2-scales DWT-SURF.

	Wavelet Filters					
# of Subjects	**db1**	**db2**	**db3**	**db4**	**db5**	**haar**
5	82.00	77.60	78.80	78.40	85.20	82.00
10	79.40	81.20	79.80	80.40	82.00	79.40
15	79.20	80.80	79.60	77.87	81.33	79.20
20	79.20	80.40	77.80	74.50	79.60	79.20
25	78.32	77.92	77.76	75.84	78.80	78.32
30	73.00	73.67	75.47	72.53	73.87	73.00
35	72.57	72.46	75.77	72.51	74.69	72.57
40	70.90	69.60	75.05	71.45	73.50	70.90

For DWT algorithm it can be concluded that there is approximately ~1% difference between 1-scale and 2-scales transformation. One can observe the differences from Figure 4.5 which shows the average performance of 1-scale and 2-scales transformations.

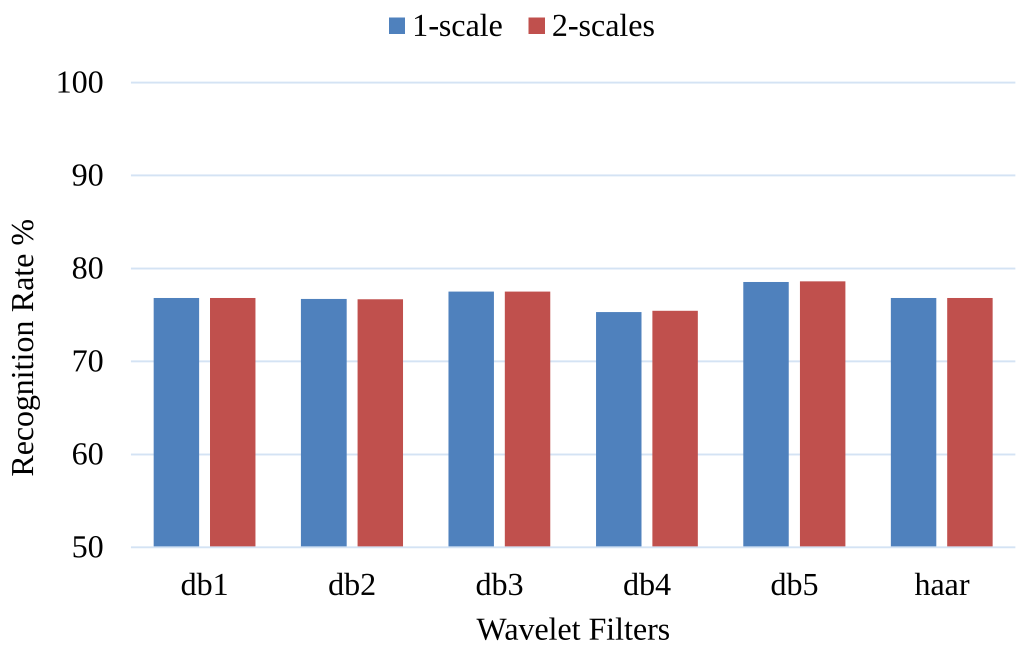


Figure 4.5. Average performance of ORL database after applying 1-scale and 2-scales DWT-SURF.

- **Results for ORL face database using SIFT**

SIFT algorithm performance rate doesn't get affected by size of subjects. As shown in Figure 4.6 for 5 subjects recognition rate is 81.20% while for 40 subjects it is 85.20%. Overall average of SIFT using ORL database is 83.15%.

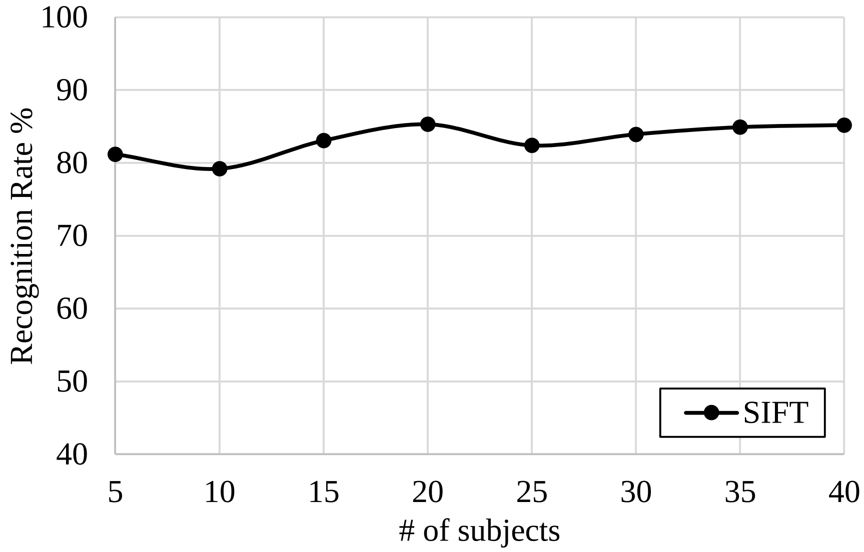


Figure 4.6. Performance of SIFT using ORL face database.

With 1-scale transformation with DWT, the performance of proposed approach is ~7% different form SIFT as listed in Table 4.3 and its better with small number of subjects. The rate of recognition varies with number of subjects, as number of subjects increase recognition rate decreases.

Table 4.3. Recognition performance of ORL database after applying 1-scale DWT-SIFT

	Wavelet Filters					
# of Subjects	**db1**	**db2**	**db3**	**db4**	**db5**	**haar**
5	90.00	84.00	78.00	85.20	87.20	90.00
10	85.20	78.00	80.40	77.80	80.60	85.20
15	84.00	75.87	78.40	69.60	78.13	84.00
20	79.90	74.70	73.60	70.70	78.00	79.90
25	77.20	73.28	74.48	72.64	75.92	77.20
30	77.47	72.07	74.20	74.07	77.47	77.47
35	79.49	73.83	75.31	74.74	78.46	79.49
40	77.80	71.70	73.85	73.55	77.65	77.80

In 2-scales transformations using DWT, the performance is ~1% different from 1-scale transformation as tabulated in Table 4.4. Figure 4.7 clarify average performance difference between them.

Table 4.4. Recognition performance of ORL database after applying 2-scales DWT-SIFT.

	Wavelet Filters					
# of Subjects	**db1**	**db2**	**db3**	**db4**	**db5**	**haar**
5	88.40	83.60	78.80	87.20	85.60	88.40
10	86.60	77.80	80.80	82.80	82.20	86.60
15	84.13	76.00	79.07	75.33	80.53	84.13
20	78.60	74.00	73.40	73.50	80.30	78.60
25	77.20	73.52	74.08	75.12	79.20	77.20
30	76.80	72.27	75.13	75.60	80.20	76.80
35	78.80	73.31	76.11	76.17	80.86	78.80
40	77.85	71.65	74.70	75.50	79.80	77.85

As a conclusion, there is approximately ~2% difference between 1-scale and 2-scales transformation on ORL face images using SIFT.

With different number of subjects, GWT-SURF outperforms DWT-SURF and SURF. The difference between GWT-SURF and SURF is approximately ~4%.

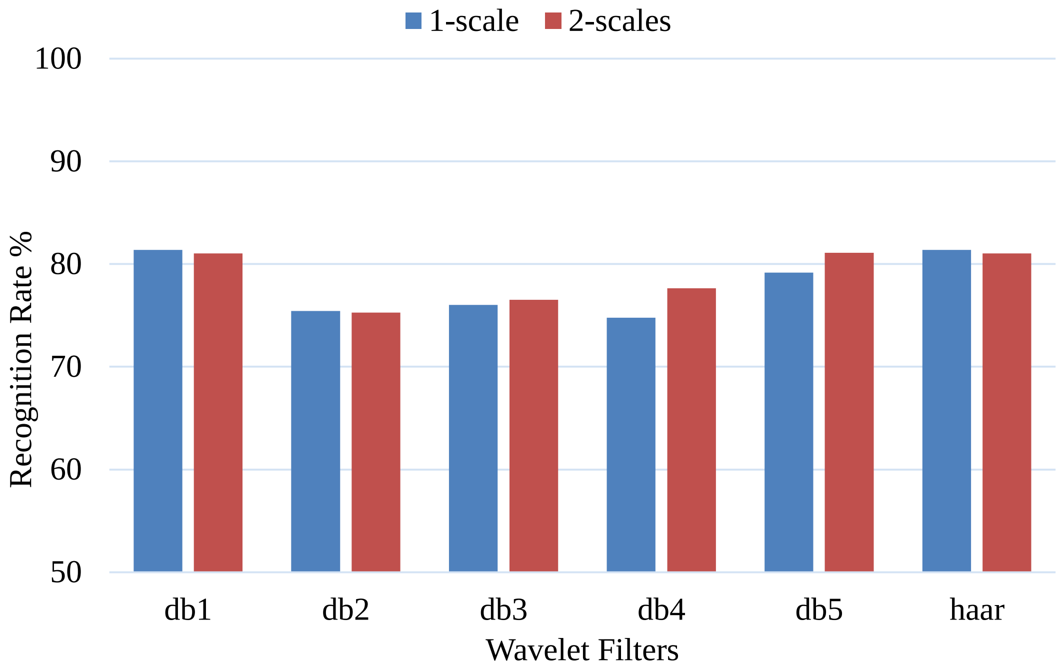


Figure 4.7. Average performance of ORL database after applying 1-scale and 2-scales DWT-SIFT.

