Enhanced Method for Handwritten Signature Recognition

طريقة محسنة لتمييز التوقيع اليدوي

Prepared by

Maha Mahmoud Al-Saidi

200810408

Supervisor

Dr.Muzhir Shaban Al-Ani

A Thesis Submitted in Partial Fulfillment of the Requirement for the Degree of Master in Computer Science

Department of Computer Science

College of Computer Science and Informatics

Amman Arab University

February, 2011

Authorization

I, the undersigned "Maha Mahmoud Al-Saidi" hereby authorize Amman Arab University to provide copies of this thesis to libraries, institutions and any other parties upon their

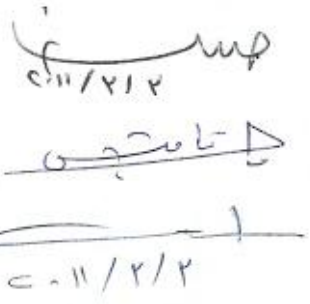
Name: Maha Mahmoud Al-Saidi

 Signature:

Date: 9/ 3 / 2011

Resolution of the Examining Committee

This thesis titled "Enhanced Method for Handwritten Signature Recognition". Has been defended and approved on 9 /3 /2011.

Examining Committee	Signature
Dr. Hussein Al-Bahadili	
Dr. Malek Kakish	
Dr. Muzhir Al-Ani	

Dedication

To my family,

My supervisor Dr. Muzhir Al-Ani for his help and support,

And to all my friends.

Acknowledgment

Initially, I want to thank ALLAH for all things I reached today. Also, I would like to thank my supervisor Dr Muzhair Al-Ani for his help to complete this thesis and especially in collection samples for the study, and I would like to thank my best friend Ala'a Khresat for her support. Finally, I would like to thank my family for their help and support in this life.

Abstract

Electronic biometric systems became available in last few decades. Handwriting signature recognition as a behavioral biometric plays an important role in our life, because of its widespread acceptance among people, especially in security transactions.

In this thesis, an enhanced method for handwritten signature recognition has been developed and implemented in a comprehensive system for handwritten signature recognition. The system includes many elements that are integrated to generate the entire system. Database of 1000 Arabic and English signatures has been introduced. The system has been separated into two important phases: training and testing.

In the training phase, preprocessing process is applied firstly for signatures enhancement, then feature extraction process is applied using discrete wavelet transform (DWT) and vectors fusion. In enrollment process, after support vector machine (SVM) is trained, an individual template is stored in the database.

In the testing phase, two recognition processes have been introduced: verification and identification. In each process, SVM classifier makes a decision for classification.

The system indicates good results for verification and identification compared to other approaches.

الخلاصة

توفرت الأنظمة البيومترية الإلكترونية منذ العقود القليلة الماضية.تقوم عملية التعرف على التوقيع اليدوي كسلوك بيومتري بدور مهم في حياتنا، وذلك لقبوله الواسع بين الناس و خاصة في المعاملات الأمنية.

تقدم هذه الأطروحة، طريقة محسّنة لتمييز التوقيع اليدوي، تم تطويرها وتنفيذها في نظام شامل لتمييز التوقيع اليدوي. يتضمن النظام عدة أساسيات إندمجت لتوليد النظام الكلي.وقد تم عرض قاعدة بيانات مكونة من 1000 توقيع باللغتين العربية والإنجليزية . وهو مكوّن من مرحلتين هامتين: التدريب و الإختبار.

قامت الباحثه في مرحلة التدريب، بدايةً بتطبيق عملية تجهيزية لتحسين التواقيع، تلتها عملية إستخراج الخصائص من خلال استخدام الموجة المقطعة الموجهه (DWT) و إندماج الناقلات (Vectors Fusion). في عملية التسجيل، بعد تدريب الجهاز الداعم الموجه (SVM) يخزن قالب الفرد في قاعدة البيانات.

في مرحلة الإختبار، تم عرض عمليتين للتعرف: التحقق و تحديد الهوية. في كل عملية يتخذ جهاز الدعم الموجه (SVM) المصنف القرار للتصنيف.

أشارت نتائج النظام إلى تحقيق مستويات جيدة في التحقق وتحديد الهوية مقارنة بالطرق الأخرى.

Contents

List of Figures

List of Tables

List of Abbreviations

AMT	Ammar Matching Technique
DTW	Dynamic Time Warping
DWT	Discrete Wavelet Transform
EER	Equal Error Rate
FAR	False Acceptance Rate
FRR	False Rejection Rate
GA	Genetic Algorithm
HMM	Hidden Markov Model
ICDAR	International Conference on Document Analysis and Recognition
ICFHR	International Conference on Frontiers in Handwriting Recognition
ICR	Intelligent Character Recognition
KNN	K-Nearest Neighbor
MDF	Modified Direction Feature
NHR	Natural Handwriting Recognition
NN	Neural Network
OCR	Optical Character Recognition

SRM	Structural Risk Minimization
SVC	Signature Verification Competition
SVM	Support Vector Machine
URL	Uniform Resource Locator
JPEG	Joint Photographic Experts Group
PPI	Pixel Peer Inch
HWSR100	Handwritten Signature Recognition 1000
HWSRFB	Handwritten Signature Recognition Fusion Based

Chapter one
Introduction

1. Introduction

Handwriting has been a medium of interaction between people across space and time, through exchanging messages and ideas for thousands of years ago [1].

In modern society, there is an increasing request for more and more reliable personal verification systems. In fact, personal verification plays a very important role not only in personal security, but also in data protection and transaction validation [2].

The analysis of handwritten documents has been a subject of intensive research for the last decades. The interest devoted to this field is not only explained from the scientific point of view, but also in terms of the social benefits that convey those systems [3].

Nowadays, even with the modern technologies, handwriting is still a useful and easy communication method because of the convenience of using pen and paper in various daily situations [1].

Handwriting recognition attracted the attention of researchers since the inception of computers. Nowadays, the technological progress made in the field of computer architectures and peripheral devices, as well as the advances of scientific research, make the development of new systems for handwriting recognition possible [4].

Signatures are a special case of handwriting subject to interpersonal variations and differences. This variability makes it necessary to analyze signatures as complete images and not as collections of letters and words [5].

Handwritten signatures and personal signs belong to the behavioral biometric characteristics as the person must become active for signing [6]. This chapter talks briefly about Handwriting Signatures as a behavioral biometric, Handwriting Signature Recognition, some related applications, statement of the problem, previous work, objectives of the thesis, and chapters' outlines.

1.1 Handwritten Signature: A Behavioral Biometric

Biometrics is the science or technology which analyzes and measures the biological data. In computer science, it refers to the science or technology that measures and analyzes

physical or behavioral characteristics of a person, for authentication [7]. Figure 1.1 illustrates two biometrics types: physiological and behavioral.

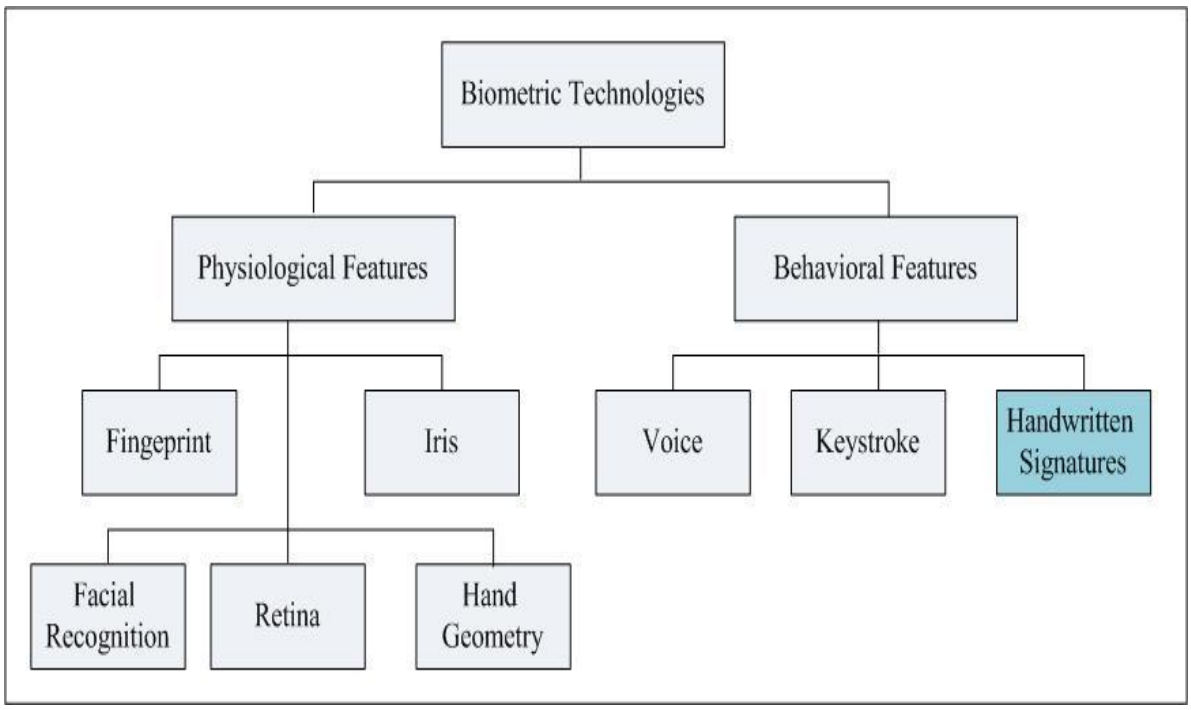


Figure 1.1 Biometrics types

Physiological biometrics are based on data derived from direct measurements of a part of the human body. Fingerprints, iris-scans, retina-scans, hand geometry, and facial recognition are all leading physiological biometrics.

Behavioral characteristics are based on an action taken by a person. Behavioral biometrics, in turn, are based on measurements and data derived from an action, and indirectly measured characteristics of the human body. Voice recognition, keystroke-scans, and signature-scans are

leading behavioral biometric technologies. One of the defining characteristics of a behavioral biometric is the incorporation of time as a metric – the measured behavior has a beginning, middle and end. Figure 1.2 illustrates examples of biometrics authentication for physiological and behavioral biometrics.

A key point is that while behavioral biometrics are based on an individual's actions, those actions are in turn influenced by physiological attributes such as the size of a person's hand (signature-scan) or the shape of their vocal chords (voice recognition) [8].

Biometric systems have been researched and tested for a few decades, but have only recently entered into the public consciousness because of high profile applications, usage in entertainment media (though often not realistically) and increased usage by the public in day-to-day activities.

Many factors must be taken into account when implementing a biometric device including location, security risks, task (identification or verification), expected number of users, user circumstances, existing data, etc. It is also important to note that biometric modalities are in varying stages of maturity [9].

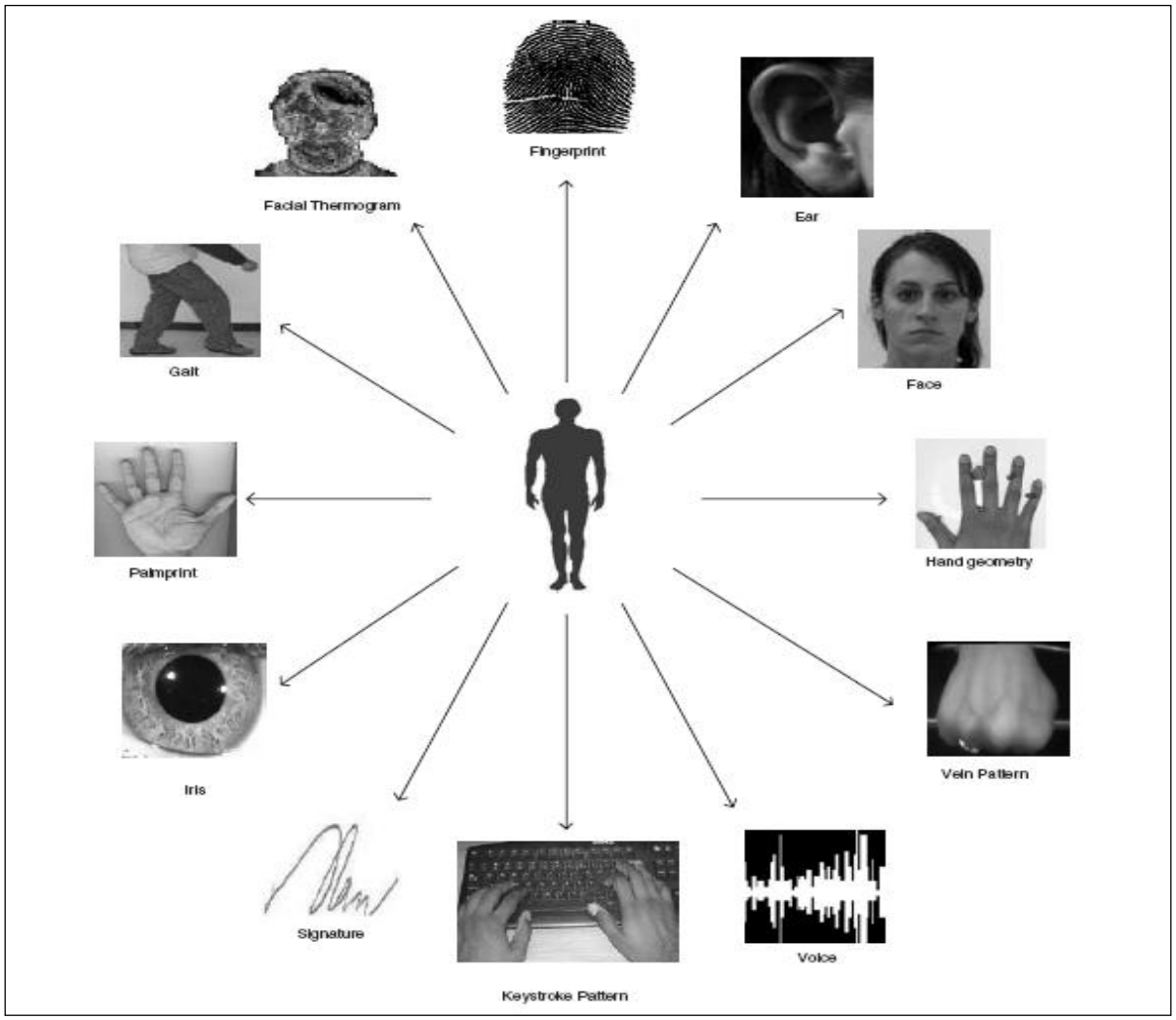


Figure 1.2 Example of biometrics for authentication

Biometrics have been used for identification and recognition for a long time. Signatures and fingerprints are commonly used for identification in banks, documents related to financial deals etc. Photographs are used for identification for passport, license etc. However, the processes are still manual and time consuming in many applications. Therefore, automatic identification using biometrics has gained lot of popularity in recent times [7].

In general, there are many types of biometrics technologies. Biometric types could be categorized into**:**

- **Recognition:** is a generic term, and does not necessarily imply either verification or identification. All biometric systems perform "recognition" to "again know" a person who has been previously enrolled.
- **Verification:** is a task where the biometric system attempts to confirm an individual's claimed identity by comparing a submitted sample to one or more previously enrolled templates.
- **Identification:** is a task where the biometric system attempts to determine the identity of an individual. A biometric is collected and compared to all the templates in a database. Identification is "closed-set" if the person is known to exist in the database. In "open-set" identification is sometimes referred to as a "watchlist". The person is not guaranteed to exist in the database. The system must determine whether the person is in the database or not [9].

1.2 Handwritten Signature Recognition

Signature is an image carrying a certain pattern of pixels that pertains to a specific individual [10].

The way a person signs his name is known to be the characteristic of that individual. Although signatures require contact and effort with the writing instrument, they are considered acceptable in many governmental, legal, and commercial transactions as a method of personal authentication [11].

A number of features that are suitable for automated comparison can be extracted from handwritten signatures. The features depend on the type of data captured and the chosen comparison method. Handwritten data could be classified into**:**

- Off-line data captured from paper after the writing process using devices such as image scanners or cameras.
- On-line data captured during the writing process using devices such as digitizing tablets, Tablet PC's, or special pens.

