Deanship of Graduate Studies

Al-Quds University

A Colour Code Three Dimensional Euclidean Space for Static Signature Verification

Majdoleen Sameer Abu Taqa

M.Sc. Thesis

Jerusalem - Palestine

1436 - 2014

A Colour Code Three Dimensional Euclidean Space

for Static Signature Verification

Prepared By:

Majdoleen Sameer Abu Taqa

B.Sc. Computer Engineering, 2010

An-Najah National University, Palestine

First Supervisor: Dr. Hanna Abdel Nour

Second Supervisor: Dr. Labib Arafeh

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Electronic and Computer Engineering/Department of Electronic and Computer Engineering/ Faculty of Engineering/ Graduate Studies.

Jerusalem - Palestine

1436 - 2014

Al-Quds University
Deanship of Graduate Studies
Master of Electronics and Computer Engineering

Thesis Approval

A Colour Code Three Dimensional Euclidean Space for Static Signature Verification

Prepared By: Majdoleen Sameer Abu Taqa

Registration No: 21010494

Supervisors: Dr. Hanna Abdel Nour and Dr.Labib Arafeh

Master thesis submitted and accepted, Date: / /2014

The name and signatures of examining committee members are as follows:

Head of committee	Dr. Labib Arafeh	Signature:………………………..
Committee member	Dr. Hanna Abdel Nour	Signature:………………………..
Internal examiner	Dr. Salaheddin Odeh	Signature:………………………..
External examiner	Dr. Radwan Tahboob	Signature:………………………..

Jerusalem - Palestine

1436 - 2014

Declaration:

I hereby certify that this thesis submitted for the degree of Master in Science, is the result of my own research, except where otherwise acknowledged, and that this study (or any part of the same) has not been submitted for a higher degree to any other university or institution.

Signed: …………………..

Majdoleen Sameer Abu Taqa

Date: …………….

ACKNOWLEDGMENT

After thanking God who honored me with this thesis I would like to thank my family for all of their love, my father, my mother the angel, who was the secret behind my success, and my mother in law for her tremendous support.

My tender thanks to my husband, Amer, who has been a constant source of support and encouragement at all times. I am truly thankful for having you in my life. My daughter, Lujain, my source of unending joy and love and who gave me the strength to go on and to be a good model for her.

I am grateful for my supervisors Dr. Hanna Abdel Nour and Dr.Labib Arafeh for devoting continuous guidance, support, help and time throughout my thesis work, and for the opportunities they have afforded me.

I would like to express my sincere gratitude to my friends, Asmaa Bshara and Kawthar Halabia, for their encouragement during my master endeavor.

Finally many thanks and gratitude goes to Prof. Mohamad Hoseyn Sigari who shared his signature database with me.

Abstract

Biometric technologies measure and analyze the human body characteristics and use an individual's unique biological traits to determine one's identity. The technologies behind biometrics are still emerging, and they provide a lot of security benefits across the spectrum. In this dissertation and in an effort to authenticate the identity of an individual through his/her handwritten signature, a novel off-line signature verification system that tries to mimic the way a person compares signatures is proposed, whereas the strength of the system is represented in its reliance on mathematical modeling and the use of an algebraic formalism in order to solve the problem. After capturing the image of the signature, a set of pre-processing steps are applied. Followed by the implementation of a colour code algorithm coupled to an overlapping (XORing) operation of every two pre-processed signatures under comparison, which results in generating a check pattern. Depending on the degree of similarity and dissimilarity between the two overlapped signatures different colours with different values appear in the generated pattern. Analyzing the generated pattern results in getting a set of features, which are the counts of coloured pixels. Those features are used in the process of verification. The verification is performed by representing all the points related to the genuine and forged signatures in a colour code three dimensional Euclidean space, and then encapsulating those points using spheres. This three dimensional parametric model is compared to the Artificial Neural Networks (ANNs) and the Adaptive Neuro-Fuzzy Inference System (ANFIS) using the same set of features every time. This former novel method outperforms the other ones in terms of FAR (False Acceptance Rate), FRR (False Rejection Rate) and CCR (Correct Classification Rate).