4.4.2. Results for ORL face database with GWT

- **Results for ORL face database using SURF**

The performance of our proposed approach was completely different when GWT was applied on face images before extracting features from it. At first, with 1-scale transformation, GWT outputs subimages in complex and SURF doesn't work properly with complex, so our proposed approaches performed well using Magnitude, Phase and both. Performance of our proposed approach listed in Table 4.5.

Table 4.5. Recognition performance of ORL database after applying 1-scale GWT-SURF

# of Subjects	Magnitude	Phase	Magnitude + Phase
5	91.20	92.00	93.60
10	90.20	92.60	94.20
15	91.73	93.60	95.20
20	90.60	92.50	93.70
25	91.28	92.88	94.16
30	88.73	92.33	93.33
35	89.31	92.00	93.60
40	89.70	91.60	93.40

The performance of proposed approach using Magnitude and Phase of transformed images is ~5% higher than the conventional SURF algorithm. The performance of recognition of our proposed approach decreases less compared to SURF itself with increasing number of subjects.

- **Results for ORL face database using SIFT**

The performance of proposed approach was not as expected with DWT, but with GWT it differs around ~12% which is much better than conventional SIFT only, as listed in Table 4.6. Overall average of SIFT is 83.15% while recognition rate of SIFT after applying GWT as a transformation was 95.06%.

Table 4.6. Recognition performance of ORL database after applying 1-scale GWT-SIFT.

# of Subjects	Magnitude	Phase	Magnitude + Phase
5	91.20	20.80	96.40
10	90.20	14.60	96.00
15	91.73	6.93	96.27
20	90.60	5.80	94.20
25	91.28	4.72	94.08
30	88.73	4.40	94.27
35	89.31	3.89	94.63
40	89.70	3.30	94.60

It was observed that Phase of complex GWT subimages didn't work well with SIFT while Magnitude and combination of Magnitude and Phase are quite more accurate compared to SIFT only.

The overall recognition performance rate for different number of subjects for SURF, SIFT, and proposed approaches are shown in Figure 4.8 and Figure 4.9.

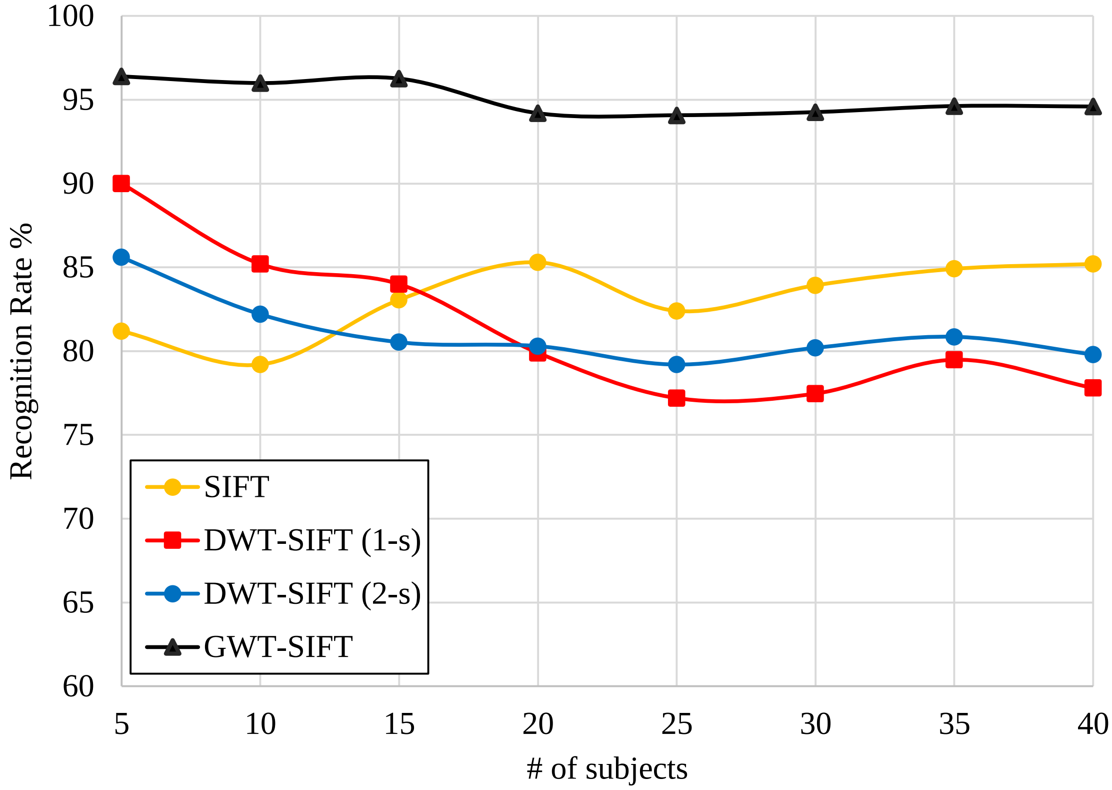


Figure 4.8. Overall recognition performance of SIFT, DWT-SIFT (1-scale), DWT-SIFT (2-scales), and GWT-SIFT on ORL database.

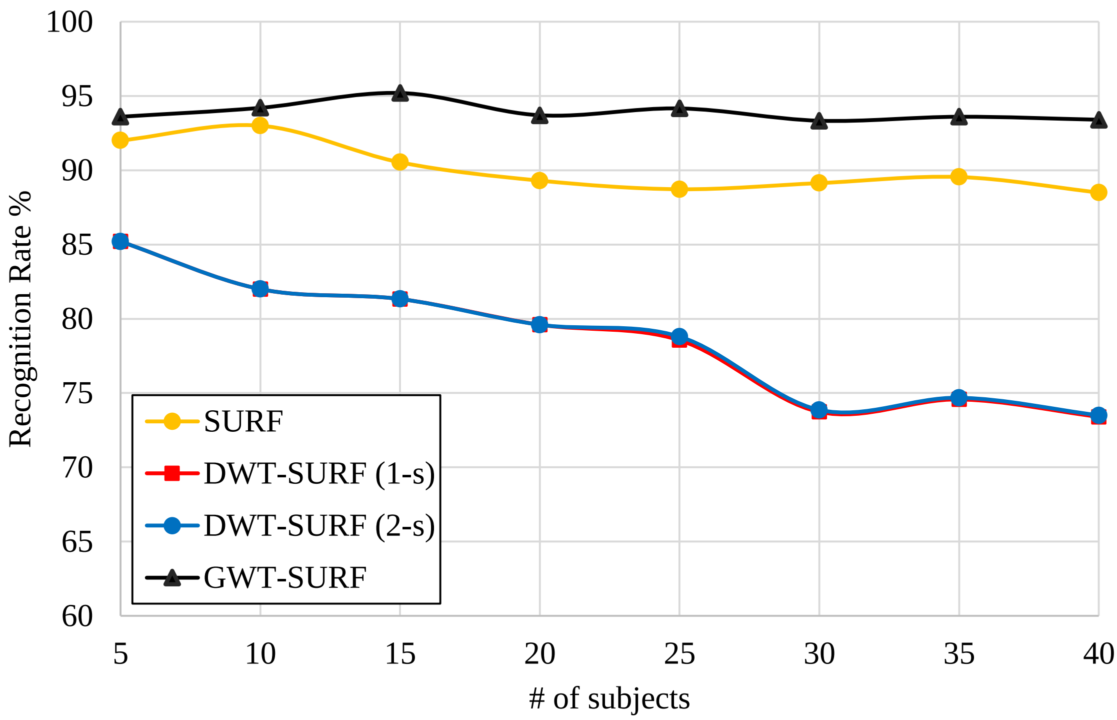


Figure 4.9. Overall recognition performance of SURF, DWT- SURF (1-scale), DWT- SURF (2-scales), and GWT- SURF on ORL database.

Adding a preprocessing stage increases performance of our proposed approach but at the same time it increases computation time especially with GWT since there is more subimages than DWT so the time for computing GWT subimages are higher than DWT subimages. Number of keypoints detected using DWT or GWT are higher than SIFT or SURF only. Table 4.7 shows the computation time (in seconds) and average number of keypoints detected for 400 images of ORL face database.

Table 4.7. Computation time and number of keypoints detected for SIFT, SURF, and proposed approach on ORL database.