1.2.1 Off-line Signatures

Off-line signature recognition is the recognition of handwritten signatures based on a two-dimensional gray image obtained by an item of equipment such as a scanner, an optical reader, or a digitizer.

Since the invention of writing in human society, the signature on a document (or picture, including monochrome brush painting) has been the most common means of authenticating the writer (or painter) of the document (or picture). Not only leaders but also persons accorded with responsibility in various capacities have had to put their signatures on paper and recognize those of others. Thus, signature recognition has naturally been done off-line. It is only with the recent development of an on-line technology for biometric recognition that the relevance of off-line recognition has been reduced. The relative characteristics of on-line and off-line methods of signature recognition will be discussed later.

The written name of the writer was originally used as signature. In the course of the traditional use of signatures, people started including symbols and distorting them in order to increase their uniqueness and beauty. In general, this made it impossible to recover the writers' names from such signatures. Furthermore, some signatures, like names, are merely personal signs that help establish authenticity. In this sense, Off-line signature recognition merely entails pattern recognition of questionable images on a two-dimensional space by referring to registered reference images, which may be reduced to a template.

1.2.2 On-line Signatures

On-line signature verification uses data obtained while a signature is being written. The data obtained during the process of writing a signature is called an on-line signature. On-line signature verification is based on the hypothesis that the writing style of a signature differs from person to person and cannot be easily forged. Figure 1.3 illustrates example of online signature tablets and special pen.

Figure 1.3 Online signature tablet PC's and special pen

On-line signature verification verifies whether an input signature is a genuine signature or a forgery. Ideally, it is a two-class partitioning problem; however, it is not an easy problem to solve, because of the following reasons**:**

- People do not reproduce their signature exactly each time. Characteristics of the writing manner of genuine writers can change over time. There is, necessarily, intra-class (intra-person) variability. In contrast, forgers attempt to make their forged signatures as similar as possible to genuine signatures, and thus inter-class (inter-person) variability decreases. Therefore, it is difficult to distinguish between genuine signatures and forgeries.

- Both the number and type of signatures available for training are often severely limited. As on-line signature verification is a two-class partitioning problem, general pattern recognition techniques can be applied if enough data is available from both classes. In practice, however, only a few genuine signatures are available from the genuine class. Moreover, there are several types of forgeries in the forgery class, but only a few types of forgery can be collected for the following reasons: Forgeries that are most similar to genuine signatures and the most difficult to distinguish from genuine signatures will be signatures that were produced by imitating genuine signatures well. Because genuine signatures differ from writer to writer, well-imitated forged signatures should be collected for every

- writer; however, this is extremely difficult. Scarcity of genuine training data exists in all biometric methods; however, scarcities of forgeries exist only in methods that must prepare for imitation attacks.

1.2.3 Applications

Handwritten signatures are generally used for verification (confirming a claimed identity through one-to-one comparisons of biometric features), but rarely for identification (finding identifiers attributable to a person through one-to-many search of biometric features in a large database). Handwritten signatures have been used for a long time for authentication purposes in many applications, such as credit cards, banking transactions, agreements, and legal documents. Off-line signatures serve as a unique means to verify the authenticity of a person through past records, such as signatures on traveler's cheques [6].

1.3 Statement of the Problem

The researcher sought to develops an electronic offline signature verification system that used to authenticate a large number of documents in limited time, because the manual verification is often unrealistic.

Signatures are often forged for the purpose of violating the privacy of person and their documents (i.e. checks).

According to the National Check Fraud Center, check fraud and counterfeiting are among the fastest-growing problems affecting the nation's financial system, producing estimated annual losses of $10 billion and losses continue to rise at an alarming rate annually [12].

Eight in ten banks (80 %) incurred check fraud losses in 2006, up from 75% in 2003. The total amount of attempted check fraud against banks' deposit accounts reached an estimated $12.2 billion in 2006. Most of the attempts (92 %) were caught by banks' prevention systems or measures before incurring any financial loss to the bank [13]. This leads to the problem of being able to correctly verify whether a signature is genuine or a forgery.

In this thesis , the researcher sought to construct a system that is more accurate and faster for recognition process comprised to existing approaches that are used to solve the problem of separation between genuine variation of an individual’s signatures and that of forgery signatures by using the concept of Discrete Wavelet Transform (DWT), fusion and Support Vector Machine (SVM), where Offline methods do not require some special hardware like digitizers,

pressure sensitive tablets to capture the dynamic features which the online methods require. DWT and its capability of multi-resolution analysis will be used in feature extraction phase to increase operation speed and system efficiency. Then, SVM classifier was employed to verify signatures, as it has a good generalization performance. However, there is no signature verification system available to be used in all financial and security daily documents authentication transactions.

1.4 Overview of Concepts

1.4.1 Discrete Wavelet Transform

The multi-resolution wavelet is used for texture analysis in literature, possibly due to its finite duration, which provides both frequency and spatial locality.

Dwt also, provides sufficient information both for analysis and synthesis of the original signal, with a significant reduction in the computation time [14]. The hierarchical wavelet transform uses a family of wavelet functions and its associated scaling functions to decompose the original signal/image into different sub bands; Figure 1.4 illustrates 2D-DWT multiresolution decomposition process. The decomposition process is recursively applied to the sub bands to generate the next level of the hierarchy. At each iteration of the DWT, the lines of the input image

(obtained at the end of the previous iteration) are low-pass filtered and high pass filtered. Then the lines of the two images obtained at the output of the two filters are decimated with a factor of 2. Next, the columns of the two images obtained are low and high pass filtered. The columns of those four images are also decimated with a factor of 2. Four new sub-images (representing the result of the current iteration) are generated. The first one, obtained after two low-passes filtering, is named approximation sub-image (or LL image). The remaining three are named detail sub-images: LH, HL and HH. The LL image represents the input for the next iteration [15]; Figure 1.5 illustrates 2D-DWT image decomposition level.

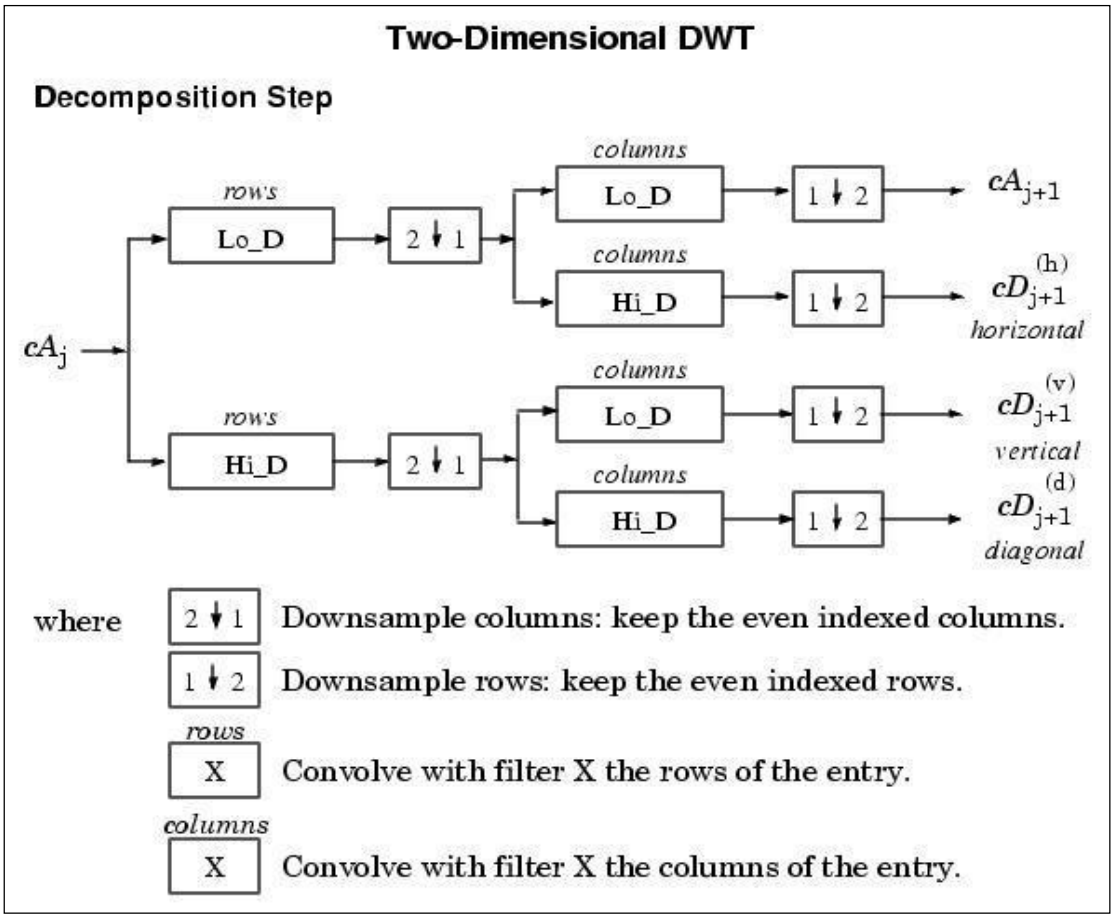


Figure 1.4 2D-DWT multiresolution decomposition

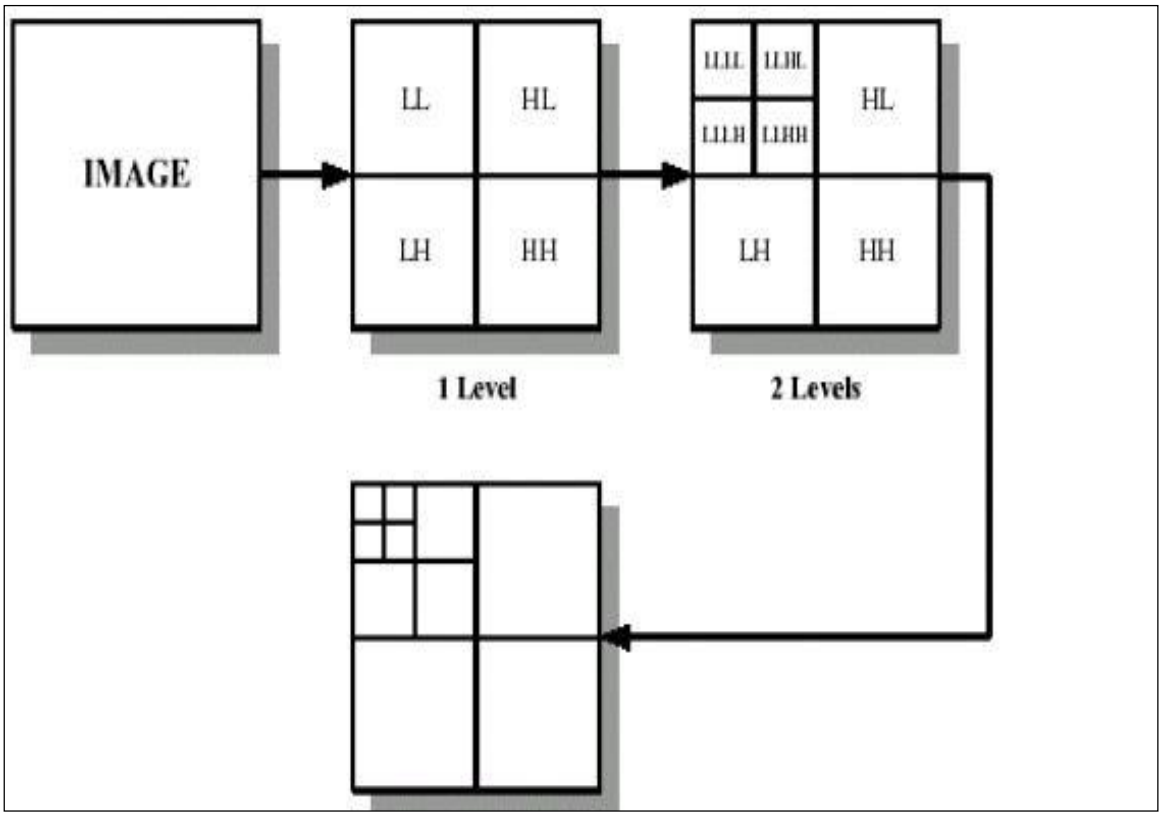


Figure 1.5 2D-DWT image decomposition

1.4.2 Fusion

Image fusion is the process of combining relevant information from multiple images of the same scene. The result of image fusion is a new image which is more suitable for human and machine perception or further image processing tasks such as segmentation, feature extraction and object recognition [16].

Objectives of Image Fusion Schemes

- Extract all the useful information from the source images
- Do not introduce artifacts or inconsistencies which will distract human observers or the following processing.
- Reliable and robust to imperfections such as mis-registration [17].

1.4.3 Support Vector Machine

SVM is a promising classification technique developed by Vapnik. It has a good generalization performance even under the conditions of small training sets [18]. SVM classification task usually involves separating the data into training and testing sets.

SVM is a technique in the field of statistical learning theory. It is based on the structural risk minimization principle (SRM). The SRM induction principle has two main objectives. The first is to control the empirical risk on the training data set. The second is to control the capacity of the decision functions used to obtain this risk value [19], where SVM classifier is an algorithm which maximizes the margin between the classes and minimizes the classification error [20].

In training sets, **L** training points we introduced, where each input $\boldsymbol{x_i}$ has a feature vector and is in one of two classes $\boldsymbol{y_i}$ = -1 or +1. Training sets are illustrated in the following equation:

$$\{\boldsymbol{x_i}, \boldsymbol{y_i}\} \quad where\ \boldsymbol{i} = 1 \ldots\ldots \mathrm{L}, \boldsymbol{y_i} \in \{-1, +1\}, \boldsymbol{x_i} \in \mathrm{R^n} \qquad (1.1)$$

Each variable belongs to two separate classes, H1 ($\boldsymbol{y_i}$ = +1) and H2 ($\boldsymbol{y_i}$ = -1), where $\boldsymbol{y_i}$ is the target output for training data $\boldsymbol{x_i}$. The SVM finds the hyperplane with maximum Euclidian distance from the training set, which is illustrated in Figure 1.6, where **w** is the maximal margin to optimal hyperplane, defined as the sum of distances from the hyperplane to the closest points of the classes [19], and $\frac{\boldsymbol{b}}{||\boldsymbol{w}||}$ is the distance from the hyperplane to the origin.

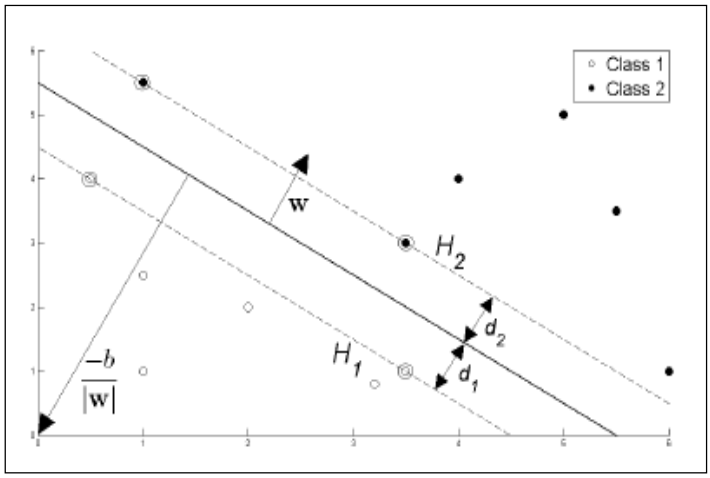


Figure 1.6 Optimal separation hyperplane with a maximal margin

All training sets satisfy the following constraints for not fully linearly separable data; this is done by introducing a positive slack variable $\Phi_i, i = 1, \ldots . L$:

$$x_i \,.w + b \;\geq\; +1 - \Phi_i \qquad for\;\; y_i = +1 \tag{1.2}$$

$$x_i \,.w + b \;\leq\; -1 + \Phi_i \qquad for\;\; y_i = -1 \tag{1.3}$$

$$\Phi_i \;\geq 0 \;\forall_i \tag{1.4}$$

This can be combined into:

$$\boldsymbol{y_i}(\boldsymbol{x_i}\,.w + b\,) - 1 + \Phi_i \geq 0 \qquad where\;\; \Phi_i \;\geq 0 \;\forall_i \tag{1.5}$$

Classification function for SVM liner kernel is illustrated as the following:

$$f(x) = \text{sign}\Big(\sum_{i=1}^{e}(\,a_i\, y_i\, K\,(x, x_i) + b)\Big) \tag{1.6}$$

Where ***f(x)*** is a decision function, $\boldsymbol{a_i}$ is the lagrance multiplier assigned to each training data whose value depends on the role of training the data in the classifier system, ***e*** denotes the number of support vectors, **K** is the linear kernel function ***b*** is the bias and $\boldsymbol{y_i}$ is the classifier output for the test data$\boldsymbol{x_i}$. Where Linear Kernel is as the follows:

$$K(X, X_i) = \; X^T \, X_i \tag{1.7}$$

1.5 Previous Work

In this thesis, offline signature recognition system has been developed based on DWT, Fusion and SVM. There are many researches related to this field, some of which are:

1- Justino Edson, et al (2004) made a comparison between SVM and HMM under two specific conditions, the first being the number of samples used for training, and the second being the use of different types of forgeries [19].