الملخص

تعمل التقنيات البيومترية على قياس و تحليل خصائص الجسم البشري و استخدام صفات الفرد البيولوجية الفريدة في تحديد هويته. لا تزال التقنيات الخاصة بهذه القياسات الحيوية في الظهور، وذلك لتوفيرها عدد من المزايا الأمنية وغيرها الكثير. في هذه الأطروحة ومن أجل التحقق من هوية شخص ما استنادا إلى توقيع خطي، نقترح هنا نظاما مستحدثا للتحقق من التوقيع المكتمل (بمعنى غير المباشر) يحاكي طريقة الإنسان في مقاربته للتواقيع. تكمن قدرة هذا النظام بأنها تعتمد على أنموذج رياضي واستخدام الشكليات الجبرية لحل المسألة. بعد تصوير التوقيع يتم على التوالي تنفيذ خطوات أولية هي تعديل حجم الصورة، إحالتها إلى هيكل يوجز الخطوط، تطبيق عدة مستويات من الاتساع، الطرح، التعبئة بألوان، والجمع. يتبع كل هذا تطبيق خوارزميات تشفير بالألوان مقرون بعملية تطابق لكل توقيعين معالجين مسبقا وخاضعين للمقارنة مما يؤدي إلى إنشاء نمط فحص. استنادا إلى درجة التشابه أو الاختلاف ما بين التوقيعين قيد المقارنة يظهر في النمط الناتج ألوان مختلفة مقرونة بعدة قيم. يفضي تحليل النمط الناتج إلى مجموعة من الخصائص هي تعدادات الوحدات الملونة. تستخدم هذه الخصائص في عملية التحقق من التوقيع. تتم عملية التحقق من خلال تمثيل التواقيع الصحيحة وتلك المزيفة على شكل نقاط في فضاء إقليدي ثلاثي الأبعاد حيث يتم تغليفهم داخل كرات. تم إخضاع هذا الأنموذج للمقارنة مع الشبكات العصبية الاصطناعية ونظام الاستنباط التكيفي العصبي-الإبهامي (الترجيحي). تتفوق الطريقة المستحدثة على الأخرتين من ناحية إحصاء القبول الخاطئ وإحصاء الرفض الخاطئ وإحصاء التصنيف الصحيح.

TABLE OF CONTENTS

List of Figures

List of Tables

CHAPTER 1

INTRODUCTION

1.1 Introduction to Signature Verification Systems

Personal identification is the process of associating a particular individual with an identity. Identification can be in the form of verification (also known as authentication) or recognition (also known as identification). A particular class of identification technologies that are used to determine the human's identity is the Biometrics technologies. These technologies measure and analyze the human body characteristics and use an individual's unique biological traits to determine one's identity. There are different types of biometrics which are the chemical and the visual ones; those two kinds include different types of traits such as DNA, fingerprints, retina and iris patterns, facial characteristics, signature and many more. The technologies behind biometrics are still emerging, and they provide a lot of security benefits across the spectrum. [1]

Handwritten signature is a behavioral biometric; that is not based on physiological properties of the individual, such as fingerprint or face, but behavioral ones. As such, one's signature may change over time and it is not nearly as unique as iris patterns or fingerprints. [2]

Handwritten signatures are commonly used to approbate the contents of a document or to authenticate a financial transaction as they are the primary mechanism both for authentication and authorization in legal transactions, then the need for research in efficient automated solutions for signature recognition and verification has increased in recent years. [3]

In addition to that, signatures are composed of special characters and therefore most of the time they can be unreadable. The variation among signatures of the same person is called Intra Personal Variation, while the variation between originals and forgeries is called Inter Personal Variation; so these differences make it necessary to analyze them as complete images and not as letters or words put together. [4]

A wide range of methods for signature verification have been reported, and showed that signature verification is a difficult pattern recognition problem; because of the fact that intra-class variations can be large, and on the other hand signatures are easier to be forged than other biometric attributes, then different types of forgeries can be carried out randomly, unskilled and skilled ones; this led to that even forensic experts cannot always tell whether a signature is authentic or not. [5]

1.2 Online and Offline Signature Verification Systems

Signature verification systems can be divided according to the acquisition of the data into online and offline ones. Online (dynamic) signature verification system uses the signatures that are captured by pressure sensitive tablets that extract dynamic properties (location, velocity, acceleration, and pen pressure as functions of time) of a signature in addition to its shape, and can be used in real time applications like credit card transactions, protection of small personal devices (e.g. PDA, laptop), authorization of computer users for accessing sensitive data or programs, and authentication of individuals for access to physical devices or buildings. [6]

Offline (static) signature verification system takes the 2D image of a signature as an input; this technique is useful in automatic verification of signatures found on bank checks and documents. Signatures in off-line systems usually may have noise, due to the scanning hardware or paper background, and contain less discriminative information since only the image of the signature is the input to the system. While genuine signatures of the same person may slightly vary, the differences between a forgery and a genuine signature may be imperceptible, which makes automatic off-line signature verification be a very challenging pattern recognition problem. Besides, the difference in pen widths and unpredictable change in signature's aspect ratio are other difficulties of the problem. [2]

1.3 Motivation and Problem Statement

One of the primary missions that is considered as a necessity for different societies all over the world in numerous daily applications is the automatic authentication and verification of the identity of a person (e.g. cheque verification in a bank), according to that; the development of a full proof signature verification scheme which can guarantee maximum possible security is the most widespread procedure that aims to achieve that mission [7]. In the last decade, a colour code algorithm [8] has been developed that seems to have the merit of being relatively fast and simple; compared to other techniques, where the algorithm attains a higher performance.