	Time (s)	**# of Keypoints**
SIFT	3.03	36
DWT-SIFT	3.75	38
GWT-SIFT	47.66	45
SURF	1.99	16
DWT-SURF	5.81	43
GWT-SURF	49.76	340

4.5. Results for the PUT Face Database

4.5.1. Results for PUT face database with DWT

- **Results for PUT face database using SURF**

With SURF only, with 5 to 100 subjects were tested, the average recognition rate was 84.84%. The average of correctly classified different number of subjects are shown in Figure 4.10. Performance of SURF decreases when number of subjects are increased.

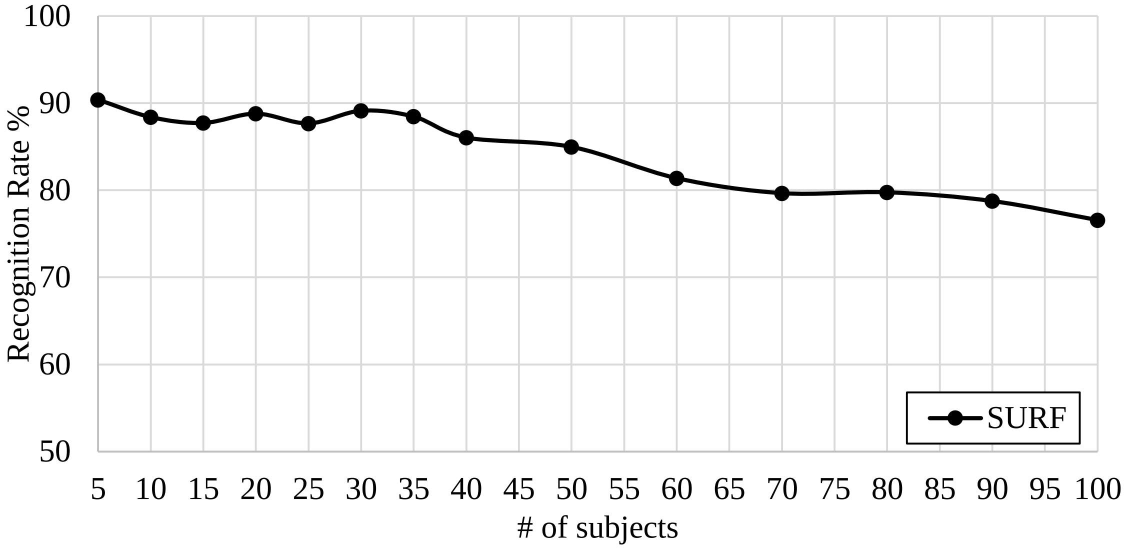


Figure 4.10. Performance of SURF using PUT face database.

The same steps that were performed on ORL database, was performed on PUT database. At first 1-scale transformation was applied on images, using different transformation filters. Results are tabulated in Table 4.8. Performance of the algorithm was not good because after transformation, SURF couldn't extract enough features to describe face images. For cA (Approximate), there was 0 SURF points for all images in gallery and probe sets. So we had to use cH (Horizontal), cV (Vertical) and cD (Diagonal) subimages of DWT.

Table 4.8. Recognition performance of PUT database after applying 1-scale DWT-SURF.

	Wavelet Filters					
# of Subjects	**db1**	**db2**	**db3**	**db4**	**db5**	**haar**
5	96.00	94.00	96.00	97.60	96.40	96.00
10	94.00	92.40	94.00	92.20	94.20	94.00
15	92.67	88.67	90.40	90.00	92.93	92.67
20	91.70	89.60	88.30	88.50	92.40	91.70
25	90.40	89.76	86.88	87.44	92.16	90.40
30	89.60	89.47	85.33	85.87	90.07	89.60
35	87.60	87.77	84.91	85.14	88.86	87.60
40	87.65	86.25	84.40	83.50	87.80	87.65
50	86.27	84.22	82.53	82.36	86.27	86.27
60	83.02	80.15	76.51	77.93	81.89	83.02
70	80.68	76.95	74.12	76.09	80.58	80.68
80	80.61	76.56	74.56	74.75	78.96	80.61
90	79.88	76.09	73.65	74.19	78.38	79.88
100	78.59	74.82	72.59	71.52	76.93	78.59

2-scales of transformations were applied on images with the same filters that we used in 1-scale transformation. Performance of 2-scale transformation is tabulated in Table 4.9.

Table 4.9. Recognition performance of PUT database after applying 2-scales DWT-SURF

	Wavelet Filters					
# of Subjects	**db1**	**db2**	**db3**	**db4**	**db5**	**haar**
5	96.40	94.00	96.00	97.60	96.40	96.40
10	94.20	92.20	94.00	92.20	94.40	94.20
15	92.80	88.53	90.67	90.13	92.93	92.80
20	91.70	89.70	88.60	88.80	92.30	91.70
25	90.48	89.84	86.96	87.60	92.40	90.48
30	89.60	89.60	85.40	86.40	90.40	89.60
35	87.60	87.83	84.69	85.71	88.97	87.60
40	87.50	86.35	84.00	84.30	88.05	87.50
50	86.18	84.27	82.36	83.29	86.49	86.18
60	82.91	80.18	76.40	78.80	82.29	82.91
70	80.62	76.95	74.37	76.98	80.92	80.62
80	80.64	76.72	74.88	75.55	79.33	80.64
90	79.98	76.38	73.98	74.80	78.82	79.98
100	78.67	75.05	72.86	72.42	77.45	78.67

For DWT algorithm we can conclude that there is approximately ~1% difference between 1 scale and 2-scales transformations. One can observe the average performance differences from Figure 4.11 which is quite close to each other.

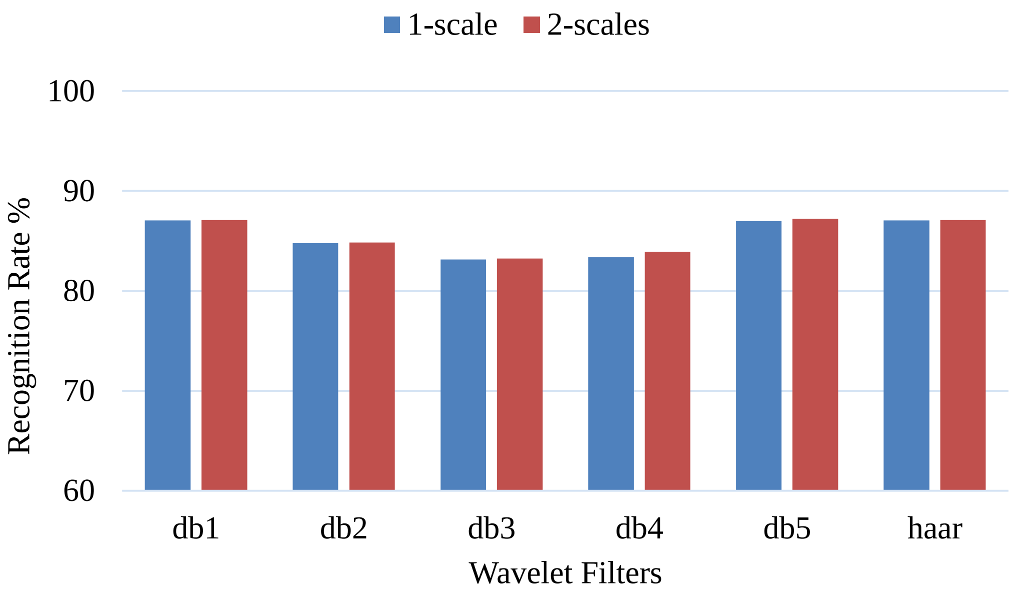


Figure 4.11. Average performance of PUT database after applying 1-scale and 2-scales DWT-SURF.

- **Results for PUT face database using SIFT**

SIFT algorithm performance rate changes when size of subject's decreases. As shown in Figure 4.12, for 5 subjects recognition rate is 98.80% while for 100 subjects it is 93.90%. Overall average of SIFT on PUT database is 96.12%.

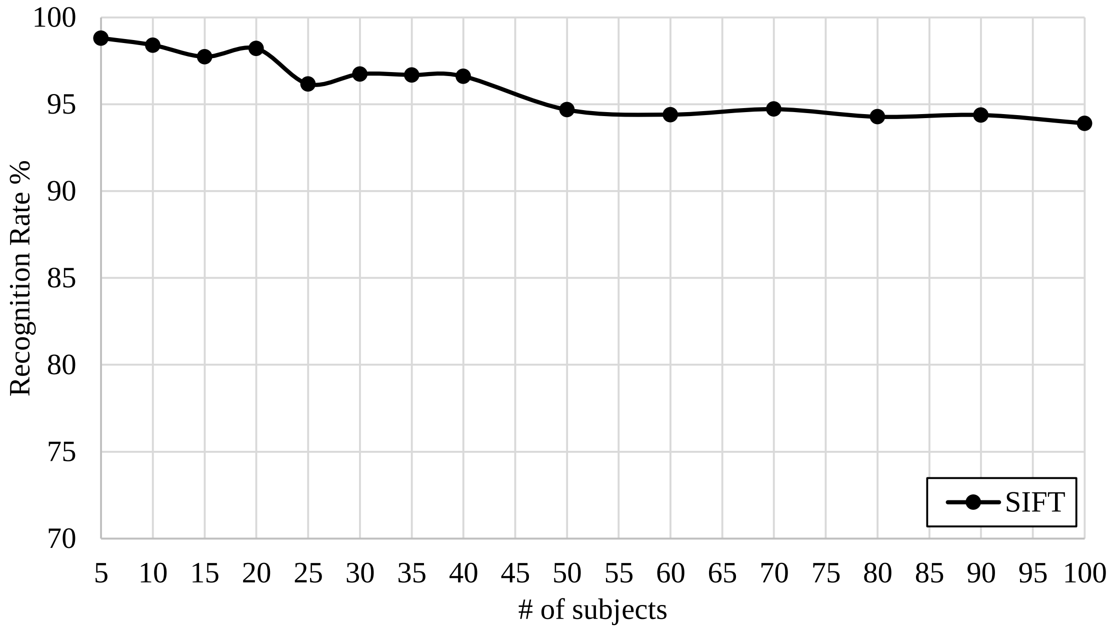


Figure 4.12. Performance of SIFT using PUT face database.