2- Wei Ji, et al (2005) decomposed the pen-position parameters of online signature into multi scale signals by wavelet transform technique, then the distance between the input signature and the reference signature of the corresponding zero-crossing representations are computed as the feature. Then build a binary SVM classifier [18].

3- Özgündüz Emre, et al (2005) developed signature verification and recognition system by using the global, directional and grid features based on SVM [21].

4- Naguyen Vu, et al (2007) used a method to perform signature verification based on intelligent techniques. Structural features are extracted from the signature's contour using Modified Direction Feature (MDF) and its extended version. Two Neural Network (NN)-based techniques and SVMs were investigated and compared for the process of

signature verification [22].

5- Kisku Dakshina, et al (2009) presented a method involving signature image, global and local features are extracted and the signatures are verified with the help of Gaussian empirical rule, Euclidean and Mahalanobis distance based classifiers. SVM is used to fuse matching scores of these matchers [23].

6- Kiani Vahid, et al (2009) developed a method that uses Radon Transform locally as a feature extractor and SVM as a classifier. The main idea of this method is using Radon Transform locally for line segments detection and feature extraction, against using it globally [24].

7- Jawarkar NP, et al. (2009) developed a system in which a person uses mouse movement to imitate/simulate normal finger movement rhythm to insert signature. This inserted signature is processed and filtered, and DWT, slope features are extracted and verification is carried out using feed-forward NN [25].

8- Ali A, et al (2009) presented a system for offline signature verification, approaching the problem as a two-class pattern recognition problem (SVM and K-Nearest Neighbor (KNN)). They used discrete Radon Transform to extract global features from the signatures [26].

9- Fauziyah S, et al (2009) developed an online signature verification system using SVM and VBTablet 2.0 to verify the input signature by comparing database [27].

10- Ghandali Samaneh, et al (2009) developed a method for off-line Persian signature identification and verification that is based on Image Registration, DWT and Image Fusion [28].

1.6 Objectives of the Thesis

The main objectives of this thesis are:

1. Construct a HWSR1000 database of signatures.

2. Develop a reliable, accurate, and quick electronic offline signatures recognition system.

3. Evaluate the performance of the developed system.

1.7 Outline of Forthcoming Chapters

Chapter2: describes Handwritten Signature Recognition system in more detail.

Chapter3: is concerned with system development, database construction, and system training and testing phases.

Chapter4: lists experimental results.

Chapter5: Based on the obtained results, a number of conclusions are drawn, and a number of recommendations for future tasks are pointed-out.

Chapter two
Handwritten signature recognition

Handwritten Signature Recognition is an important issue in many fields. This section includes some reviews about image processing, as signatures are considered as images, and a brief review about Handwriting Recognition, Signature Recognition, and many Signatures Recognition Approaches.

2.1 Image Processing

Images are pictures: a way of recording and processing information visually. Pictures are important to us because they can be an extraordinarily effective medium for the storage and communication of information [29].

A digital image is composed of pixels which can be thought of as small dots on the screen. A digital image is an instruction of how to color each pixel. A typical size of an image is 512-by-512 pixels. In the general case image is of size m-by-n if it is composed of **m** pixels in the vertical direction and **n** pixels in the horizontal direction [30].

Image Processing is a technique to enhance raw images received from cameras/sensors placed on satellites, space probes and aircrafts or pictures taken in normal day-to-day life for various applications.

Various techniques have been developed in Image Processing during the last four to five decades. Most of the techniques are developed for enhancing images obtained from unmanned spacecrafts, space probes and military reconnaissance flights. Image Processing systems are becoming popular due to easy availability of powerful personal computers, large size memory devices, graphics software etc. The common steps in image processing are image scanning, storing, enhancing and interpretation.

2.1.1 Methods of Image Processing

There are two methods available in Image Processing:

***(a) Analog Image Processing*:**

Analog Image Processing refers to the alteration of image through electrical means. The most common example is the television image.

The television signal is a voltage level which varies in amplitude to represent brightness through the image. By electrically varying the signal, the displayed image appearance is altered. The brightness and contrast controls on a TV set serve to adjust the amplitude and reference of the video signal, resulting in the brightening, darkening and alteration of the brightness range of the displayed image.

(b) Digital Image Processing:

In this case, digital computers are used to process the image. The image will be converted to digital form using a scanner – digitizer and then processed. It is defined as subjecting numerical representations of objects to a series of operations in order to obtain a desired result. It starts with one image and produces a modified version of the same. It is therefore a process that takes an image into another.

The term digital image processing generally refers to processing of a two-dimensional picture by a digital computer .In a broader context; it implies digital processing of any two-dimensional data. A digital image is an array of real numbers represented by a finite number of bits.

The principal advantage of Digital Image Processing methods is their versatility, repeatability and the preservation of original data precision [31].

The various Image Processing techniques are:

- Image representation
- Image preprocessing
- Image enhancement
- Image restoration
- Image analysis
- Image reconstruction

• Image data compression

Techniques and applications in the areas of image processing and pattern recognition are growing at an unprecedented rate. Containing the latest state-of-the-art developments in the field, Image Processing and Pattern Recognition present clear explanations of the fundamentals as well as the most recent applications. It explains the essential principles so readers will not only be able to easily implement the algorithms and techniques, but also lead themselves to discover new problems and applications [32]. Signatures are analyzed as a complete image and not as a letters or words.

2.2 Handwriting Recognition

Handwriting recognition is the ability of a computer to receive and interpret intelligible handwritten input from sources such as paper documents, photographs, touch-screens and other devices. The image of the written text may be sensed "off line" from a piece of paper by optical scanning (Optical Character Recognition (OCR)) or Intelligent Character Recognition (ICR). Alternatively, the movements of the pen tip may be sensed "online", for example by a pen-based computer screen surface [33].

Also handwriting recognition can be defined as the task of transforming text represented in the spatial form of graphical marks into its symbolic representation. Not only is useful for making digital copies of handwritten documents, but also in many automated processing tasks, such as automatic mail sorting or cheque processing. In automated mail sorting, letters are directed to the correct location by recognition of the handwritten address. Similarly, cheque processing involves recognizing the words making up the cheque amount [34]. The field of handwriting recognition can be split into two different approaches, offline and online.

2.2.1 Offline Recognition

Off-line handwriting recognition involves the automatic conversion of text in an image into letter codes which are usable within computer and text-processing applications. The data obtained by this form is regarded as a static representation of handwriting. Off-line handwriting recognition is comparatively difficult, as different people have different handwriting styles. OCR engines are primarily focused on machine printed text and ICR for hand "printed" text. There is no OCR/ICR engine that supports handwriting recognition till now.

2.2.2 Online Recognition

On-line handwriting recognition involves the automatic conversion of text as it is written on a special digitizer or PDA, where a sensor picks up the pen-tip movements as well as pen-up/pen-down switching. That kind of data is known as digital ink and can be regarded as a dynamic representation of handwriting. The obtained signal is converted into letter codes which are usable within computer and text-processing applications. Figure 2.1 illustrates online handwriting recognition elements.

The elements of an on-line handwriting recognition interface typically include**:**

- A pen or stylus for the user to write with.
- A touch sensitive surface, which may be integrated with, or adjacent to, an output display.
- A software application which interprets the movements of the stylus across the writing surface, translating the resulting strokes into digital text [33].

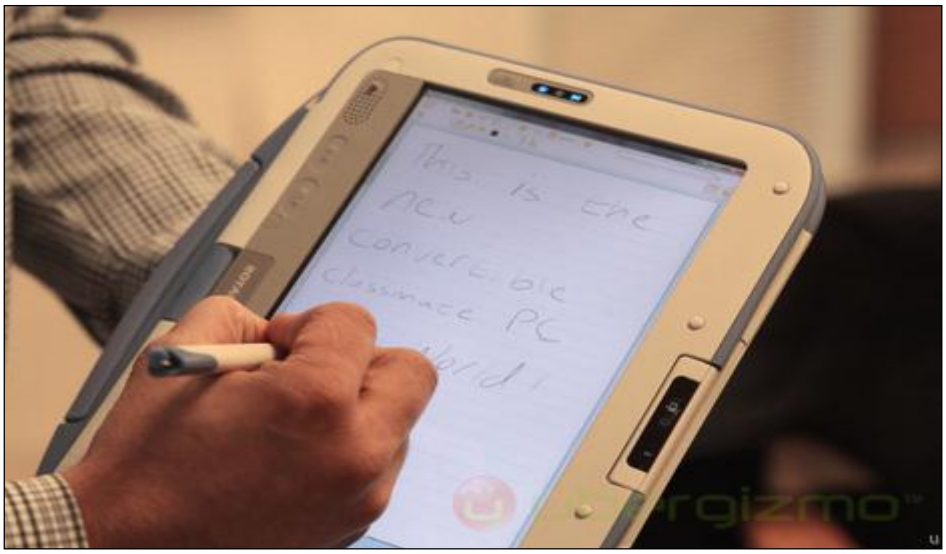


Figure 2.1Online handwriting recognition

2.2.3 Natural Handwriting Recognition

OCR and ICR still have limits to what they can read. OCR only reads machine printed characters. ICR has severe limitations when it comes to human handwriting. Today, new technology has arrived that will have a major impact on reading documents. This technology is known as Natural Handwriting Recognition (NHR). NHR allows computers to read and recognize handwriting with a high degree of accuracy. Unlike its cousins OCR/ICR, NHR uses a complex series of algorithms to compare and recognize each character. Knowledge of what it is reading along with a field image is essential for success in reading the characters, just as a human would not be able to read this article if he had no

knowledge of the language being used. NHR needs a dictionary of possible values of the field and a definition of the field type, such as name. The following outlines the basic NHR tasks:

- **Form Identification:** This task consists of identifying certain expected features on each form image presented for recognition. The output of the task is either the rejection of the form as unrecognizable or a set of locations of key features that identify the form as acceptable for further processing.
- **Field Isolation:** This task consists of extracting the text image of each data field from the form. The output of this task consists of one or more images of text minus the surrounding portions of the form.
- **Segmentation:** This task consists of breaking each image of text into smaller units for recognition. The output of this task is one or more image segments. Each segment is either the image of an isolated character, an image of an isolated piece of a character, or the image of an isolated group of connected or otherwise under segmented characters. In most cases, this will be images of an isolated character or symbol.

- **Recombining Segments:** This task consists of selecting various combinations of segments as plausible candidates for isolated character images. The output of this task is one or more isolated character-image candidates.
- **Recognition:** This task consists of assigning relative confidences to all of the allowed classes for each character-image candidate. The output of this task is either a single class or a set of ordered pairs consisting of character class and associated confidences. This output is called raw HWR to emphasize that it has been generated without the help of any context other than the one existing in the isolated character-image candidates.
- **Organizing Character Candidates:** This task consists of organizing the output of tasks, Recombining Segments, and Recognition into a form useful for the dictionary input stage. The output of this task is just the output of those tasks, in a format suitable for the particular dictionary look-up method being used.
- **Dictionary-Based Correction:** This task consists of selecting the dictionary entries that best match the properly organized character-image candidates according to some set criteria. The output of this task is the hypothetical answer provided by the HWR system

- as its final result and a confidence for that field.
- **Level of Acceptance:** This task consists of comparing the final result confidence level with set accuracy levels. The output of this task will be a rejection (the confidence level is too low) or acceptances of the final result as the conversion of the written data.
- **Rejection of the Result:** This task consists of transmitting the field image, the final result, to a workstation for human correction or acceptances [35].

2.2.4 Handwriting Recognition Research

Handwriting Recognition has an active community of academics studying it. The biggest conferences for handwriting recognition are the International Conference on Frontiers in Handwriting Recognition (ICFHR), held in even-numbered years, and the International Conference on Document Analysis and Recognition (ICDAR), held in odd-numbered years. Both of these conferences are scrutinized by the IEEE. Active areas of research include:

- Online Recognition
- Offline Recognition
- Signature Verification

- Postal-Address Interpretation
- Bank-Check Processing

2.2.5 Handwriting Recognition: Brief History

- 1888**:** U.S. Patent granted to Elisha Gray on electrical stylus device for capturing handwriting.
- 1915**:** U.S. Patent on handwriting recognition user interface with a stylus.
- 1942**:** U.S. Patent on touchscreen for handwriting input.
- 1957**:** Stylator tablet: Tom Dimond demonstrates electronic tablet with pen for computer input and handwriting recognition.
- 1961**:** RAND Tablet invented: better known than earlier Stylator system.
- 1962**:** Computer recognition of connected/script handwriting
- 1969**:** GRAIL system: handwriting recognition with electronic ink display, gesture commands.
- 1973**:** Applicon CAD/CAM computer system using the Ledeen recognizer for handwriting recognition.
- 1980s**:** Retail handwriting-recognition systems: Pencept and CIC both offer PC computers for the consumer market using a tablet and handwriting recognition instead of a keyboard and mouse. Cadre

- System markets Inforite point-of-sale terminal using handwriting recognition and a small electronic tablet and pen.
- 1989**:** Portable handwriting recognition computer: GRiDPad from GRiD Systems [33].

2.3 Signature Recognition

Signatures are a special case of handwriting in which special characters and flourishes are viable. Signature Verification is a difficult pattern recognition problem as no two genuine signatures of a person are precisely the same. Its difficulty also stems from the fact that skilled forgeries follow the genuine pattern unlike fingerprints or irises where fingerprints or irises from two different persons vary widely. Ideally interpersonal variations should be much more than the intrapersonal variations. Therefore it is very important to identify and extract those features which minimize intrapersonal variation and maximize interpersonal variations [10].

In biometric applications, there are two types of identity recognition methods: verification (authentication) and identification.

In the signature verification/identification, same set of features have been used. In verification, the individual claims his identity which is verified by comparing these feature

vectors of the individual which he claimed to be. If the matching score crosses the threshold, then the system verifies the individual as authentic user.

In identification, the feature vectors of the individual are compared with the feature vectors of every individual stored in the database. If the highest matching score crosses the threshold, then it identifies the individual as the person whose matching score is the highest, otherwise the system suggests few top most matches [7]. Figure 2.2 illustrates the architecture of biometric system for signature recognition.