Our signature verification system differs from other ones; as it mimics the way a person (e.g. a bank employee) compares signatures. Where in an attempt to imitate how a person compares signatures, the colour code algorithm [8] overlaps the two pre-processed signatures under comparison and accordingly decides on the degree of their similitude. At first, different dilations [9] are undertaken, followed by the generation of a check pattern. The analysis of this pattern gives a set of features that are used as entries in the process of verification. The singularity and the robustness of our algorithm is represented in its reliance on mathematics of modelling and the use of an algebraic formalism to find out the results; where a purely mathematical problem will be formulated in order to be solved, finding a result that can be used in the verification process using a new algebraic algorithm that can provide comparable or better performance than already established offline signature verification schemes.

Moreover and after testing our mathematical model, another two models were designed for the verification process. The first one has been designed based on the Artificial Neural Networks (ANNs), while the other is based on the Adaptive Neuro-Fuzzy Inference System (ANFIS). The same entries (set of features) were used as inputs for both models; providing the ability of making a reliable comparison between the different results we get, and then deciding which one of them is the best. Figure 1.1 shows our proposed signature verification system.

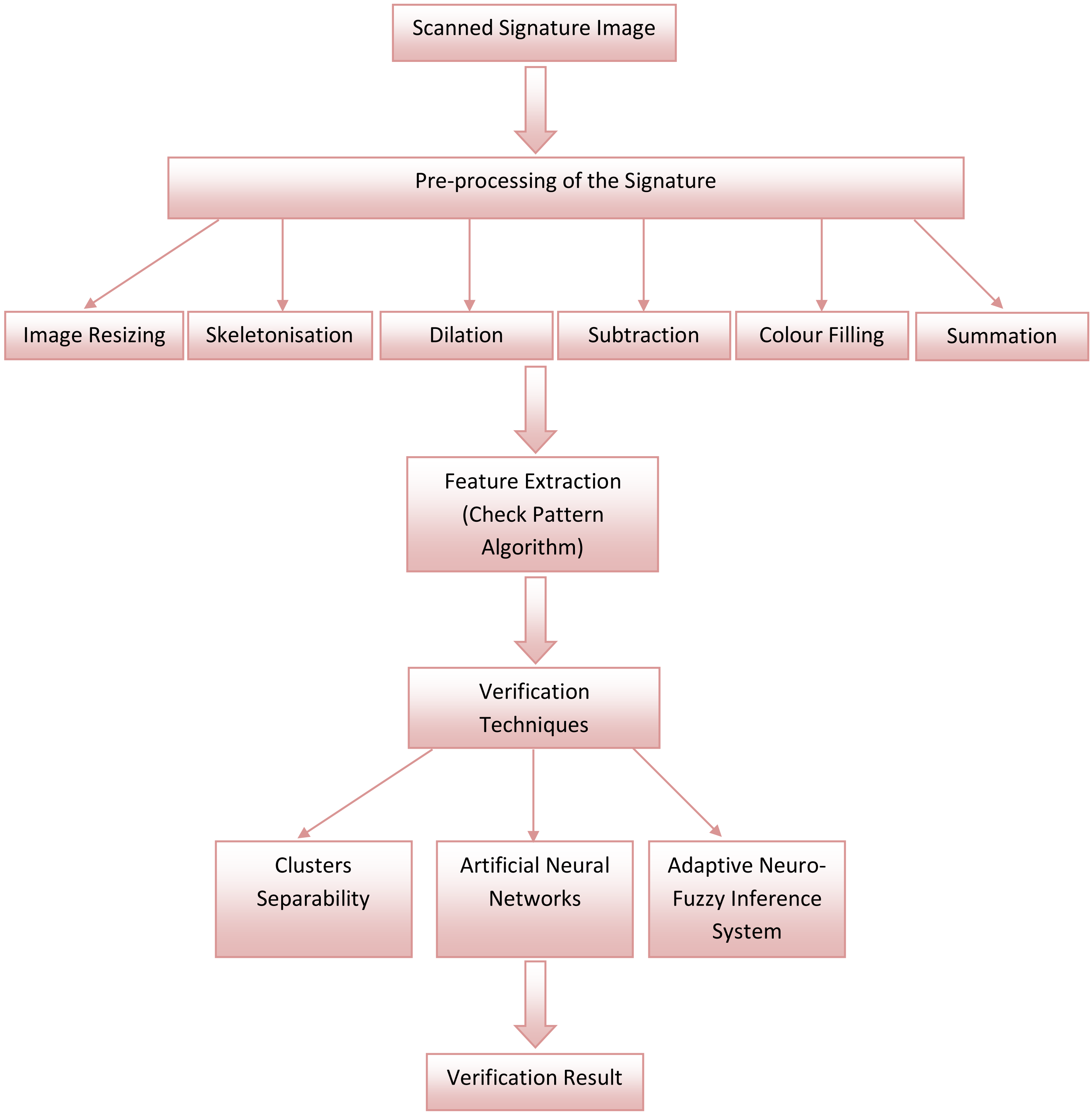


Figure 1.1: Block diagram of our proposed signature verification system.