The same steps were performed that was performed before on ORL database. At first 1-scale transformation was applied on face images, using different transformation filters. Results are tabulated in Table 4.10. Performance of our proposed approach was not good because after transformation, SURF couldn't extract enough features to describe face images. For cA (Approximate), there was 0 SURF points for all images in test and train sets. So we had to use cH (Horizontal), cV (Vertical) and cD (Diagonal) subimages of DWT.

Table 4.10. Recognition performance of PUT database after applying 1-scale DWT-SIFT.

	Wavelet Filters					
# of Subjects	**db1**	**db2**	**db3**	**db4**	**db5**	**haar**
5	97.60	89.60	89.20	90.80	92.80	97.60
10	91.00	80.80	82.00	87.60	86.00	91.00
15	91.87	82.80	87.07	90.80	89.33	91.87
20	92.20	83.80	87.40	90.20	90.20	92.20
25	92.16	84.48	88.88	91.20	91.20	92.16
30	91.80	86.60	89.47	92.07	91.60	91.80
35	91.09	86.29	89.77	91.89	91.37	91.09
40	88.95	84.05	88.35	89.30	89.45	88.95
50	87.07	82.22	86.71	88.62	88.98	87.07
60	85.09	76.80	84.58	84.73	86.84	85.09
70	83.88	77.60	84.86	84.92	87.75	83.88
80	84.48	78.69	85.09	84.80	87.65	84.48
90	83.08	76.49	82.87	84.05	86.00	83.08
100	82.88	75.92	82.61	83.81	85.77	82.88

2-scales transformation was applied on face images using the same filters that was used in 1-scale transformation. Performance of 2-scales transformation is tabulated in Table 4.11.

Table 4.11. Recognition performance of PUT database after applying 2-scales DWT-SIFT.

	Wavelet Filters					
# of Subjects	**db1**	**db2**	**db3**	**db4**	**db5**	**haar**
5	98.40	93.60	91.20	94.40	94.40	98.40
10	91.80	83.40	87.60	91.00	90.60	91.80
15	94.13	86.67	90.80	93.07	93.07	94.13
20	94.70	86.60	90.40	92.70	93.80	94.70
25	94.72	86.96	91.20	93.20	94.24	94.72
30	94.40	88.53	91.47	93.47	94.53	94.40
35	93.54	88.57	91.77	92.51	94.17	93.54
40	91.65	86.45	90.30	89.75	93.25	91.65
50	89.42	85.24	88.49	89.16	92.93	89.42
60	87.31	81.49	87.13	87.56	90.33	87.31
70	86.00	82.00	86.74	88.22	90.83	86.00
80	85.68	82.91	86.80	88.00	90.51	85.68
90	84.87	80.59	84.94	87.18	88.94	84.87
100	84.80	80.51	84.15	86.84	88.53	84.80

For DWT algorithm it can be concluded that there is approximately ~2% difference between 1-scale and 2-scales transformation. One can observe the average performance differences from Figures 4.13.

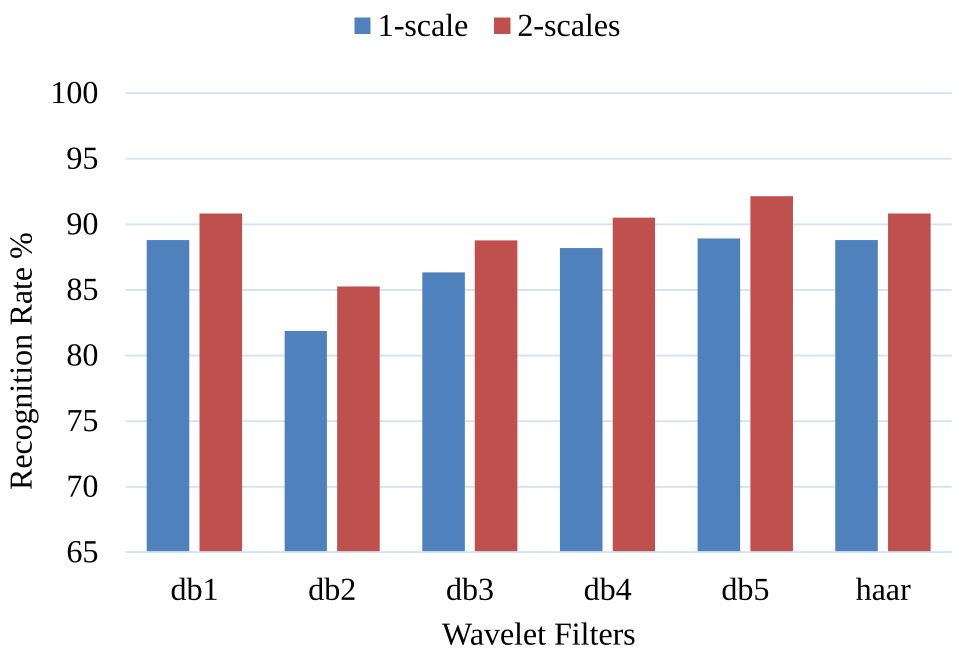


Figure 4.13. Average performance of PUT database after applying 1-scale and 2-scales DWT-SIFT.

4.5.2. Results for PUT face database with GWT

- **Results for PUT face database using SURF**

The performance of our proposed approach was completely different when GWT was applied on face images before extracting features from it. At first 1-scale transformation was used. GWT outputs vectors in complex and SURF doesn't work properly with complex, so Magnitude, Phase and both of them were used in the proposed approach. Performance of proposed approach is listed in Table 4.12.

Table 4.12. Recognition performance of PUT database after applying 1-scale GWT-SURF.

# of subjects	Magnitude	Phase	Magnitude + Phase
5	99.60	97.60	97.60
10	98.80	98.20	98.40
15	99.20	97.60	98.67
20	99.30	98.10	99.00
25	99.44	98.32	99.12
30	99.53	98.60	99.27
35	99.03	98.74	99.37
40	98.90	98.60	99.20
50	99.02	98.76	99.29
60	99.16	98.62	99.13
70	99.23	98.77	99.26
80	99.20	98.83	99.36
90	99.15	98.73	99.29
100	98.82	98.63	99.18

The performance of proposed approach using Phase of transformed images is ~3% higher than conventional SURF algorithm. The rate of recognition performance of proposed approach decreases slowly compared to SURF itself.

- **Results for PUT face database using SIFT**

The performance of proposed approach was not as expected using DWT, but using GWT it differs around ~13% which is much better than conventional SIFT only, as listed in Table 4.13. Overall average recognition rate of SIFT is 83.15%. On the other hand recognition rate of SIFT after applying GWT as a transformation is 95.06%.

We observe that Phase of complex GWT algorithm doesn’t work properly with SIFT while magnitude and combination of Magnitude and Phase were quite more accurate compared to SIFT itself.

Table 4.13. Recognition performance of PUT database after applying 1-scale GWT-SIFT.

# of subjects	Magnitude	Phase	Magnitude + Phase
5	99.60	24.80	100.00
10	99.80	17.00	100.00
15	98.93	15.33	99.07
20	99.10	12.60	99.00
25	99.20	10.64	99.12
30	99.27	9.00	99.20
35	99.31	8.23	99.26
40	99.25	7.80	99.15
50	99.24	6.93	99.16
60	99.27	6.15	99.16
70	99.35	5.32	99.29
80	99.39	4.69	99.36
90	99.22	4.05	99.25
100	99.07	3.81	99.09

The overall recognition performance rate for different number of subjects for SURF, SIFT, and proposed approaches are shown in Figure 4.14 and Figure 4.15.