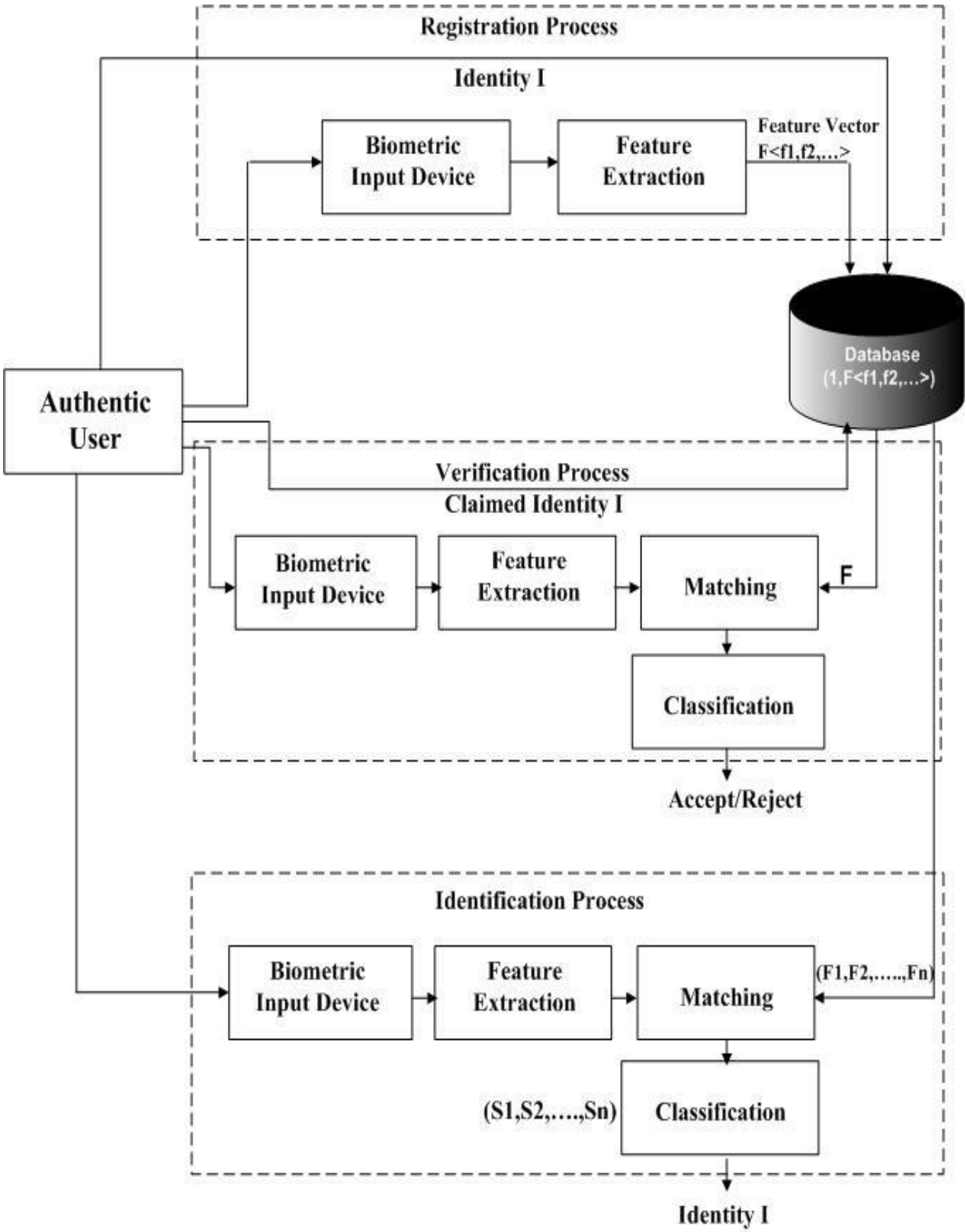


Figure 2.2 Biometric system architecture of signature recognition

2.3.1 Signature Recognition: Brief History

- 1677**:** England passed an act to prevent frauds and perjuries by requiring documents to be signed by the participating [27].

- The first published work on automatic signature verification seems to be Mauceri.
- 1977**:** Popular development of Herbst and Liu in, which summarized the state-of-the-art up to that date [36].
- Since 1989**:**
    - M. Ammar introduces a new technique for static signature verification which he calls AMT (Ammar Matching Technique).
    - The research of J. C. Pan and S. Lee centers on representing the signature image. Using base heuristics, the authors represent a signature as a series of elements that simulate the process of generating a handwritten stroke by a human.
    - R. Sabourin, M. Cheriet and G. Genest are evaluating a shade-coding method to eliminate random forgeries.

- R. Sabourin and R. Plamondon are defining and evaluating a number of relational similarity measures taken between relational vectors representing spatial distances between the reference profile and pairs of test primitives.
- F. Nouboud and M. J. Revillet apply dynamic programming to the envelope of the signature image [37].

- 1997**:** the studies of both offline and online signature verification algorithms were published [27].
- 2000**:** the popular methods of Dynamic Time Warping (DTW), and Hidden Markov Models (HMM) were successfully applied to on-line sig-nature verification, and the search for good global features was significantly advanced.
- 2004**:** the organization of the First International Signature Verification Competition (SVC) [36].

2.4 Another Handwritten Signature Approaches

There are many approaches of handwritten signature recognition, as shown below**:**

1-Hidden Markov Models (HMMS)

HMMS are used to model sequence of observations and their relationship to each other, and is a stochastic approach to pattern recognition.

Fierrez et al. (2007) used a set of time sequences and HMMs. The system is compared to other state-of-the-art systems based on the results of the First International Signature Verification Competition (SVC 2004). Training strategy had shown that training with multi-session data remarkably improves the verification performance and 5 training signatures are enough for obtaining robust models. Verification performance results are 0.74% and 0.05% Equal Error Rate (EER) for skilled and random forgeries [38].

Coetzer et al. HMM-based system is implemented on the randomly selected test signatures, to make a comparison between human and machine performance. The results show that one human verifier, with a FRR of 9.0% and a FAR of 10.5%, performed significantly better than HMM-based system. The EER for HMM-based system is 12.6% [39].

2- Neural Networks (NN)

An NN is a massively parallel computing system that consists of a large number of simple processors with many interconnections. The main characteristic of an NN is that it has the ability to learn complex non-linear input-output relationships, use sequential training procedures, and adapt itself to the data. An NN model attempts to use organizational

principles in a network of weighted directed graphs, in which the nodes are artificial neurons and the directed edges are connections between neuron outputs and neuron inputs.

Horváth et al. aimed to construct an efficient off-line signature analyzer, which can reconstruct the signing method and several hidden features like velocity or strokes, and use these features by a classifier based on a neural network. . The results demonstrated that local features can successfully be used with NN classification systems, to distinguish original signatures from forgeries [40].

Sisodia et al. (2009) evaluated the performance of an Error Back Propagation (EBP) ANN for authentication.

The results show that the system verification rate that stands at 94.27 % may be further improved by rigorous evaluation and a feedback network needs to be considered to limit the possibility of over training of the neural network [41].

3-Genetic Algorithm (GA)

GA is non deterministic methods which apply the rules of selection, mutation and recombination to a population of subjects each of them representing a possible solution to the problem. The acceptability level of each solution is computed according to some optimization criteria that give the best

individuals a higher probability of surviving to the next generation. Following an iterative process a near to optimal solution is reached.

Galbally et al. (2007) used two different GA architectures which were applied to a feature selection problem in on-line signature verification. The standard GA with binary coding was first used to find a suboptimal subset of features that minimizes the verification error rate of the system. The curse of dimensionality phenomenon was further investigated using a GA with integer coding. The results showed that different dimension subspaces were found in which the recognition rate of the system was improved compared to the original 100 dimensional space. Features regarding speed and acceleration information of the signatures were the most suitable for the skilled forgeries scenario, while those dealing with temporal information should be used in the random forgeries case [42].

4-Graph Matching

ABUHAIBA (2007) presented a simple and effective signature verification method that depends only on the raw binary pixel intensities and avoids using complex sets of features. The method looks at the signature verification problem as a graph matching problem.

The results show that equal error rate of 26.7% and 5.6% for skilled and random forgeries, respectively, was achieved at size 32 × 64 pixels. A positive property of this algorithm is that the false acceptance rate of random forgeries vanishes at the point of equal false rejection and skilled forgery false acceptance rates [43].

Chapter three
handwritten signature recognition fusion based system

Offline signature recognition has widespread acceptance by the public so it is our consideration in this system for both verification and identification processes. This chapter includes: database construction, a brief explanation about the proposed system, and system phases in training and testing for verification and identification.

3.1 HWSR1000 Database Construction

The focus of this thesis is off-line signature recognition, so the researcher constructed her own database.

3.1.1 Collecting Database Samples

Database is important to store personal unique features that can be retrieved later for comparison in recognition processes. Our database is constructed based on data collected manually over five months from 100 persons from different age groups. Each person received a form and signed 10 times. Theses samples were scanned using optical scanner on Resolution 100 PPI, and color JPEG format. After that, each sample was segmented to obtain signatures.

3.1.1 Database Tables

Our database contains two tables**:** SIGNATURES table and FEATUREVECTORS Table. SIGNATURES table contain thirteen fields: PERSONID (INT , PK), FIRSTNAME (NVARCHAR (MAX)), LASTNAME (NVARCHAR (MAX)), SIGNATURE1 (NVARCHAR (MAX)), SIGNATURE2 (NVARCHAR (MAX)), SIGNATURE3 (NVARCHAR (MAX)), SIGNATURE4 (NVARCHAR (MAX)), SIGNATURE5 (NVARCHAR (MAX)), SIGNATURE6 (NVARCHAR (MAX)), SIGNATURE7(NVARCHAR (MAX)), SIGNATURE8 (NVARCHAR (MAX)), SIGNATURE9 (NVARCHAR (MAX)), SIGNATURE10 (NVARCHAR (MAX)).

Signature images were not directly stored on the database, but rather as location of signatures images (Uniform Resource Locator (URL)), because storing the database location reduces the size of database greatly as well as simplifying updating or replacing the image instead of massive update/insert/delete in database [44].

FEATUREVECTORS table contains three fields: FEATUREID (INT, PK), TEMPLATE (NVARCHAR), PERSONID (INT, FK), where the researcher just stores the URL of templates exactly as Signatures images. The database was constructed using Microsoft SQL Server 2008. Figure 3.1 illustrates database tables which were used in proposed system.

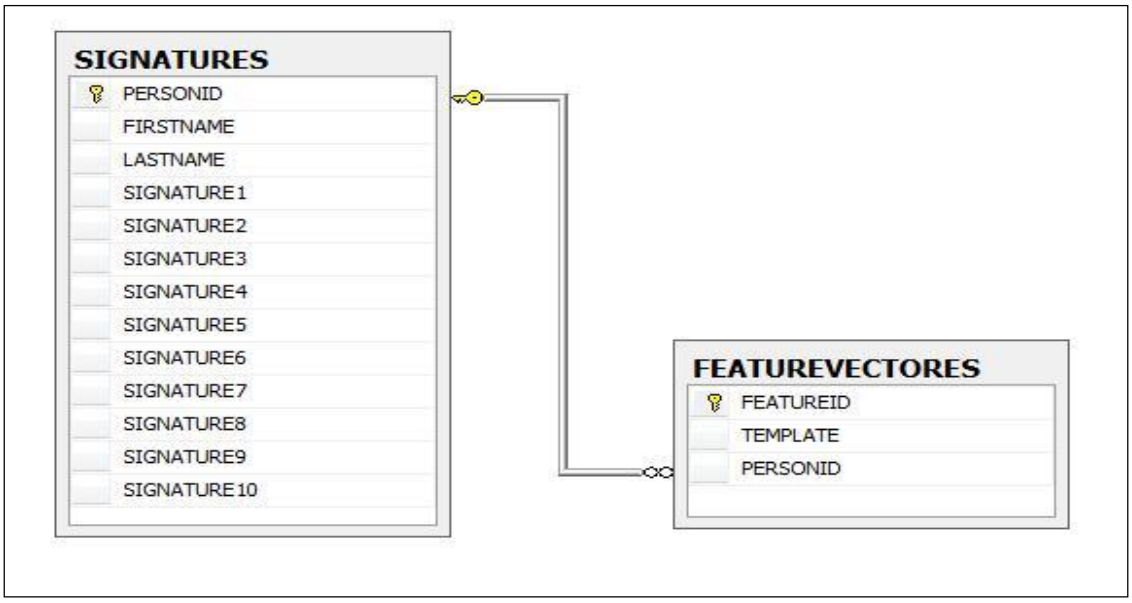


Figure 3.1 Database tables

3.2 Handwritten Signature Recognition Fusion Based System

The developed system of offline signature recognition is based on the concept of DWT, Fusion and SVM. The implementation was done using MATLAB program for signature recognition. The Handwritten Signature Recognition Fusion Based System (HWSRFB) system is carried out in two steps**:** the training phase and the testing phase which include Verification and Identification processes. HWSR1000 database of 1000 signatures was obtained, passed through an optical scanner and stored as a digital form. This static information was used for both training and testing phases. Figure 3.2 illustrates the developed system.

Training phase: this phase is important to initialize the system to be ready for use in testing phase. This phase starts with preprocessing steps for enhancement, then feature extraction processes using DWT and Fusion, in enrollment process after SVM training individual template is stored in HWSR1000 database.

Testing phase**:** this phase includes two processes of recognition: verification and identification. In each process SVM classifier made a decision for classification.

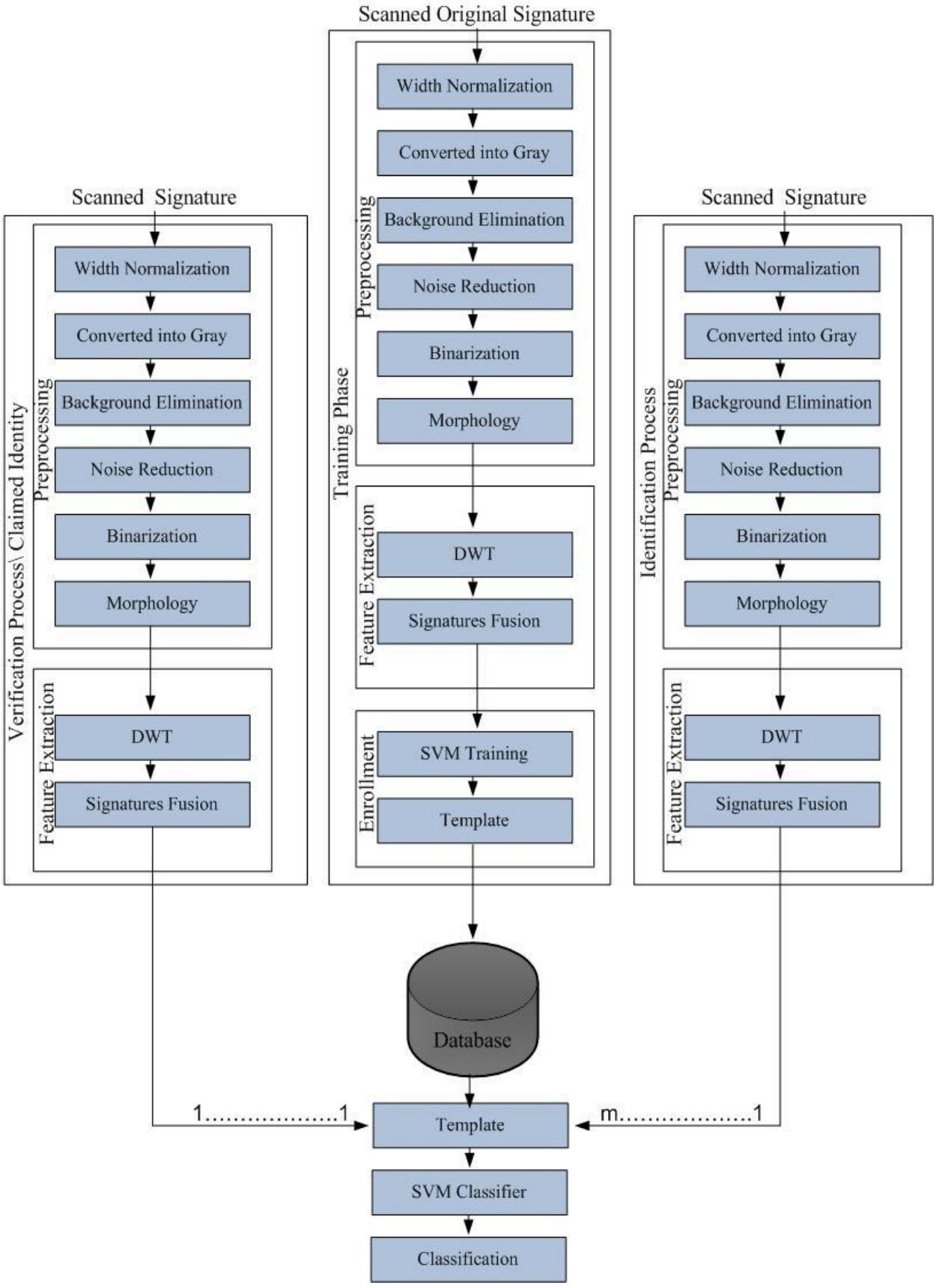


Figure 3.2 The developed systems

3.3 Training Phase

For every person participating in HWSRFB system, it is necessary initially to store feature vectors in the HWSR1000 database, which is considered as the identity for this user, to be used later in verification or identification processes.