1.4 Thesis Organization

This thesis is organized as follows. Chapter 2 provides an overview of the existing offline signature verification systems. In the third chapter, we clarify that for any inputted signature image that will be compared with other signatures either genuine or forged ones, a set of pre-processing stages must be applied on them. These stages are represented in resizing the signature image, finding its skeleton, followed by generating four different contours through dilation, colour filling, and subtraction processes, finally a fully pre-processed signature image is prepared by summing those coloured contours.

Chapter 4 explains the Check Pattern Algorithm which aims to generate a check pattern by overlapping (XORing) every two pre-processed signatures images, where it always takes the reference signature image as the first one and the other one is the genuine or forged signature image under comparison, and accordingly decides on the degree of their similitude by analyzing the generated pattern, and getting a set of features that are used in the process of verification later. Chapter 5 discusses the details of the different proposed signature verification techniques. The subsequent chapter discusses the details of measuring the performance of a signature verification system, the evaluation of the proposed approaches and the results achieved. Chapter 7 concludes the thesis and presents the scope for future work.

CHAPTER 2

LITERATURE SURVEY

2.1 Offline Signature Verification Systems

Offline Signature Verification is still an open research problem. In the past years, several attempts have been made, where different techniques and methodologies were tried out to solve the problem. Manasjyoti Bhuyan, Kandarpa Kumar Sarma and Hirendra Das [10] worked on Offline Signature Recognition and Verification using Hybrid Features and Clustered Artificial Neural Network ANNs, where a set of preprocessed training signatures with specific characteristics and of three different languages, are used to extract and calculate a set of several features such as Euclidian distances from horizontal and vertical sectioning, Standard deviation of the row wise sum and column wise sum, Standard deviation, mean and median for each projection of the signature. These features are the input to a set of Artificial Neural Networks (in this case four different ANNs) that are already trained with similar features extracted from other signatures. The primary objective of this process is to reduce the two crucial parameters which are the False Acceptance Rate (FAR) and False Rejection Rate (FRR) with lesser training time. This method deals with tilted and forged signatures successfully, as it captures finer variations in the signature, and it is expected to extend this work in order to include a wide class of signature and form an effective verification system.

Dr. S. Adebayo Daramola and Prof. T. Samuel Ibiyemi [11] proposed a signature recognition system that basically depends on using Discrete Cosine Transform (DCT) and Hidden Markov Model (HMM) which is a probabilistic pattern matching technique that has the ability to absorb both the variability and the similarity between signature samples. In this system after acquiring the signature, several preprocessing steps are performed on it, then using horizontal and vertical splitting techniques, the signature image was vertically segmented into four blocks, where each block is segmented into 16 smaller cells, after that the DCT coefficients are calculated and used to form the observation vector, which is then quantized, finally, Hidden Markov Models (HMM) represent a signature as a sequence of states. In each state an observation vector can be generated, then the HMM probabilities, or parameters are trained using observation vector extracted from a representative sample of signature data. Recognition of an unknown signature is based on the probability that a signature was generated by the HMM. The recognition performance they achieved in this method is 99.2%, as out of 500 signature images tested only four signatures were not recognized.

Kai Huang and Hong Yan [12] suggested a new method that achieved a 90% correct classification rate after being tested on a database composed of over 3000 signature images, this method was for off-line signature verification that depends on the use of multi-resolution feature extraction and multiple expert voting techniques. Through it different features of the signature are found in order to describe its shape; such as its core, outline, ink area distribution and its frontiers. A rough shape comparison is performed to filter out the less skilled forgery inputs. Then an Artificial Neural Network classifier that is trained with the genuine signature shape at multiple resolutions is used to simultaneously examine the geometric features of an input signature image. After that the acceptance and rejection space are effectively controlled by perturbation parameters.

Mehdi Radmehr, Seyed Mahmoud Anisheh, Mohsen Nikpour and Abbas Yaseri [13] worked on off-line signature recognition, and designed a new system based on Radon Transform, Fractal Dimension (FD) and Support Vector Machine (SVM). Firstly, Radon Transform (RD) is applied on the original signatures with angles of 0°, 45°, 90° and 135°, in order to produce four projections (vectors) of it along four specified directions. Then calculating these vectors Fractal Dimensions will result in finding a feature vector for each signature in which their values indicate the complexity of a pattern, or the quantity of information embedded in them. These feature vectors will be used finally as inputs for the Support Vector Machine Classifier (SVM) that is used for the recognition of these signatures. Experimental result indicates that this method achieved high accuracy rate in signature recognition, but the statistics were omitted.

Vinayak Balkrishana Kulkarni [8] proposed a different technique that deals with the recognition of the signature, as human operators generally make the work of signature recognition. Hence the algorithm simulates human behavior, to achieve perfection and skill through Artificial Intelligence. Signature recognition is done by applying several dilations of various levels on the pre-processed signature, in order to generate contours around the signature image where each contour is of different radii of the other one; these contours are then coloured with different colour bands that are black, blue, red and green. After that a detection process is done using the XOR operation between three standard signature templates of the same person, and a test signature, which will generate a check pattern. This technique is used with a neuro-fuzzy classifier for signature verification and tested on various operating systems. They found that it works very well with an accuracy of 80% to 90%in recognition process.