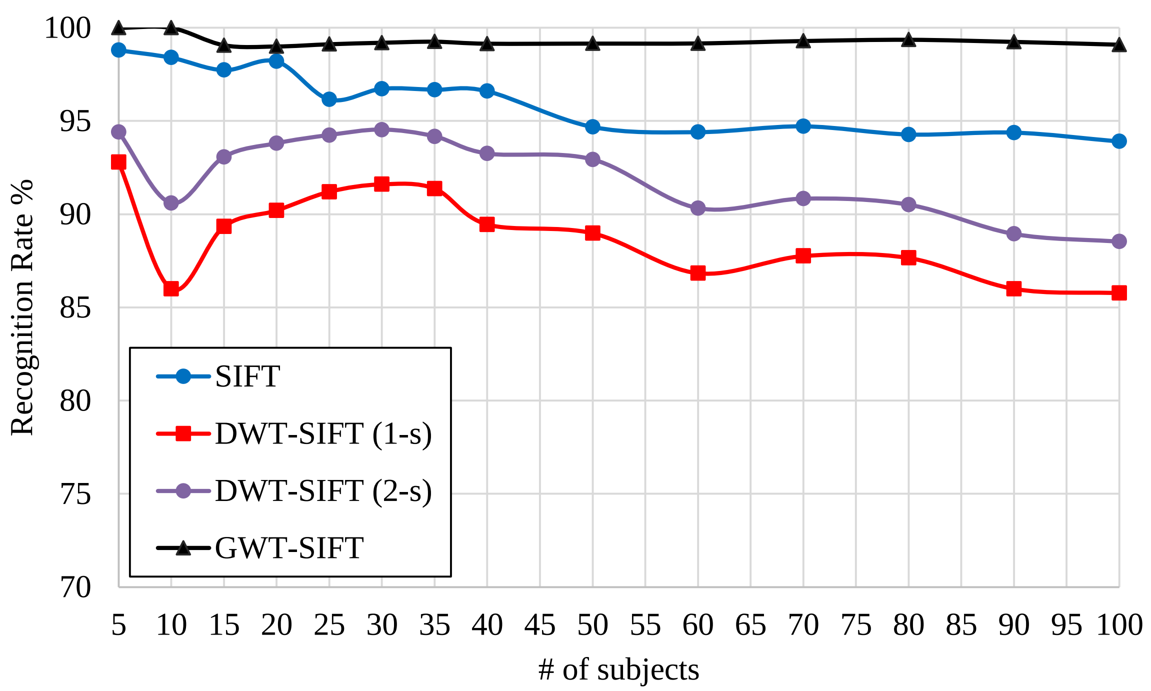


Figure 4.14. Overall recognition performance of SIFT, DWT-SIFT (1-scale), DWT-SIFT (2-scales), and GWT-SIFT on PUT database.

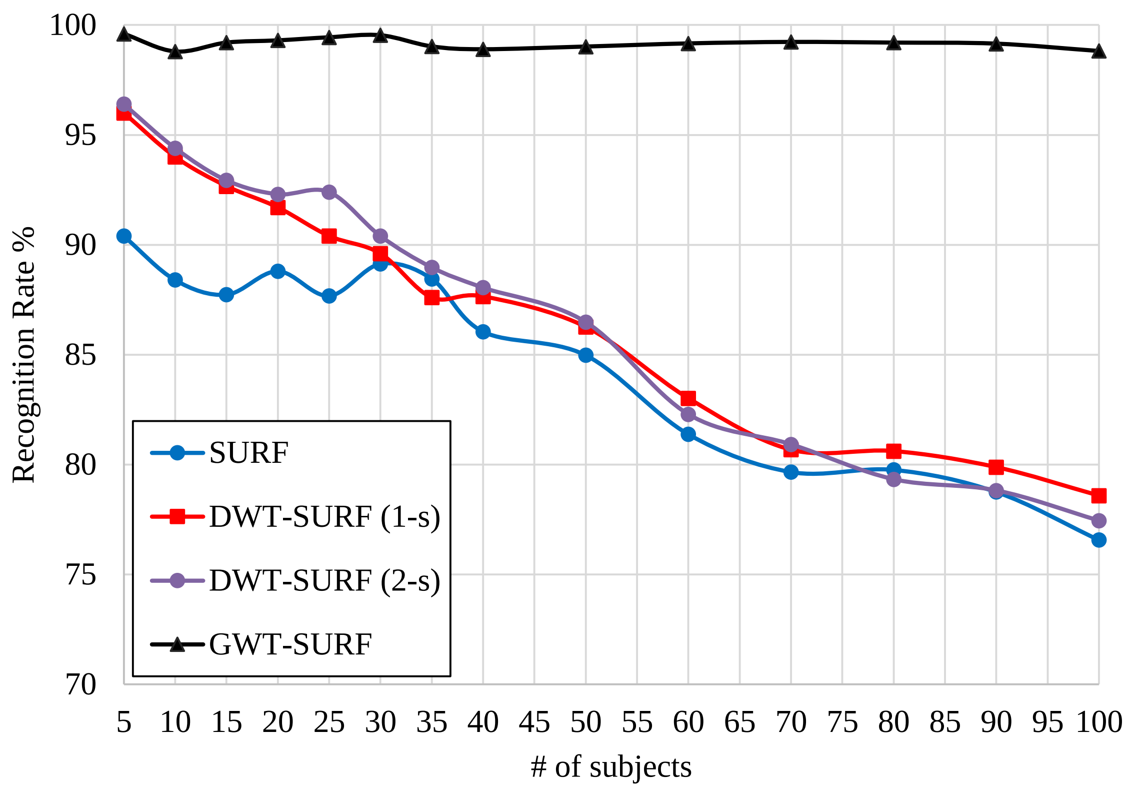


Figure 4.15. Overall recognition performance of SURF, DWT- SURF (1-scale), DWT- SURF (2-scales), and GWT- SURF on PUT database.

Size of face images affects overall performance face recognition system and computation time. Since the size of PUT face images are bigger than ORL face images so performance of our proposed approach is higher and at the same time computation time is also higher. Especially, with GWT since there is more subimages than DWT so the time for computing GWT subimages are higher than DWT subimages. Table 4.14 shows the computation time and average number of keypoints detected for 1000 images of PUT face database.

Table 4.14. Computation time and number of keypoints detected for SIFT, SURF, and proposed approach on PUT database.

	Time (s)	**# of Keypoints**
SIFT	10.67	45
DWT-SIFT	12.69	54
GWT-SIFT	259.08	75
SURF	5.21	11
DWT-SURF	15.55	87
GWT-SURF	254.78	490

4.6. Results for Different Sizes of Gallery Set

In the following experiments, different number of images per subject from gallery and probe set are used. The number of subjects vary from 1 to 9 (1,2,3,4,5,6,7,8,9). For example, 1 subject from gallery and the rest, which is 9, will be in probe set. Our proposed approaches were applied and tested on both of the ORL and PUT databases.

The results of the experiments are described briefly in the following sections.

4.6.1. Results for PUT face database

In general, with PUT database face images experiments show that, the performance of our proposed approaches DWT-SURF, GWT-SURF and GWT-SIFT are higher and better compared to those of SURF and SIFT.

- **Results using SIFT**

  In this experiment we used SIFT as a feature extraction algorithms and the results show that, GWT-SIFT outperforms DWT-SIFT and SIFT. The difference between GWT-SURF and SIFT is approximately ~6%. The difference of recognition performances are shown in Figure 4.16.

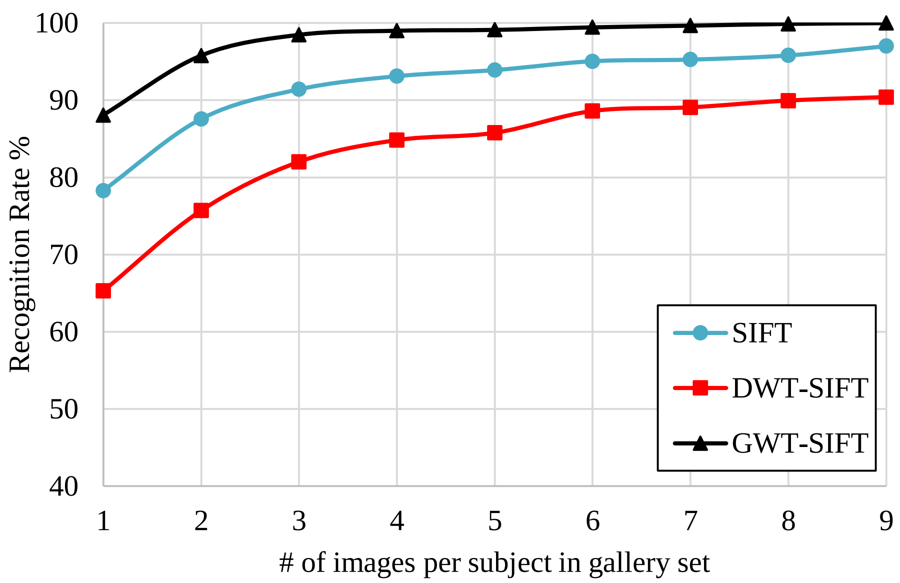


Figure 4.16. The performance of SIFT, DWT-SIFT and GWT-SIFT using PUT with different number of images per subjects.