This stage can be divided into three steps: preprocessing, feature extraction and enrollment.

3.3.1 Preprocessing

The preprocessing stage includes six steps: Normalization, Converted signatures to gray scale, Background elimination, Noise reduction, Binarization, and Morphology. Figure 3.3 illustrates samples of signatures.

Figure 3.3 Signatures sample

(a) *Normalization (resizing)*:

Signature dimensions may have intrapersonal and interpersonal differences, so the image size is adjusted to a uniform size. To achieve best results and to reduce the number of operations, the signatures must have the same size, which means normalized one. In this approach, the reference dimension is 90×180.

(b) *Converted signatures to gray scale*:

This stage aims to standardize signatures and get them ready for feature extraction. Color has no meaning in the case of signature recognition. The grayscale intensity is stored as an 8-bit integer giving 256 possible different shades of gray from black to white. Figure 3.4 illustrates signatures after conversion to grayscale.

Figure 3.4 Grayscale signatures

(c) *Background elimination*:

In this stage based on Otsu's method or basically gray threshold computed using global threshold (T). Thresholding was chosen to capture signature from the background. The threshold of signature image g(*x*, *y*) is defined as:

$$g(x,y) = \begin{cases} 0, & \text{if } (x,y) \geq T \\ 1, & \text{if } (x,y) < T \end{cases} \qquad (3.1)$$

Pixels(*x*, *y*) of the signature would be (0) and the other pixels(*x*, *y*) which belong to the background would be (1). This process is done by choosing a threshold automatically. MATLAB function called (graythresh) was used to compute a threshold which based on Otsu's method.

(d) *Noise reduction*:

A noise reduction filter is applied to signatures in order to eliminate single black pixels on white background. This noise may be caused by dusty scanner surface. Noise is removed using median filters. The median filter replaces the central value of neighborhood pixels with the value of median pixel; which illustrates in following equation:

$$f(x,y) = median\{I(x,y)\}(3.2)$$

Where *f* (*x, y*) is the output image and {*I* (*x, y*)} is the current pixel. This kind of filtering is useful for removing salt and pepper noise. Figure 3.5 illustrates signatures before filtering with noise, and figure 3.6 illustrates signatures after filtering.

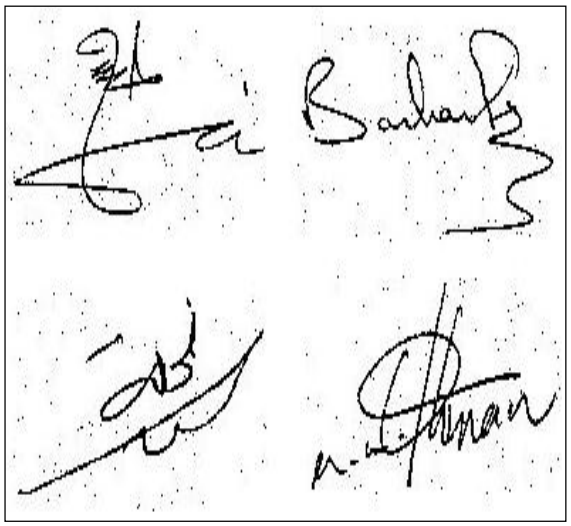


Figure 3.5 Signatures with noise Figure 3.6 Signatures after filtering

(e) *Binarization*:

In this stage, the signatures are converted to binary form (black and white pixels). All pixels of input signature greater than the threshold are replaced with (1) as a white pixel and all other pixels are replaced with (0) as a black pixel, and the threshold e equals 0.5 by default. Working with binary form is more useful than any other form; it is easy to work with 2 bits representation of signatures. Figure 3.7 illustrates signatures after binarization.

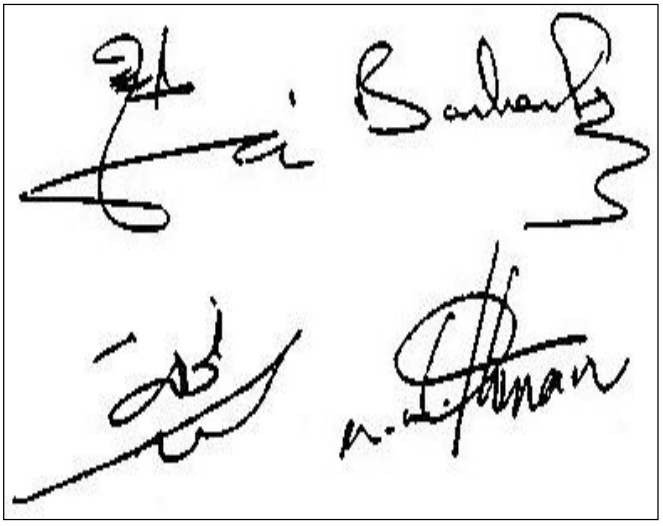


Figure 3.7 Binary signatures

(f) *Morphological operators*:

Mathematical morphology is a tool for extracting image components useful in the representation and description of

region shapes [45]. Morphological operations are used to understand the structure or form of an image. This usually means identifying objects or boundaries within an image. Morphological operations are usually performed on binary images [46]. In this system, an operation based on combinations of dilations and erosions was used. A bridge operation was used to connect pixels separated by single pixel gaps, then remove operation was applied to remove interior pixels keeping the boundaries of signature [47]. Using remove and bridge operations save the greatest amount of information, as they keep the boundary of signatures contrast using thinning operation. Figure 3.8 illustrates morphological signatures.

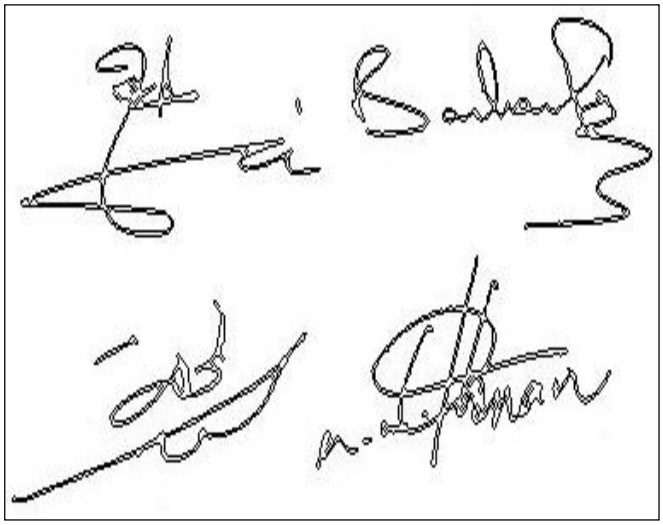


Figure 3.8 Morphological signatures

3.3.2 Feature Extraction

In the feature extraction phase, at first, DWT is applied to preprocessed signatures to obtain sub-images that contain low frequency bands. Then, different signature samples of each person are fused to obtain a pattern of his signatures. The result of the fusion is used as the feature vector.

(a) *Discrete Wavelet Transform*:

In this stage, the researcher used level 4 Mallat transform, where low pass filter 1/2 [1 1] , high pass filter 1/2 [1 -1], and only the fourth level of LL image were used for the analysis, as these contain most of the important texture information. This process was done by using MATLAB Wavelet Toolbox [48].

(b) Signatures Fusion:

This stage aimed to merge several DWT feature vectors for the same person. Each vector has different information, so these different vectors were fused in order to have new and improved information, and to generate the adequate final feature vector for any person. This stage can be achieved by a set of strategies. The most simple is to take the average of different vectors to be merged. Initially the researcher convert several DWT feature from 2-dimension to 1-dimension (row vector) for each person, then pixel-by-

pixel average fusion of those vectors is used as feature vector. Signatures fusion is described in algorithm1.

Algorithm1: Fusion Algorithm

Step 1: Convert several DWT features from 2-D to 1-D for each person.

Step 2: Compute the feature vector as a following:

For (I = 1; I <= length (DWT feature); I++)

{

Feature vector [1, I] = (DWT feature1 [1, I] + DWT feature2 [1, I] + ……._ DWT featureN [1, I]) / Number of input DWT features

}

3.3.3 Enrollment

For every person participating in HWSRFB system, individual distinct characteristic must be stored in HWSR1000 database, after being extracted from signatures samples. This stage involves two steps: SVM training sets and signatures template.

(a) SVM Training Sets:

For each person **(p)** participating in the enrollment process of **(N)** persons, the researcher trains one SVM using the

samples of **(p)** as positive samples and all other samples of the remaining **(N-1)** person as negative samples. As such, entire systems consist of **(N)** SVMs, one for each enrolled person. This process is generated by using MATLAB SVM Toolbox, where a support vector machine classifier is trained using training sets to generate SVMStruct. This Struct contains information about the trained classifier, including the support vectors that are used by SVMClassify in testing phase [49].

(b) Signatures Template:

A biometric template is a digital representation of an individual's distinct characteristics, representing information extracted from a biometric sample. Biometric templates are actually compared in a biometric recognition system [9].

These templates contain distinct characteristics for every participant, and these characteristics are used in recognition stage to measure their similarity to input signature features. In this system, a template was generated as a MATLAB file which had feature vector and SVMStruct for each person [50].

3.4 Testing Phase

Once a SVM is trained, the researcher simply determines on which side of the decision boundary a given test pattern **x** lies and assign the corresponding class label. Testing can be done for different recognition processes. In verification process, claimed signatures are accepted or rejected. In the identification process, the identity of the signature owner will be found.

3.4.1 Verification

In verification processes, a person's claimed signature sample, after preprocessing and feature extraction, is compared to the claim identity template that is retrieved from HWSR1000 database. SVM classifier decides to accept or reject the person's claimed identity. In general, verification is considered a 1 to 1 process.

This process uses MATLAB SVM Toolbox, where the researcher retrieves SVMStruct from the claimed person template which was created in the training phase. For testing, the SVM classifier takes this SVMStruct and claimed signature and compare them to make a decision [51]. This process is demonstrated in Figure 3.9.

3.4.2 Identification

In identification processes, a person's signature sample, after preprocessing and feature extraction, is presented to all templates stored in the HWSR1000 database. In each iteration, SVM classifier makes decision to accept or reject a person's sample. If no SVM accepts the sample or more than one SVM accepts the sample, the person is rejected. If only one SVM accepts the sample, the user is identified as the corresponding person. In general, identification is considered a 1 to m process.

This process is generated by using MATLAB SVM Toolbox. In each iteration; the researcher retrieves SVMStruct from different templates, then SVM classifier takes this SVMStruct and input signature and compare them to make a decision. If the SVM classifier accepts the sample the iteration stops, else the researcher takes a different SVMStruct from another template until the SVM classifier makes an accept decision or the researcher iterates all templates in HWSR1000 database [51] .This process is demonstrated in Figure 3.10.