The vertical splitting, horizontal splitting, and diagonal splitting are representing the Geometric Center Features that were studied by Sepideh Afsardoost , Siamak Yousefi, and Mohammad Ali Khorshidi [14] and used for offline signature verification. The FAR and FRR values found based on those features were 10% and 15% respectively.

Ramachandra C, Jyothi Srinivasa Rao, K B Raja, K R Venugopla, and L M Patnaik [15] achieved new values for both of FAR and FRR as they were 4.6% and 5.4% respectively, as they developed a signature verification system based on the maximum horizontal and vertical histogram, horizontal and vertical centers of signature, and finally the aspect ratio and edge points of the signature.

The Artificial Neural Networks (ANNs) were used to verify the signatures by Ali Karouni, Bassam Daya, and Samia Bahlak [16], as the performance rate of the system wa about 93%, the FAR= 1.6% and the FRR= 3%.

Table 2.1: Offline Signature Verification Systems.

No.	The Approach	Description	FAR	FRR	CCR	Reference
1.	Hybrid Features and Artificial Neural Networks (ANNs).	1st ANN trained using the features extracted from vertical sectioning of the signatures.	15%	25%		[10]
		2nd ANN trained using the features extracted from horizontal sectioning of the signatures.	30%	13%		
		3rd ANN trained using row wise, column wise, diagonal sum and the projections features.	20%	25%		
		The last ANN was trained using the skeleton.	15%	15%		
2.	Discrete Cosine Transform (DCT) and Hidden Markov Model (HMM).				99.2 %	[11]
3.	Geometric Feature Extraction and Neural Network Classification.				90%	[12]
4.	Radon Transform, Fractal Dimension (FD) and Support Vector Machine (SVM)					[13]
5.	Neuro Fuzzy Logic.				80 % to 90%	[8]
6.	The Geometric Center Features.		10%	15%		[14]

No.	The Approach	Description	FAR	FRR	CCR	Reference
7.	Robust Signature Verification based on Global Features.		4.6%	5.4%		[15]
8.	Artificial Neural Network (ANN).		1.6%	3%		[16]

CHAPTER 3

PRE-PROCESSING OF THE SIGNATURES

3.1 Introduction

The proposed offline signature verification algorithm uses as input the image of the scanned signature, where all samples of the signatures are assumed to be correctly rotated and of the same skew angle before starting doing anything. This image is then subject to a set of pre-processing steps; resizing, finding its skeleton, dilating the skeletonised signature four times with four different margins, applying a subtraction process to find four different contours, filling each contour with a specified colour, and finally applying a summation process onto the image related to the four coloured contours. The result of this pre-processing algorithm is a colour coded signature. The idea of this colour coding is similar to what has been done in [8], but the code is completely the work of the researcher. This algorithm is applied to all images of the signatures in the database, which results in a colour coded database of signatures.

3.2 Resizing the image of the signature

The algorithm accepts as input an image of signature of a standardized size of 256×256 pixels. This is why the first step is resizing the image to fit this standard.

Figures 3.1 and 3.2 show an example of the resizing process. The binary image is resized using bicubic interpolation and antialiasing for shrinking the image methodologies.

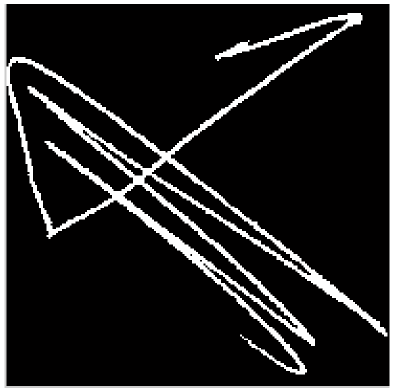


Figure 3.1: The original size of the image of a signature.

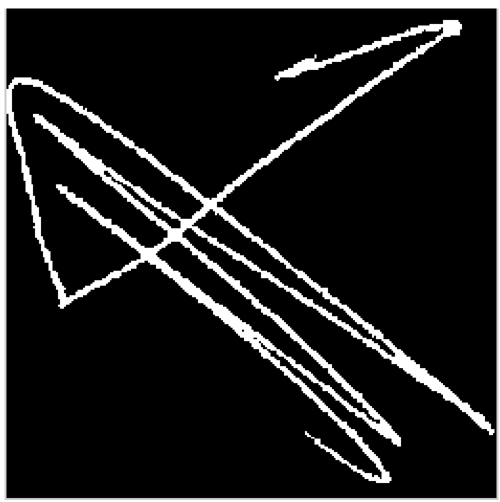


Figure 3.2: The resized image of the signature in figure 3.1.