- **Results using SURF**

  In this experiment SURF was used as a feature extraction algorithm and the results show that, GWT-SURF and DWT-SURF outperforms SURF. The difference between GWT-SURF and SURF is approximately ~22% and difference of DWT-SURF compared to SURF is approximately ~2%. The detailed differences of recognition performances are shown in Figure 4.17.

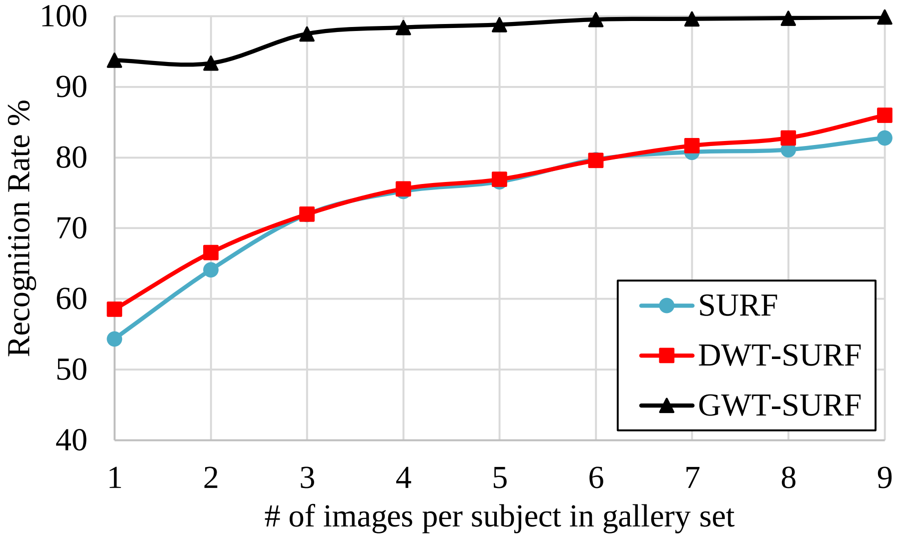


Figure 4.17. The performance of SURF, DWT-SURF, and GWT-SURF using PUT with different number of subjects.

4.6.2. Results for ORL face database

In general with ORL database, the performances of GWT-SURF and GWT-SIFT are higher and better compared to those of DWT-SURF, DWT-SIFT, SURF, and SIFT.

- **Results using SIFT**

  In this experiment SIFT was used as a feature extraction algorithm and the results show that, GWT-SIFT outperforms DWT-SIFT and SIFT. The difference between GWT-SURF and SIFT is approximately ~8%. The difference of recognition performances are shown in Figure 4.18.

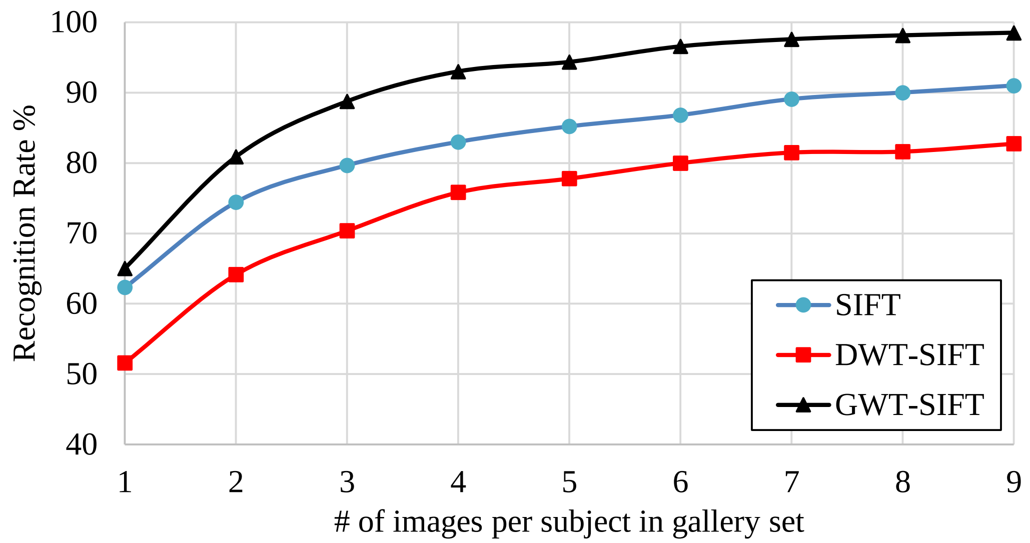


Figure 4.18. The performance of SIFT, DWT-SIFT and GWT-SIFT using ORL with different number of subjects.

- **Results using SURF**

  In this experiment SURF was used as a feature extraction algorithm and the results show that, GWT-SURF outperforms DWT-SURF and SURF. The difference between GWT-SURF and SURF is approximately ~4%. The differences of recognition performances are shown in Figure 4.19.

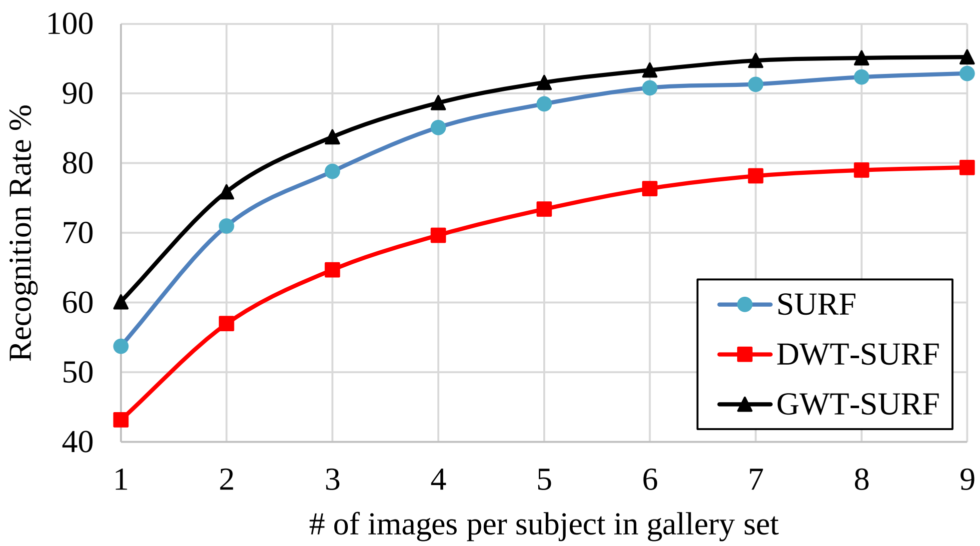


Figure 4.19. The performance of SURF, DWT-SURF and GWT-SURF using ORL with different number of subjects.

4.7. The Proposed Approach Comparison with Other Methods

The proposed approach was compared with conventional PCA. It can be observed that, generally our proposed approach shows superiority in performance over PCA [27] as described in the following experiments.

4.7.1. Results for PUT face database

In the first experiment, SURF, DWT-SURF and GWT-SURF are compared with PCA and the results show that performance of GWT-SURF and DWT-SURF are higher than that of PCA, as shown in Figure 4.20.

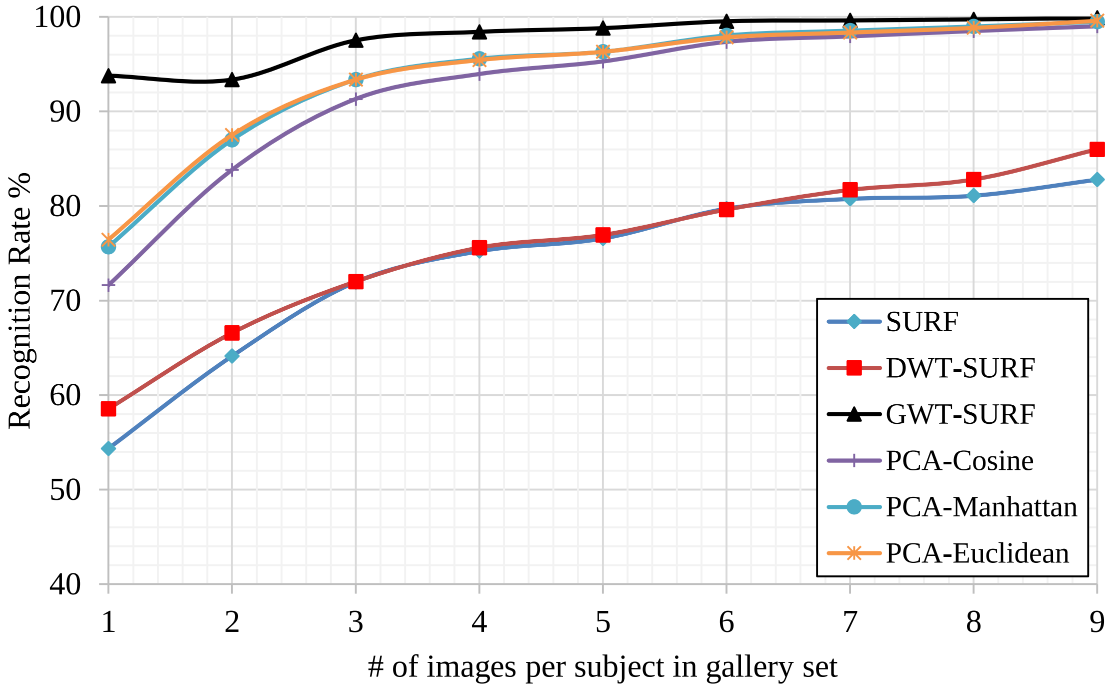


Figure 4.20. Recognition performance of SURF, DWT-SURF, GWT-SURF and PCA.