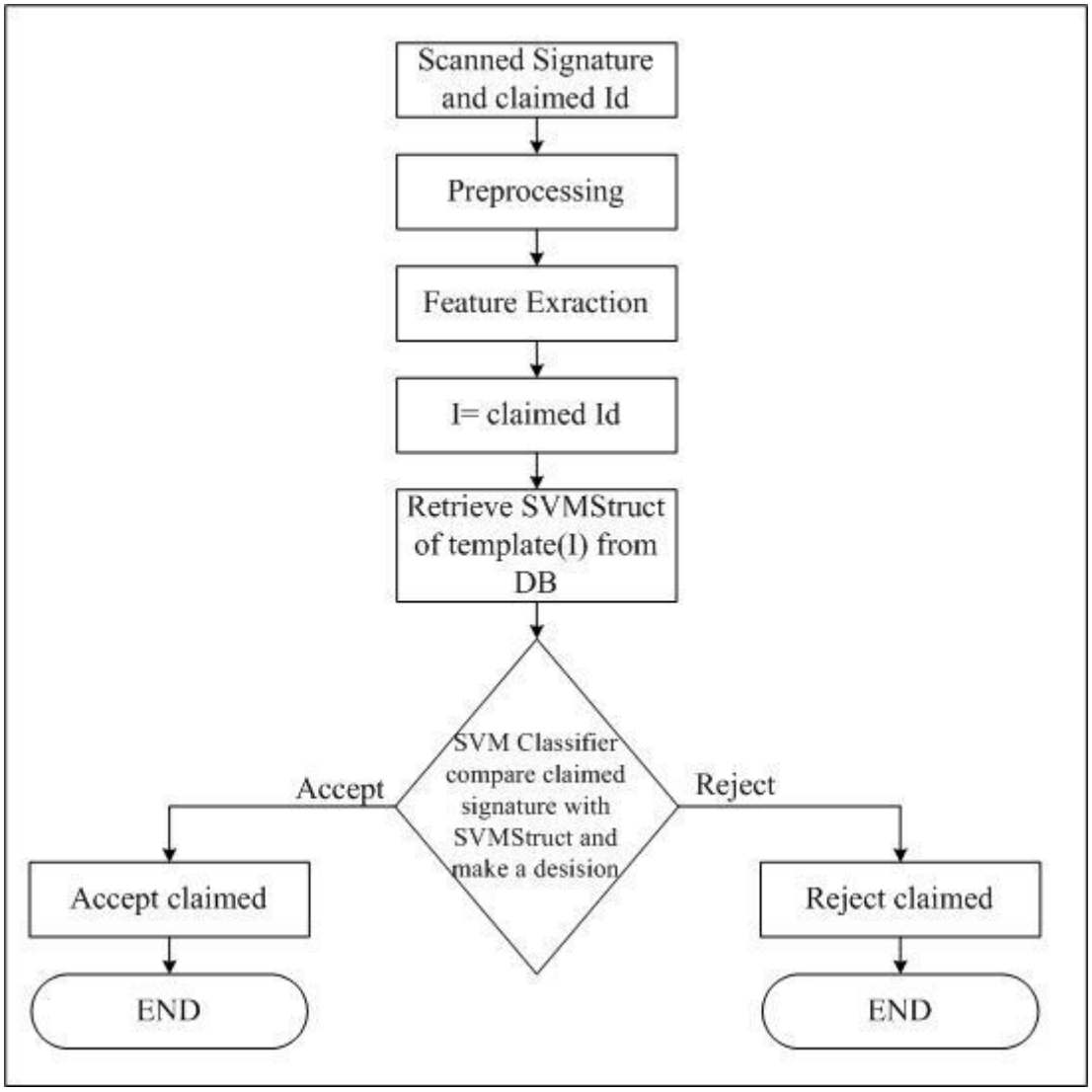


Figure 3.9 Verification process

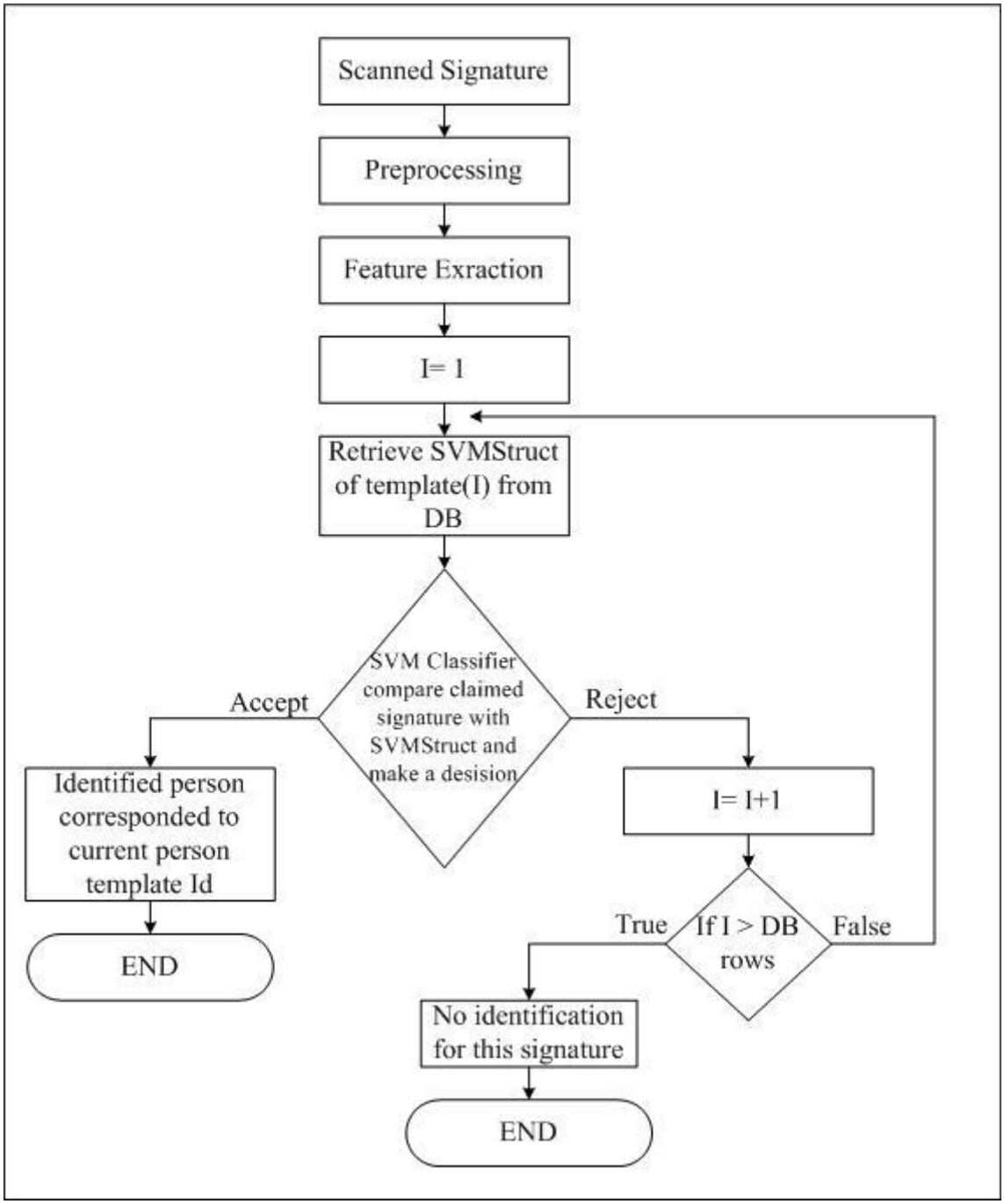


Figure 3.10 Identification process

Chapter four
Results and discussion

Results obtained from the developed system are presented in this chapter. Signatures kinds, percentage of error for identification and verification and their related curves are also discussed here.

4.1 Experimental Results

A database of 1000 Arabic and English signatures was collected from 100 persons, where each person signed 10 times. Genuine signatures and random forgeries were used, because obtaining actual forgeries is difficult.

Categorizations of forged signatures are not standard. Forgeries types are listed as following:

(1) Random Forgeries: the signatures are signed without knowledge about the name and genuine signature of the owner.

(2) Simple Forgeries: define the signatures where the name of signature owner is known.

(3) Skilled Forgeries: the aim is to make an almost exact copy of the genuine signature by using an existing sample [52].

Initially, 10 signatures for each person were scanned, and then several signatures samples of each person were taken and introduced to the next step. Samples of Arabic and English signatures are illustrated in Figures 4.1 and 4.2 respectively.

Figure 4.1 Samples of Arabic signatures

The first step is signatures enhancement by six preprocessing stages (Normalization, Converted signatures to gray scale, Background elimination, Noise reduction, Binarization, and Morphology).

After that the feature extracting stage will be started, where each enhanced signature is processed by 4 levels of DWT decomposition. Then, signatures sample are converted to row vector and average fusion made, where the result consider as feature vector. In the enrollment process, SVM training and signatures template are used to specify individual distinct characteristic'. Finally, samples are stored in the HWSR1000 database.

Figure 4.2 Samples of English signatures

In the testing phase, SVM classifier is used to make decisions. A template of several signatures fusion is retrieved to be used for comparison with the signatures samples are constructed from.

Results are affected with signatures quality obtained for each person. As such, signatures can be classified into two kinds**:** good signatures and bad signatures. Samples of good and bad signatures are illustrated in Figures 4.3 and 4.4.

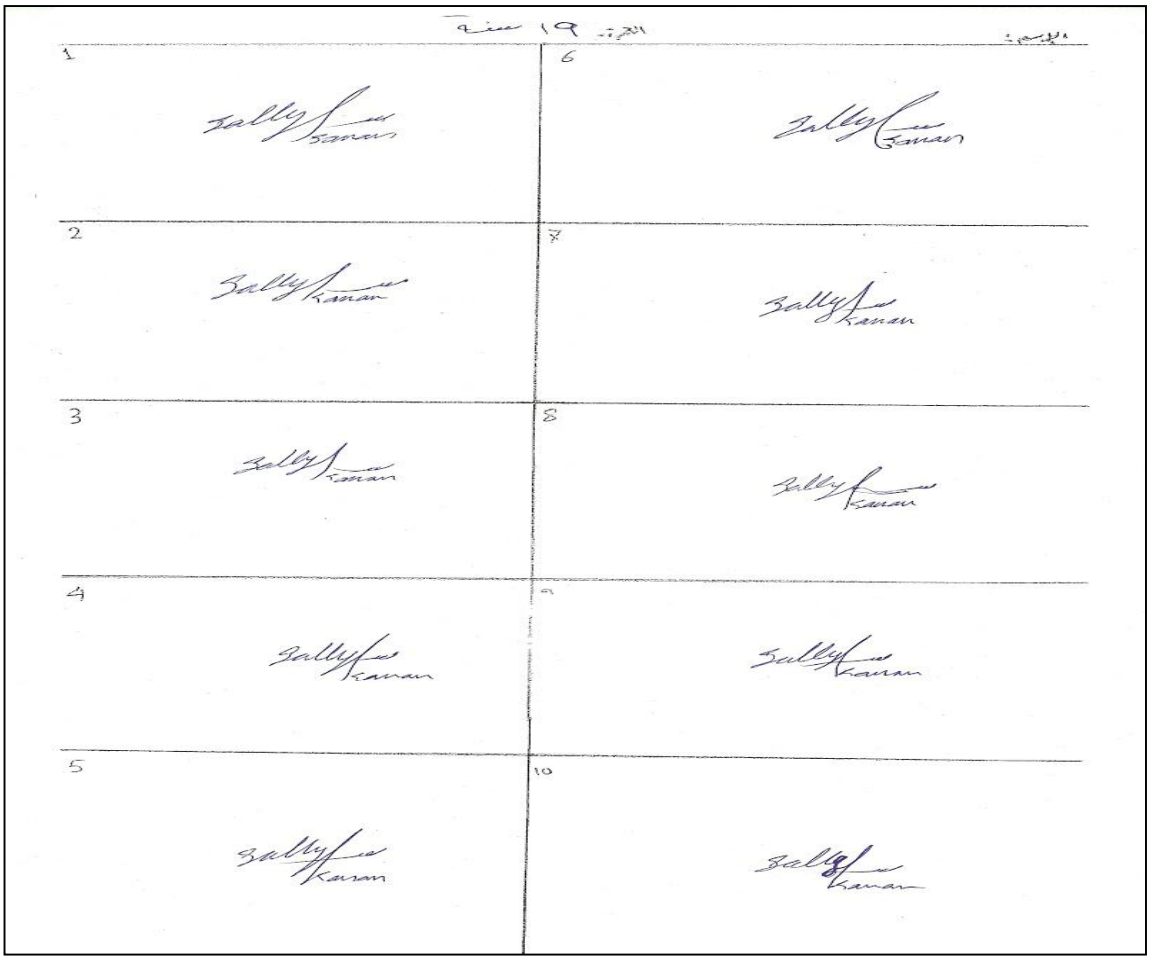


Figure 4.3 Samples of good signatures

Good signatures give best results in verification and identification processe, and increase the performance of the system.

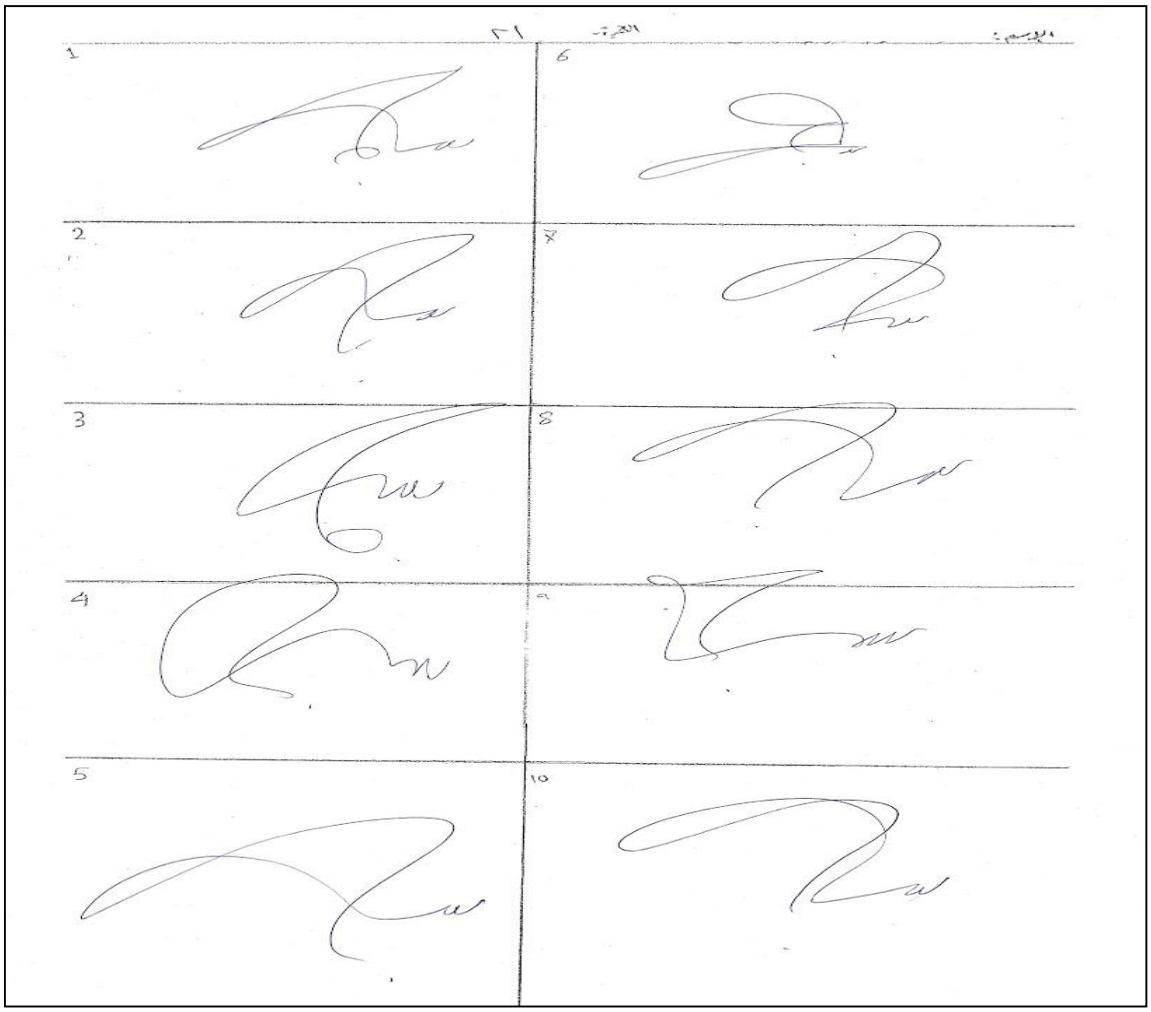


Figure 4.4 Samples of bad signatures

Bad signatures, on the other hand, reduce the performance of verification and identification processes, and so the effectiveness of the system decreases.

The results obtained are as the following:

Table 4.1 shows the verification error for 3, 4 and 5 signatures fusion, which are 2.33%, 4.00% and 5.40% respectively. And the best verification percentage is 97.67% for 3 fusions.

Table 4.2 shows the identification error for 3, 4 and 5 signatures fusion. These are 6.00%, 9.25% and 14.40% respectively. And the best identification percentage is 94.00% for 3 fusions.

4.2 Identification Process

Table 4.1 Verification Process

	% of verification error	% of verification
3 fusion	2.33	97.67
4 fusion	4.00	96.00
5fusion	5.40	94.60

Table 4.2 Identification Process

	% of identification error	% of identification
3 fusion	6.00	94.00
4 fusion	9.25	90.75
5fusion	14.40	85.60

Figure 4.5 illustrates verification and identification error, and Figure 4.6 illustrates verification and identification percentage.