3.3 Skeletonisation

A morphological operation called Skeletonisation [17] is used to reduce the foreground regions in the scaled image to skeletal residues, by throwing away most of the original foreground pixels, simultaneously maintaining the extent and connectivity of the original region, as shown below.

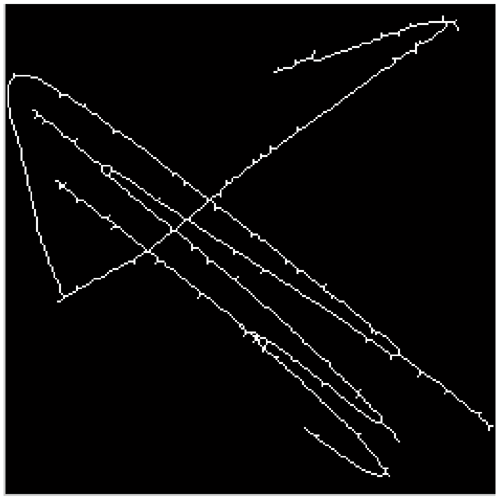


Figure 3.3: Skeletonised image of the signature in figure 3.1.

3.4 Dilation

The next step is represented in the application of another morphological operation called dilation [18]; this process aims to expand the shape of the skeletonised signature, through the use of a structuring element, generating a contour around the external boundary of that shape. Since we are in need to have four different contours of four different radii for each signature, various levels of dilations are used, such that a contour of a higher radius is drawn around the original shape each time.

Four levels of dilation are used [19], generating four contours with $r_1 = 3$ pixels, $r_2 = 6$ pixels, $r_3 = 10$ pixels, and $r_4 = 16$ pixels radius; consequently we get for each skeletonised signature four structures with bands of varying thicknesses, as can be seen in figure 3.4. Dilation diameters (sizes of margins) are directly related to the strength of the algorithm, its security, and its ability to verify the signature**.**

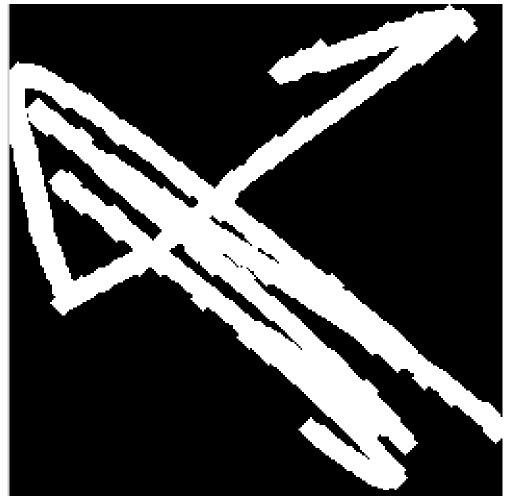
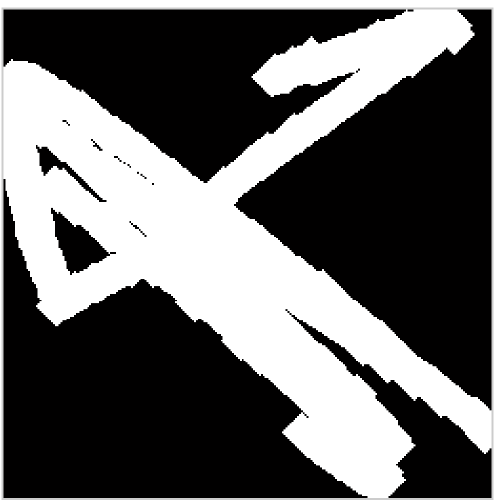
	
First level of dilation.	**Second level of dilation.**
	
Third level of dilation.	**Fourth level of dilation.**

Figure 3.4: The four levels of dilations.

3.5 Subtraction

This stage works on the four dilated images and the skeletonised one, by finding the difference between them; so again we will get four new images as shown in figure 3.5, where each image represents a contour of a different size generated for the signature and representing its external boundary, which will be worked on later. More clearly, the first image here is the outcome of subtracting the image of the first level of dilation and the skeletonised image, the second image is the outcome of subtracting the image of the second level of dilation and the image of the first level of dilation, as well the third image is the outcome of subtracting the third and the second level of dilation images, finally by subtracting the image of the fourth level of dilation from the third level of dilation image we get the last one.

3.6 Colour Filling

Figure 3.6 shows that the four previously generated contours will be filled with different colours: black, red, green, and blue; the choice of those colours is randomly done where changing them is possible taking into account that we have to change the resulting colours in the generated pattern later. But since the background colour of all the images is still black, an arbitrary colour (purple) is selected and used to fill the first contour region leaving it as it is until the following step. The colours are represented in RGB format, where Black is (0, 0, 0), Red is (255, 0, 0), Green is (0, 255, 0), Blue is (0, 0, 255), and the arbitrary colour (Purple) is (255, 0, 255).