In the second experiment, SIFT, DWT-SIFT and GWT-SIFT are compared with PCA and the results show that performance of GWT-SIFT is the best while DWT-SIFT is the worst, as shown in Figure 4.21.

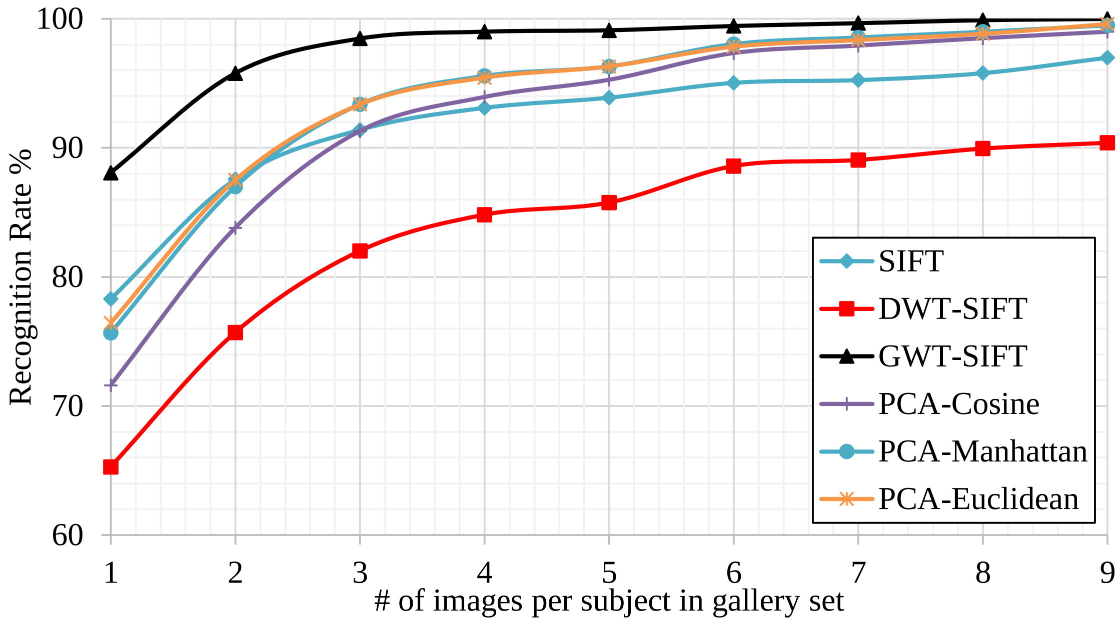


Figure 4.21. Recognition performance of SIFT, DWT-SIFT, GWT-SIFT and PCA.

4.7.2. Results for ORL face database

In the first experiment, SURF, DWT-SURF and GWT-SURF are compared with PCA and the results show that performances of GWT-SURF and DWT-SURF are lower than that of PCA, as shown in Figure 4.22, on ORL face database.

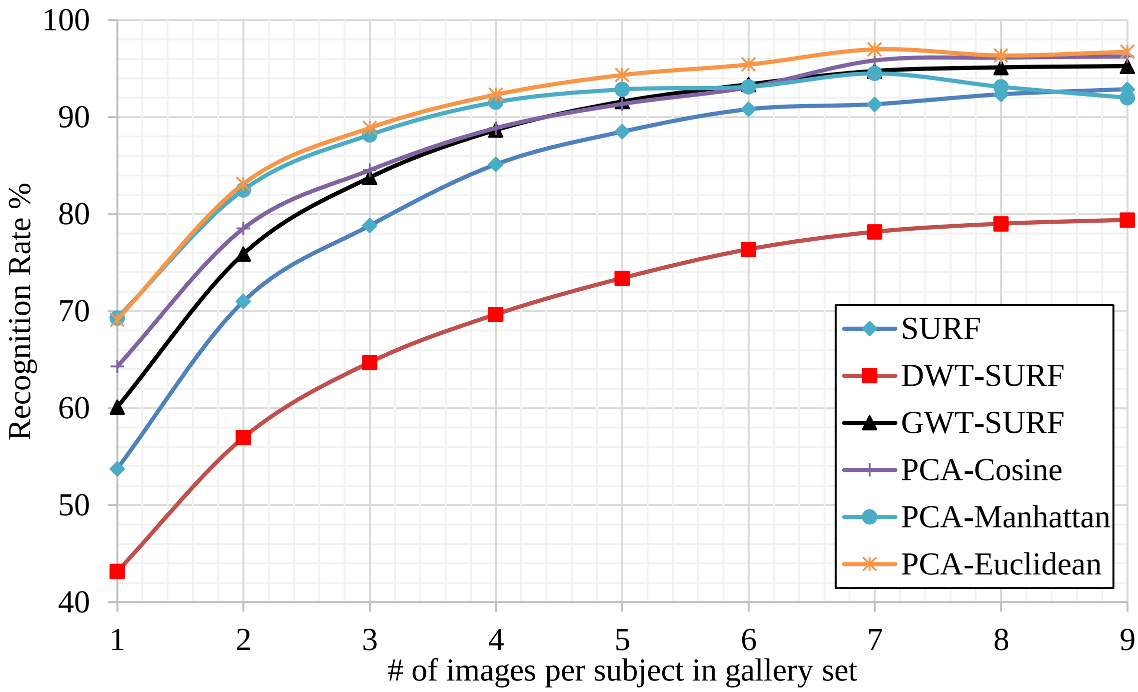


Figure 4.22. Recognition performance of SURF, DWT-SURF, GWT-SURF and PCA.

In the second experiment, SIFT, DWT-SIFT and GWT-SIFT are compared with PCA and the results show that performance of GWT-SIFT is almost highest while DWT-SIFT given the worst performance, as shown in Figure 4.23.

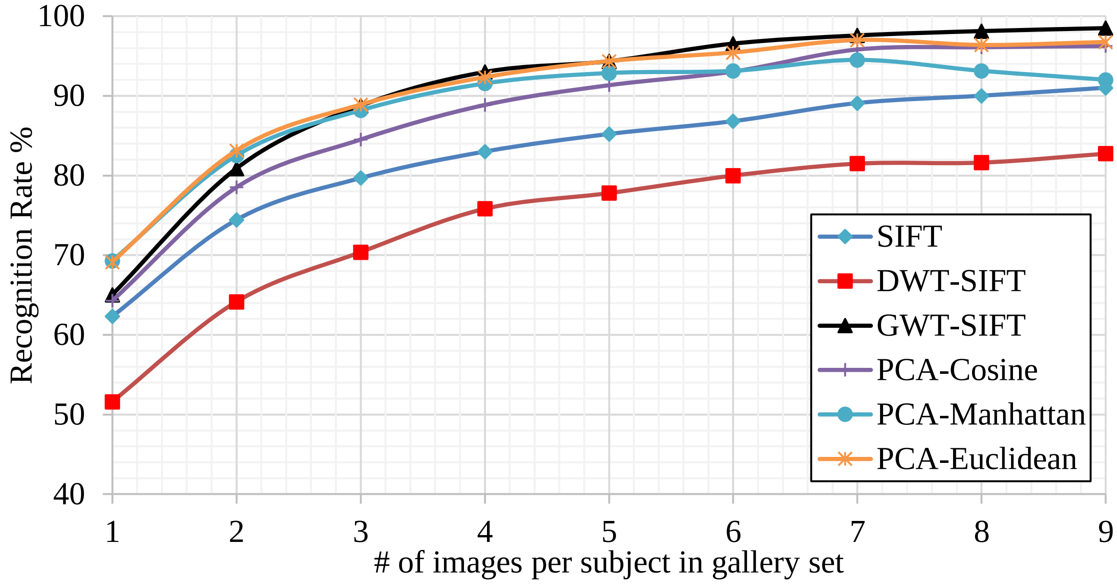


Figure 4.23. Recognition performance of SIFT, DWT-SIFT, GWT-SIFT and PCA.

4.8. General Discussion of Results

In reference to the above observations, it can be clearly stated that, the use of transformation on face images before extracting features significantly improves the recognition rates of the studied face recognition system. In general, the DWT-SIFT, GWT-SIFT, DWT-SURF or GWT-SURF outperforms the SIFT or SURF in terms of recognition performance using PUT database and GWT with SIFT or SURF but in case of computation time it was slower than SIFT or SURF.