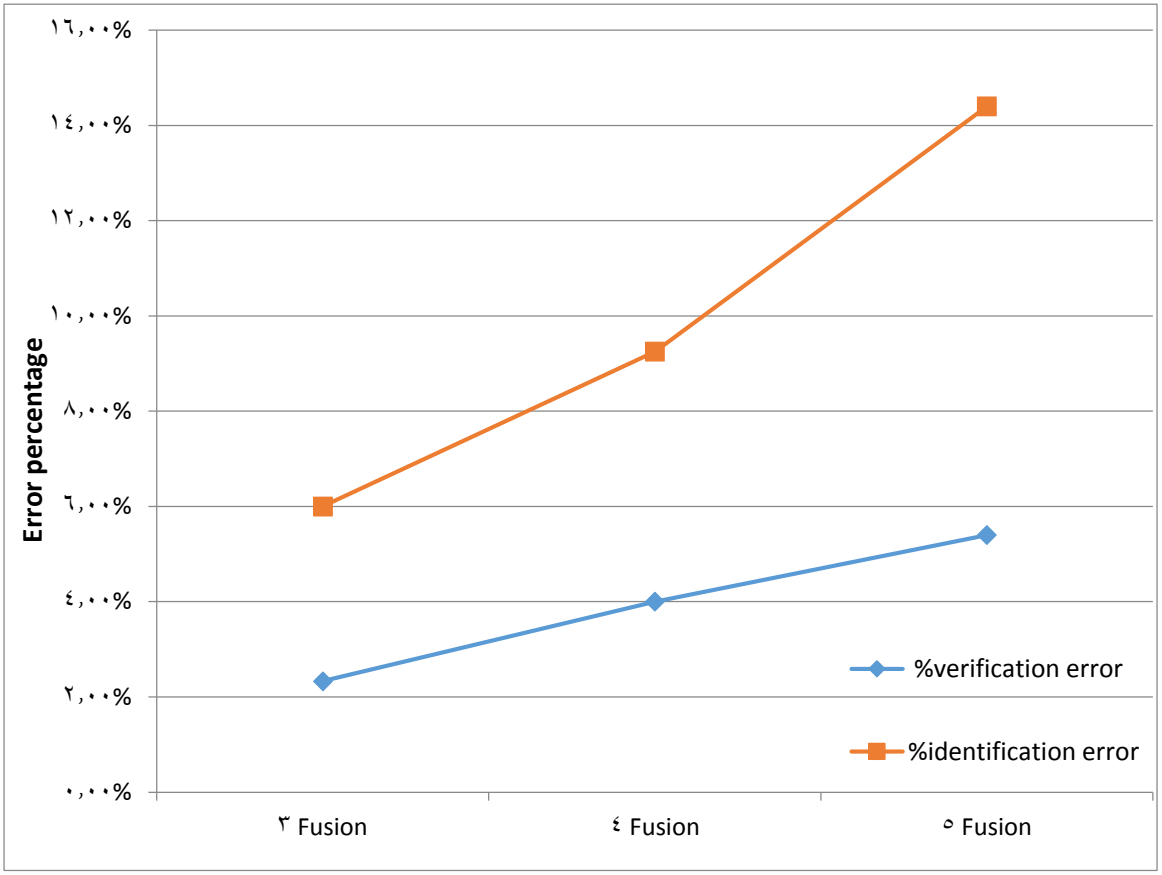


Figure 4.5 Verification and identification error

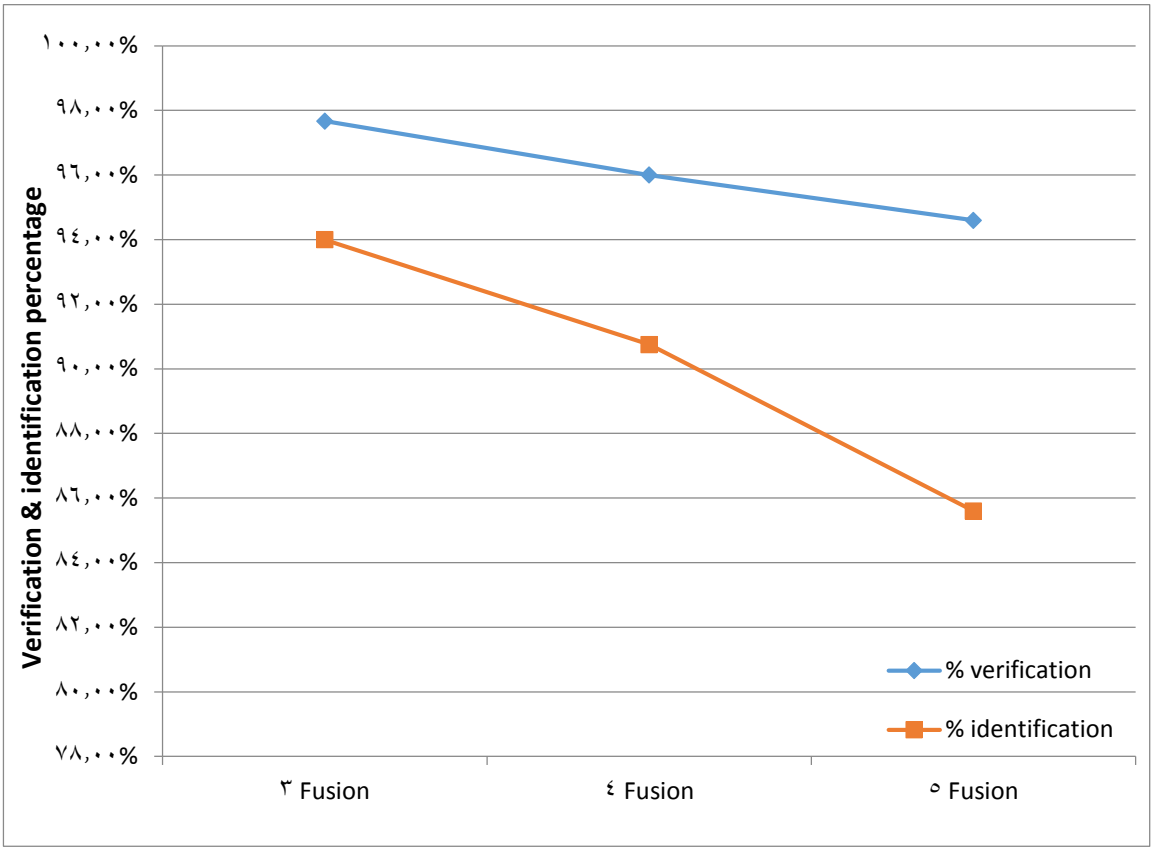


Figure 4.6 Verification and identification percentage

Based on these results, the researcher can deduce the reasons that lead to them. When a person is asked to sign for several times, his focus will decrease, so this person will start signing differently. Therefore, the increased signatures samples are used in fusion process comparison with the signatures samples are constructed from will lead to an increase in the error rate.

Signatures Recognition results are reported in three terms. False Acceptance Rate (FAR), which means a forgery signature, is considered as a genuine signature.

False Rejection Rate (FRR), which means a genuine signature is considered as a forgery signature, and average error rate is the average of FAR and FRR [28].

10 genuine signatures and 20 random forged signatures were selected randomly from the HWSR1000 database to test every participant. Random forgeries selected from different ages are: (under 20-29) (30-39) (40-49) (50-more).

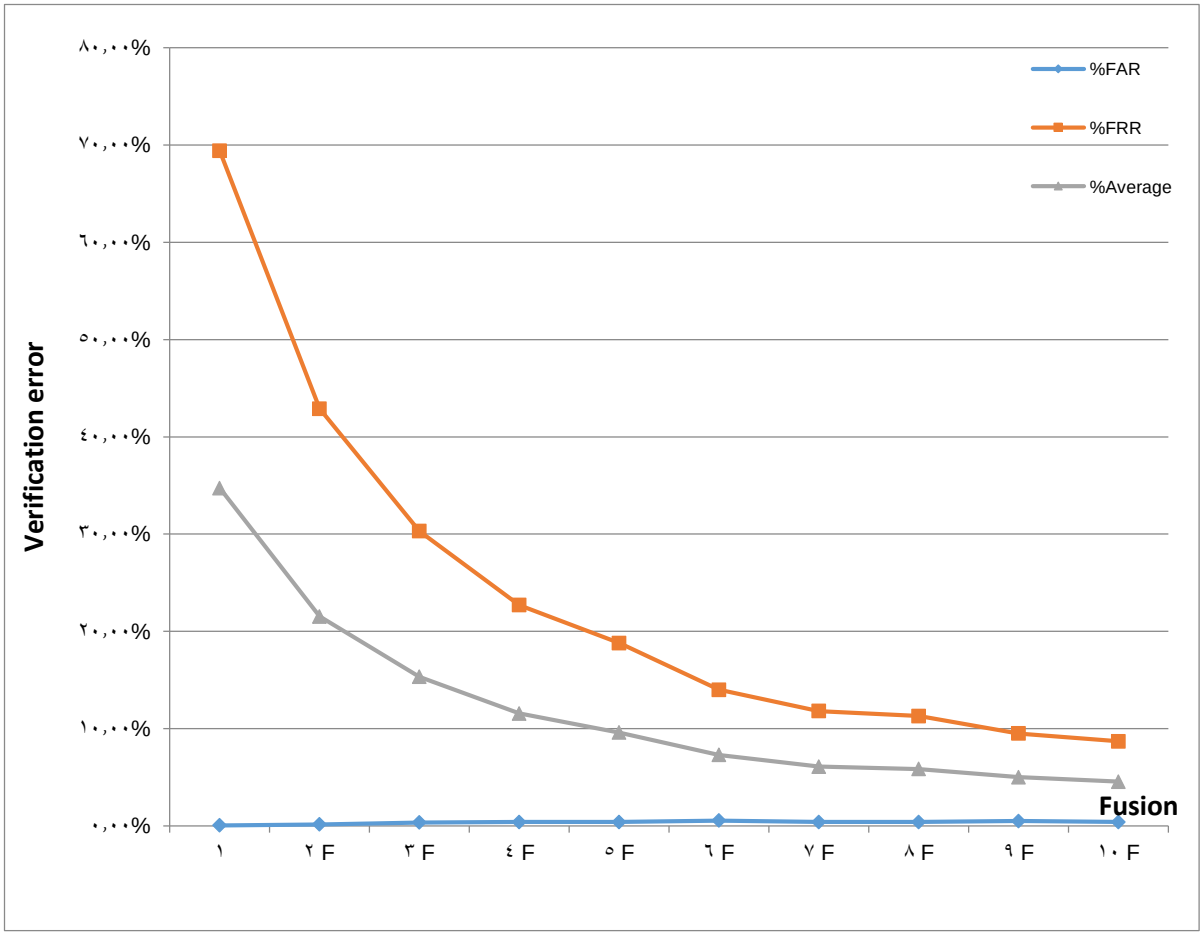
Table 4.3 and 4.4 show the performance of the HWSRFB system, by comparing every fusion level with 10 genuine signatures and 20 random forgeries in the verification phase and 10 genuine signatures in the identification phase for each person.Table 4.3 shows the average rate for verification from 1 to 10 signatures fusion and the best is 4.55% for 10 signatures fusion, with verification rate 96.83%. Figure 4.7 illustrates verification error and Figure 4.8 illustrates verification percentage.

Table 4.4 shows the percentage of error for identification from 1 to 10 signatures fusion and the best is 27.50% for 9 signatures fusion, with identification rate 72.50%. Figure 4.9 illustrates identification error and Figure 4.10 illustrates identification percentage.

Table 4.3 Verification phase

	FAR	FRR	Average	% of verification
1 signature	0.05	69.40	34.73	76.83
2 Fusion	0.15	42.90	21.53	85.60
3 Fusion	0.35	30.30	15.33	89.97
4 Fusion	0.40	22.70	11.55	92.17
5 Fusion	0.40	18.80	9.60	93.46
6 Fusion	0.55	14.00	7.30	94.96
7 Fusion	0.40	11.80	6.10	95.80
8 Fusion	0.40	11.30	5.85	95.97
9 Fusion	0.50	9.50	5.00	96.50

10 Fusion	0.40	8.70	4.55	96.83



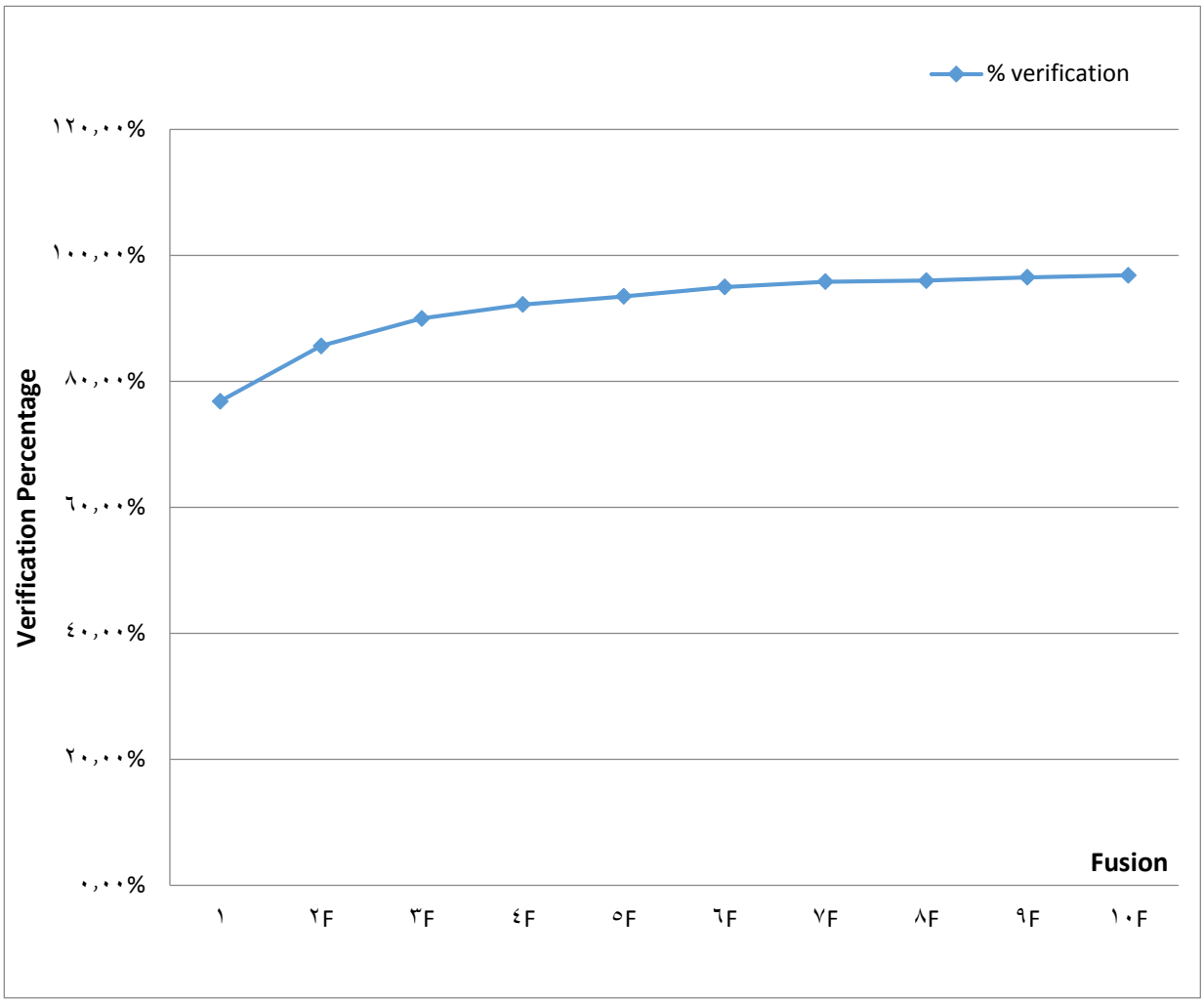


Figure 4.8 Verification percentage

Table 4.4 Identification phase

	% of error identification	% of identification
1 signature	69.90	30.10
2 Fusion	47.90	52.10
3 Fusion	37.80	62.20
4 Fusion	33.20	66.80
5 Fusion	29.460	70.40
6 Fusion	28.30	71.70
7 Fusion	28.90	71.10
8 Fusion	28.70	71.30
9 Fusion	27.50	72.50
10 Fusion	29.00	71.00