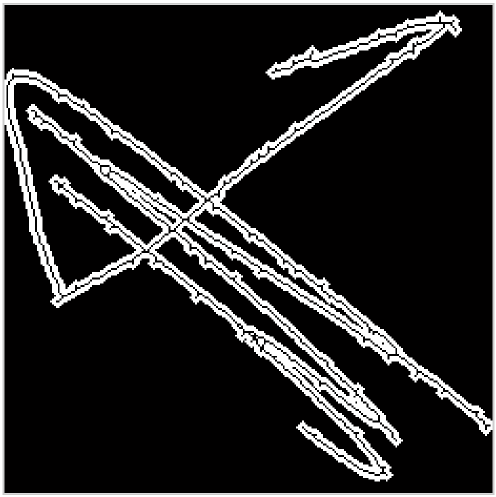
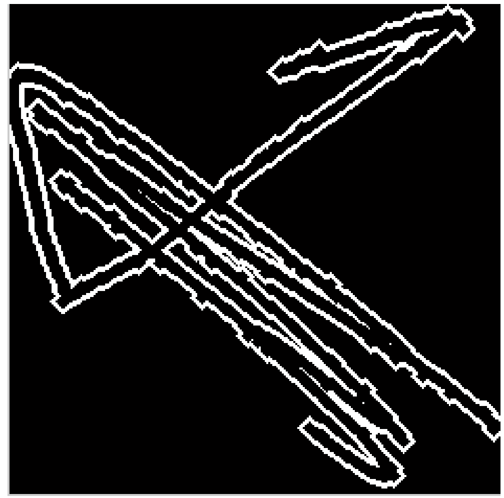
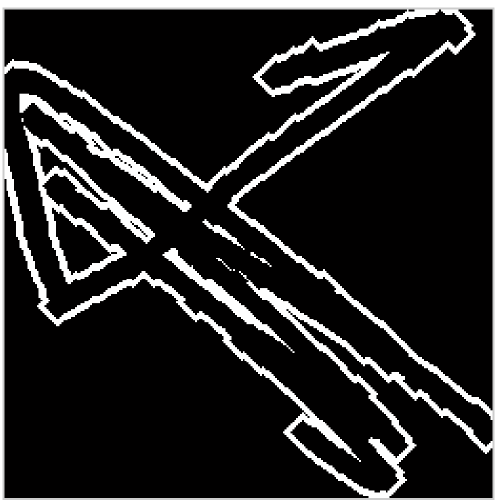
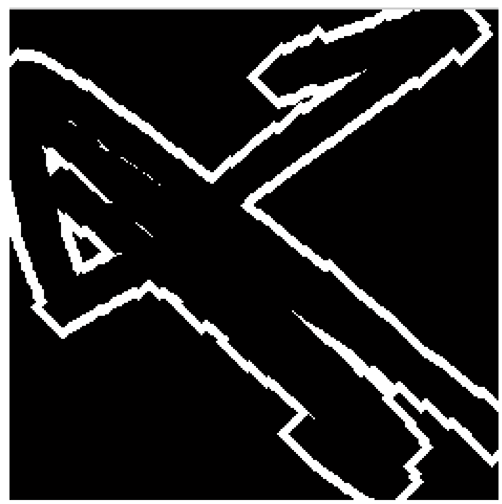
	
The first contour	**The second contour**
	
The third contour	**The fourth contour**

Figure 3.5: The resulting images of the subtraction stage.

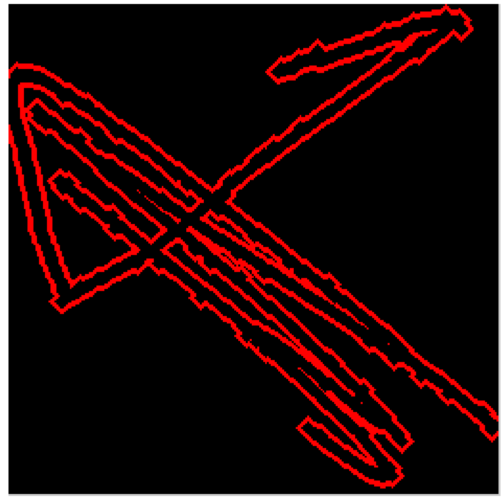
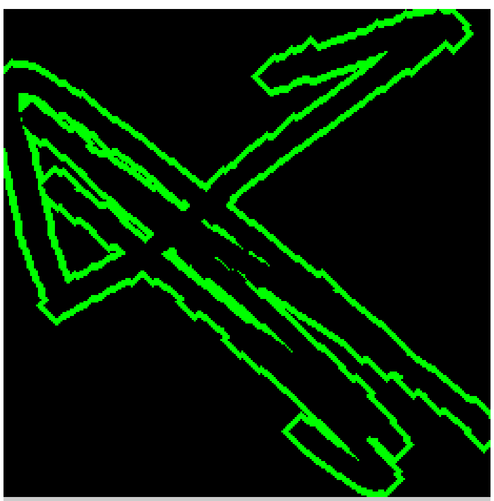
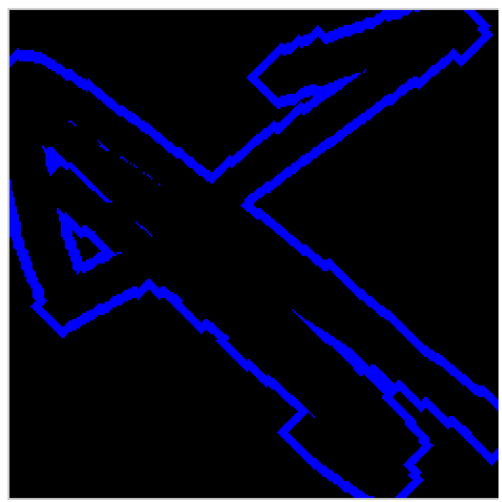
	