Gabor wavelets possess many properties which make them attractive for many applications. Directional selectivity is one of the most important of these properties. The Gabor wavelets can be oriented to have excellent selectivity in any desired direction. They respond strongly to image features which are aligned in the same direction and their response to other feature directions is weak. Invariance properties to shifts and rotations also play an important role in their success. In order to accurately capture local features in face images, a space frequency analysis is desirable. Gabor functions provide the best tradeoff between spatial resolution and frequency resolution. The optimal frequency-space localization property allows Gabor wavelets to extract the maximum amount of information from local image regions. This optimal local representation of Gabor wavelets makes them insensitive and robust to facial expression changes in face recognition applications.

While DWT has limited directionality and it filters most of the textual information from the small size of images. This makes the DWT to decrease the performance of proposed approaches.

Our proposed approach was good in performance but for computation time since it adds a preprocessing stage it affects computation time.

Finally, it can be concluded that, image pre-processing on the acquired images before recognition stage, has a potential of increasing the recognition performance of face recognition system, while increasing computation time.

CHAPTER FIVE

CONCLUSION AND FUTURE WORK

5.1. Conclusion

Automatic face recognition is a main research area in computer vision and image processing which tries to detect and recognize human face without human interaction. Important developments in this field proves that in many of the face recognition application areas the automated techniques perform better than human. Face recognition encompasses many fields and study areas, such as pattern recognition, bioinformatics, and machine vision, and has become one of the hottest research areas. The main problem in face recognition systems is how to find a distinctive feature set to identify and recognize a face. So far many algorithms about feature extraction have been developed and proposed, which mostly include three aspects: face geometry, statistical and facial features.

In this thesis work, SURF or SIFT are used to extract features from face images. However, after SURF or SIFT are successfully applied for the feature detection and description, two approaches are proposed to improve the results. The first approach is based on DWT with SURF or SIFT namely DWT-SURF or DWT-SIFT. The second approach is based on GWT with SURF or SIFT namely GWT-SURF or GWT-SIFT. The DWT or GWT is applied to the image as a preprocessing stage before conventional SURF or SIFT algorithm. The recognition results obtained using this technique show substantial improvements, especially, in the recognition performance.

The performances of the two proposed approaches have been measured using widely used databases ORL and PUT. Performance of the first proposed approach has been tested on both ORL and PUT face images with different number of images per subject and gallery sets. The proposed approach is found to perform well in face recognition. Performance of the second proposed approach has been tested on ORL and PUT face images. The results obtained show that our proposed approach performs better than the SURF, SIFT, and conventional algorithms.

5.2. Future Work

In this work, DWT or GWT is used with the conventional SIFT or SURF to propose two approaches to improve the performance of face recognition. As a future work, the followings may be considered:

- In all the experiments, for the 1-scale and 2-scales, it is noticed that the performances of the proposed approach are close to each other. Hence, it can be applied for different scales, and the results may be compared to that of the SURF or SIFT.
- ORL and PUT databases were used in all of the experiments; more databases can be tested and the results may be compared with the proposed approach and conventional algorithms.

REFERENCES

[1] R. Gottumukkal and V. K. Asari, "An Improved Face Recognition Technique based on Modular PCA Approach", Pattern Recognition Letters, Elsevier, pp. 429-436, Dec. 2013.

[2] Z. Liu, Z. Xu and L. Hu, "A Comparative Study of Distance Metrics used in Face Recognition", Journal of Theoretical & Applied Information Technology, vol. 5 no. 2, 2009.

[3] V. Atamuradov, A. Eleyan and B. Karlik, "Performance Evaluation for Face Recognition using Wavelet-Based Image De-Noising", IEEE International Conference on Technological Advances in Electrical, Electronics and Computer Engineering, pp. 284-287, May 2013.

[4] M. Pietikäinen, A. Hadid, G. Zhao and T. Ahonen, "Local Binary Patterns for Still Images", Computer Vision Using Local Binary Patterns, Springer London, 2011.

[5] K. Baek, B. A. Draper, J. R. Beveridge and K. She, "PCA vs. ICA: A Comparison on the FERET Data Set", Jornal of Computer Information System, pp. 824-827, Mar. 2002.

[6] B. A. Draper, K. Baek, M. S. Bartlett and J. R. Beveridge, "Recognizing Faces with PCA and ICA", Computer Vision and Image Understanding, vol. 91, no. 1, pp. 115-137, 2003.

[7] A. Eleyan and H. Demirel, "PCA and LDA based Neutral Networks for Human Face Recognition", Face Recogniton, Kresimir Delac and Mislav Grgic (Ed.), ISBN: 978-3-902613-03-5, inTech. 2007.

[8] R. Steinhoff, B. Wongsaroj, C. Smith, A. Gallon and J. Franscis, "Improving Eigenface Face Recognition by using Image Registration Preprocessing Methods", Dept. of Computer Science and Mathematics, Florida Memorial University, Miami, Florida 33054, USA.

[9] M. Turk and A. Pentland, "Eigenface for Recognition", Journal of Cognitive Neuroscience, vol. 3, no. 1, 1991.

[10] A. P. Ginsburg, "Visual Information Processing based on Spatial Filters Constrained by Biological Data", AMRL tech. Rep., pp. 78-129, 1978.

[11] L. D. Harmon, "The recognition of Faces", Scientific American, vol. 229, pp. 71-82, 1973.

[12] J. Sergent, “Microgenesis of Face Perception”, Aspects of Face Processing, H. D. Ellis, M. A. Jeeves, F. Newcombe, and A. Young Eds. Dordrecht: Nijhoff, 1986.

[13] R. Chellappa, C. Wilson, and S, Sirobey, “Human and Machine Recognition of Faces: a Survey”, Proceedings of IEEE, vol. 83, May 1995.

[14] V. Bruce and A.W. Young, “Understanding Face Recognition”, British Journal of Psychology, vol. 77, pp. 305-327, 1986.

[15] S. Lawrence, C. Giles, A. Tsoi, and A. Back, “Face Recognition: A Convolutional Neural Network Approach”, IEEE Trans. on Neural Networks, vol. 8, pp. 98-113, 1997.

[16] M. Lades, J Vorbruggen, J, Buhmann, J. Lange, von der Malsburg, and R. Wurtz, “Distortion Invariant Object Recognition in the Dynamic Link Architecture”, IEEE Trans. on Computing, vol. 42, no. 3, pp. 300-311, 1993.

[17] R. Steinhoff, B. Wongsaroj, C. Smith, A. Gallon and J. Franscis, “Improving Eigenface Face Recognition by using Image Registration Preprocessing Methods”, Dept. of Computer Science and Mathematics, Florida Memorial University, Miami, Florida 33054, USA.

[18] Lowe, D. G., “Distinctive Image Features from Scale-Invariant Keypoints”, International Journal of Computer Vision, vol. 60, no. 2, pp. 1-28, 2004.

[19] H. Bay, T. Tuytelaars, L. Gool, “SURF: Speeded Up Robust Features”, Proc. 9th European Conference on Computer Vision, pp. 404 417, 2006.

[20] N. G. Kingsbury, “Complex Wavelets for Shift Invariant Analysis and Filtering of Signals”, Applied and Computational Harmonic Analysis, vol. 10, no. 3, pp. 234–253, 2001.

[21] D. Gabor, “Theory of Communication”, J. IEE, vol. 93, pp. 429-459, 1946.

[22] A. Eleyan, H. Özkaramanli, and H. Demirel, “Complex Wavelet Transform-based Face Recognition”, EURASIP Journal on Advances in Signal Processing, 2008.

[23] B.S. Manjunath, R. Chellappa, and C. von der Malsburg, “A Feature Based Approach to Face Recognition”, Proc. of International Conf. on Computer Vision, 1992.

[24] L. Wiskott, J. M. Fellous, N. Krüger and Christoph von der Malsburg, “Face Recognition by Elastic Graph Matching”, In Intelligent Biometric Techniques in fingerprint and Face Recognition, CRC Press, Chapter 11, pp. 355-396, 1999.

[25] A. Kasiński, A. Florek and A. Schmidt, “The PUT Face Database”, Image Processing & Communications, vol. 13, no. 3-4, pp. 59-64, 2008.

[26] F. Samaria, A. Harter, "Parameterization of a Stochastic Model for Human Face Identification", Proceedings of 2nd IEEE Workshop on Applications of Computer Vision, Sarasota FL, December 1994.

[27] A. Eleyan, H. Demirel, "Face Recognition System based on PCA and Feedforward Neural Networks", Computational Intelligence and Bioinspired Systems, pp. 935-942, Springer 2005.

PUBLICATIONS

- M. M.Ameen, G. Eleyan, A. Eleyan, “Wavelet Transform based Face Recognition using SURF Descriptors”, 3rd International Conference on Electrical and Electronics Engineering (ICEEE), Turkey, 2016. (**Submitted**)
- M. M.Ameen, A. Eleyan, “Wavelet Transform based Score Fusion for Face Recognition using SIFT Descriptors”, 24th IEEE Conference on Signal Processing and Communications Applications (SIU), Turkey, 2016. (**Submitted**)