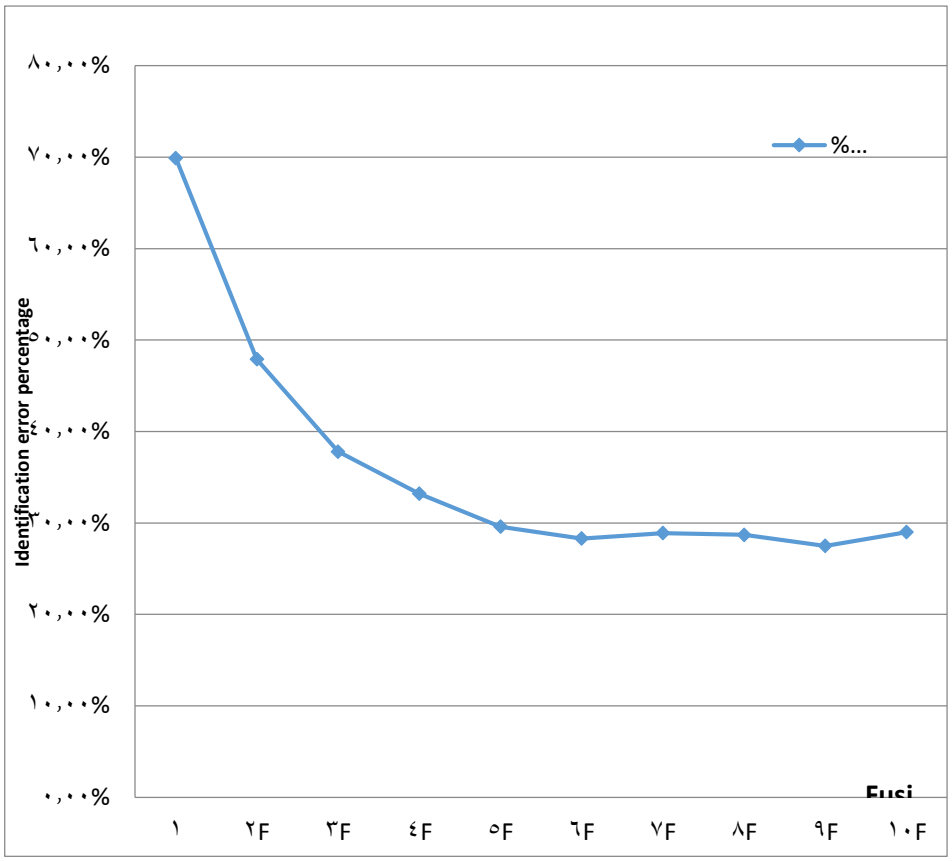


Figure 4.9 Identification error

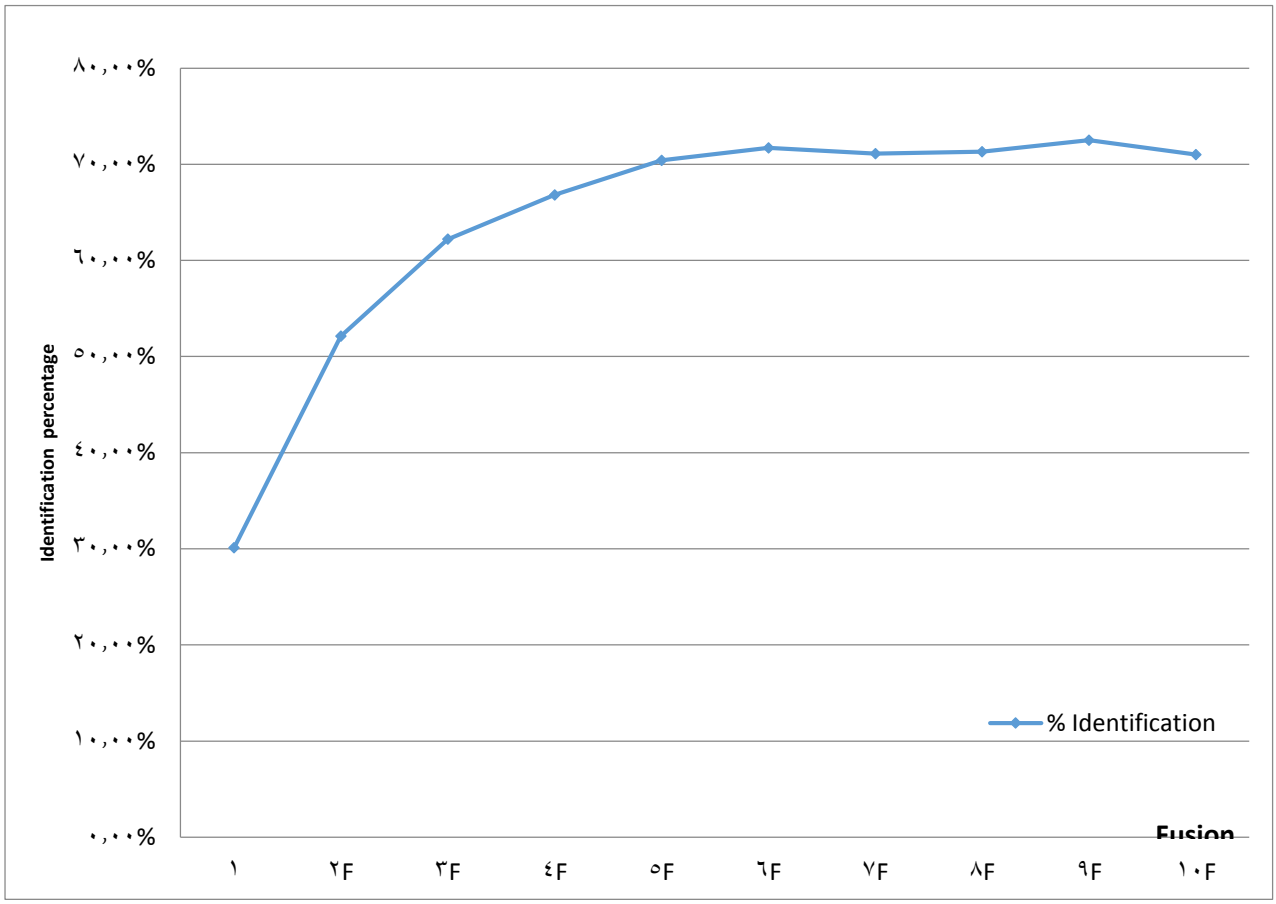


Figure 4.10 Identification percentage

The results show, in general, that when the researcher increased the number of signatures sample fusion, the performance of the system on verification or identification processes would be better. In general results can be considered good compared to some previous methods, show in table 4.5.

Table 4.5 Previous methods

Methods	FAR	FRR
Signature Registration and Fusion [23]	7.25	11.1
Local Radon Transform and SVM [19]	2 causal 22 skilled	19
MDF Conjunction with NN and SVM [17]	0.16	17.78
SVM with Radial Basic Function [16]	0.11	0.02

Chapter five
Conclusion and recommended future tasks

5.1 Conclusions

In this thesis, the researcher was mainly concerned with the development of a reliable, accurate and quick electronic offline signatures recognition system. To reach that, an approach based on DWT, vectors fusion and SVM concepts was presented. Two phases are employed in HWSRFB system: training and testing. And the results illustrated the following:

1-The increase in signatures samples used in fusions compared to signatures samples constructed from would increase error rate.

2- The increase the number of signatures fusion increases the HWSRFB system performance and decrease the error percentage for verification and identification.

3- In verification process, a 10 signatures fusion presented the best result with verification rate 96.83%, and in identification process 9 signatures fusion presented the best result with identification rate 72.50%.

4- The performance of verification was better than identification on all the levels of the fusion test.

5- The obtained results are good compared with existing algorithms.

5.2 Recommended Future Tasks

Future work will be mainly focused on:

1- Modify this approach for online signatures recognition instead of offline signatures, in hope to present a more accurate and reliable system.

2-Implement and improve other approaches for offline signatures recognition.

3- Implement many techniques for personal identification and verification.

References

[1] Oh Jong (2001). *An On-Line Handwriting Recognizer with Fisher Matching, Hypotheses Propagation Network and Context Constraint Models.* New York University.

[2] Impedovo Donato, et al (2008). *Handwritten Signature and Speech: Preliminary Experiments on Multiple Source and Classifiers for Personal Identity Verification*. Springer-Verlag Berlin Heidelberg, pp. 181–191.

[3] Forn´es Alicia, et al (2007). *Handwritten Symbol Recognition by a Boosted Blurred Shape Model with Error Correction*. Springer-Verlag Berlin Heidelberg, pp. 13–21.

[4] Dimauro G, et al. *Handwriting Recognition: State of the Art and Future Trends*. Dipartimento di Informatica, University degli Studi di Bad, Italy.

[5] Al-Mahadeen Bassam, et al (2010). *Signature Region of Interest using Auto cropping*. IJCSI International of Computer Science Issues, Vol. 7, Issue 2, No. 4.

[6] Henniger Olaf, et al (2009). *Signature Recognition*. Encyclopedia of Biometrics, Springer Science&Business Media.

[7] Mangal Sachin (2006). *Personal Identification Based on Handwriting*. Indian Institute of Technology, Kanpur.

[8] authentec (2009). *Biometrics Overview*. Retrieved on 25/10/2010 from: http://www.authentec.com/technology-biometrics-overview.cfm

[9] National Science & Technology Council (NSTC), Committee on Technology (COT), Committee on Honmeland and National Security, Subcommittee on Biometrics. *Biometrics "Foundation Documents"*. http://www.biometricscatalog.org/Introduction/Default.aspx?sindex=0

[10] Advanced Matlab (2009). *Off-Line Signature Recognition* .Retrieved on 14/9/2010 from: http://www.advancedmcode.org/off-line-signature-recognition.html

[11] Biometric System Laboratory. *Signature*. Retrieved on 25/10/2010 from: http://biolab.csr.unibo.it/research.asp?organize=Activities&select=&selObj=444&pathSubj=444&Req=&

[12] National Check Fraud Center. *Check Fraud Statistics*. Retrieved on 25/10/2010 from: http://www.ckfraud.org/statistics.html

[13] American Bankers Association (2007). *2007 Deposit Account Fraud Survey Report* .Retrieved on 25/10/2010 from: http://www.aba.com/Surveys+and+Statistics/SS_Depositfraud.htm

[14] THE WAVELET TUTORIAL. *MULTIRESOLUTION ANALYSIS: THE DISCRETE WAVELET TRANSFORM*. Retrieved on 25/10/2010 from: http://users.rowan.edu/~polikar/WAVELETS/WTpart4.html

[15] Thaiyalnayaki K (2010). *Finger Print Recognition using Discrete Wavelet Transform*. 2010 International Journal of Computer Applications,Vol. 1 , No. 24

[16] Pajares Gonzalo (2004). *A wavelet-based image fusion tutorial*. The Journal of the Pattern Recognition Society, pp.1855-1872.

[17] Lehigh University. *Investigations of Image Fusion.* Retrieved on 20/2/2011 from: http://www.ece.lehigh.edu/SPCRL/IF/image_fusion.htm#If_introduction

[18] Wei Ji, et al (2005). *Signature Verification Using Wavelet Transform and Support Vector Machine*. Springer-Verlag Berlin Heidelberg, pp.671-678

[19] Justino Edson, et al. (2004). *A comparison of SVM and HMM classifiers in the Offline Signature Verification*. Elsevier B.V. Pattern Recognition Letters, pp.1377-1385. http://www.sciencedirect.com

[20] Shubhangi D.C. et al. (2009). **Handwritten English Character And Digit Recognition Using Multiclass SVM Classifier And Using Structural Micro Features.** ***International Journal of Recent Trends in Engineering, Vol.2, No.2, pp.193-195.***

[21] Özgündüz Emre, et al (2005). *Offline Signature Verification and Recognition by Support Vector Machine*. Eusipco-2005, 4-8 September, 2005, Antalya, Turkey, pp.113-116.

http://www.eurasip.org/Proceedings/Eusipco/Eusipco2005/defevent/papers/cr2010.pdf

[22] Naguyen Vu, et al (2007). *Offline Signature Verification Using Enhanced Modified Direction Features in Conjunction with Neural Classifiers and Support Vector Machines*. Proc. 9th Int Conf on document analysis and recognition, Vol. 2, pp. 734-738.

[23] Kisku Dakshina, et al. *Fusion of Multiple Matchers Using SVM for Offline Signature Identification*. Security Technology, Vol. 58 (2009), pp. 201-208.

[24] Kiani Vahid, et al (2009). *Offline Signature Verification Using Local Radon Transform and Support Vector Machines*. International Journal of Image Processing, Vol.3, pp.184-194.

[25] Jawarkar NP, et al. (2009). *Signature Verification using DWT and Neural Network.* IE(I) Journal-CP, Vol.90, pp.41-46.

[26] Ali A.et al (2009). *Offline Signature Verification Using Radon Transform and SVM/KNN Classifiers*. Vladimir State University.

[27] Fauziyah S, et al (2009). *Signature Verification System Using Support Vector Machine*. MASAUM Journal of Basic and Applied Sciences, Vol.1, No.2, pp.291-294.

[28] Ghandali Samaneh, Moghaddam Mohsen (2009). *Off-line Persian Signature Identification and Verification Based on Image Registration and Fusion*. JOURNAL OF MALTIMEDIA Vol. 1, No. 2.

[29] Efford Nick (2000). *Digital image Processing, A Practical Introduction Using Java*. PEARSON Education Limited, Addison Wesley.

[30] University of Colorado at Boulder (2000). *Introduction to image processing in Matlab 1*. Retrieved on 03/11/2010 from: http://amath.colorado.edu/courses/5720/2000Spr/Labs/Worksheets/Matlab_tutorial/matlabimpr.html

[31] Rao K.M.M. *OVERVIEW OF IMAGE PROCESSING*. Deputy Director, National Remote Sensing Agency, Hyderabad, India.

[32] infibeam com. *Image Processing and Pattern Recognition: Fundamentals and Techniques Book Description*. Retrieved on 03/11/2010 from:

http://www.infibeam.com/Books/info/f-y-shih/image-processing-pattern-recognition-fundamentals-techniques/9780470404614.html

[33] Wapedia. *Handwriting recognition*. Retrieved on 03/11/2010 from: http://wapedia.mobi/en/Handwriting_recognition

[34] Burrow Peter (2004). *Arabic Handwriting Recognition*. School of Informatics, University of Edinburgh.

[35] infotivity. *An overview of Natural Handwriting Recognition (NHR)*, by C. Edward Rawson. Retrieved on 26/11/2010 from: http://www.infotivity.com/hwr.htm

[36] Fierrez Julian, Ortega-Garcia Javier. *On-Line Signature Verification*. Handbook of Biometrics, pp.190-209.

[37] LECLERC FRANCK, PLAMONDON REJEAN (1993). *AUTOMATIC SIGNATURE VERIFICATION: THE STATE OF THE ART—1989-1993*. Laboraioire Scribens, Ecole Polytechnique de Montreal, Canada.

[38] Fierrez Julian et al. (2007). *HMM-Based On-Line Signature Verification: Feature Extraction and Signature Modeling.* Preprint submitted to Elsevier. Vol. 18, No. 16, pp.2325-2334.

[39] Coetzer J el al. *Off-Line Signature Verification: A Comparison between Human and Machine Performance.* University of Stellenbosch, South Africa.

[40] Horváth Ádám et al. *Usability of Neural Networks in Off-line Signature Verification.* Magyar Kutatók 10. Nemzetközi Szimpóziuma. 10th International Symposium of Hungarian Researchers on Computational Intelligence and Informatics. pp.207-215

[41] Sisodia Kshitij et al. (2009). *Off-line Handwritten Signature Verification using Artificial Neural Network Classifier.* International Journal of Recent Trends in Engineering, Vol 2, No. 2, pp.205-207.

[42] Galbally Javier et al. (2007). *Feature Selection Based on Genetic Algorithms for On-Line Signature Verification.* Biometric Recognition Group–ATVS, EPS, Universidad Autonoma de Madrid, Spain.

[43] ABUHAIBA Ibrahim. (2007). *Offline Signature Verification Using Graph Matching.* Turk J Elec Engin, VOL.15, N0.1, pp.89-104.

[44] Journey to SQL Authority with Pinal Dave (2007). **SQL SERVER – Do Not Store Images in Database – Store Location of Images (URL)*. Retrieved on 9/11/2010 from: http://blog.sqlauthority.com/2007/12/13/sql-server-do-not-store-images-in-database-store-location-of-images-url/***

[45] P. Peterlin, *Morphological Operations: An Overview*. Retrieved on 14/9/2010 from: http://www.inf.u-szeged.hu/~ssip/1996/morpho/morphology.html

[46] ENDS489 Course Notes - Fall 2000. *Morphological Operations*. Retrieved on 25/10/2010 from: http://www.viz.tamu.edu/faculty/parke/ends489f00/notes/sec1_9.html

[47] Gonzalez R.C., et al. ***Digital Image Processing*** *using MATLAB*, Addison-Wesley.

[48]MATLAB; Help: Wavelet Toolbox, 2-D Discrete Wavelets. http://www.mathworks.com/help/toolbox/wavelet/ref/dwt2.html

[49] MATLAB; Help, svmtrain: Functions (Bioformatics Toolbox). http://www.mathworks.com/help/toolbox/bioinfo/ref/svmtrain.html

[50]MATLAB; Help, save: Functions (MATLAB Function Reference). http://www.mathworks.com/help/techdoc/ref/save.html?BB=1

[51] MATLAB; Help, svmclassify: Functions (Bioformatics Toolbox). http://www.mathworks.com/help/toolbox/bioinfo/ref/svmclassify.html

[52**]** Özgündüz Emre, et al (2005). *Efficient Offline Verification and Identification of Signatures by Multiclass Support Vector Machine*. Springer-Verlag Berlin Heidelberg, pp.799-805.

طريقة محسنة لتمييز التوقيع اليدوي

إعداد

مها محمود الصعيدي

200810408

إشراف

الدكتور مزهر العاني

قدمت هذه الرسالة لاستكمال متطلبات الحصول على درجة الماجستير في

علم الحاسوب

قسم علم الحاسوب

كلية العلوم الحاسوبية والمعلوماتية

جامعة عمان العربية

شباط، 2011