The first coloured contour	The second coloured contour
	
The third coloured contour	The fourth coloured contour

Figure 3.6: Colour filling of the contours.

3.7 Summation

The final pre-processing stage is the generation of the colour code image of the signature we started with. This colour coded image is outfitted on three successive phases; the first one is done through the summation of the four different coloured contour images we get from the previous step, keeping both the background colour and the first coloured contour as they are. The result is illustrated in figure 3.7.

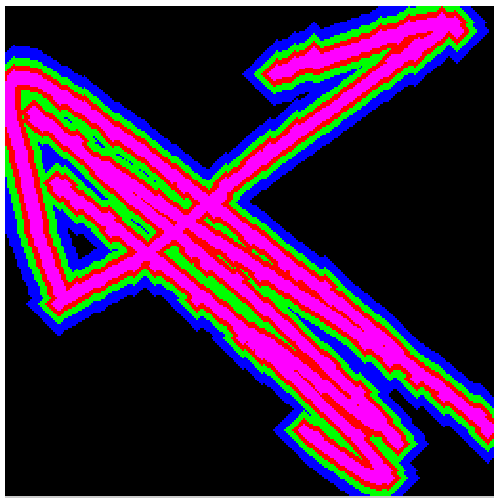


Figure 3.7: Sum of four coloured contour images of the signature.

It is wanted to use the black colour to represent the contour of the dilated skeleton of the signature. This is why the background colour is first changed into a new colour (dark teal) that equals in the RGB format (0, 100, 96), as shown is figure 3.8. Figure 3.9 clarifies that finally the pixels of the first contour with the arbitrary colour (purple) are changed into black.

We have to note that either for training or testing purposes; all standard images of signatures for the same person will go through all the pre-processing steps illustrated in this chapter.

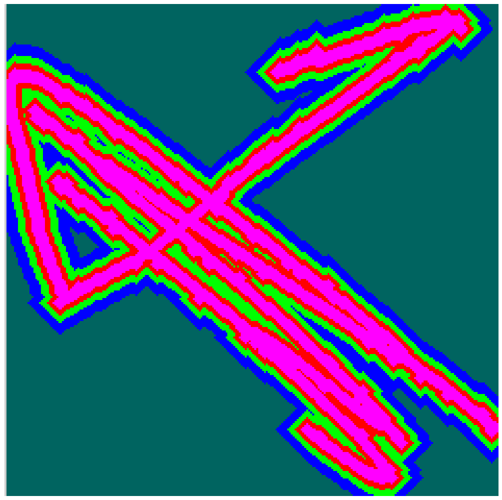


Figure 3.8: The second version of the final signature image.

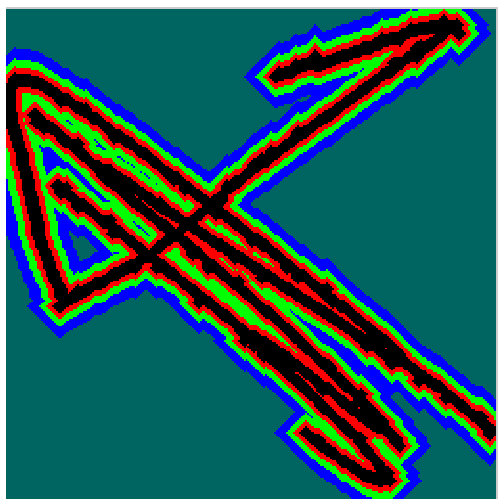


Figure 3.9: The final signature image.

CHAPTER 4

CHECK PATTERN ALGORITHM

4.1 Introduction

In an attempt to imitate how a person compares signatures, the check pattern algorithm overlaps a given signature with a reference signature and accordingly decides on the degree of their similitude. The check pattern is generated by overlapping (XORing) the colour coded signatures. The analysis of the pattern gives a set of features that are later used in the process of verification.

4.2 Check pattern generation

Our Signature Verification automated system differs from other ones, as it mimics the way a bank employee compares signatures; in the sense that it generates a check pattern that represents a comparison between a given new signature and a reference signature of the same person. This is done by overlapping (XORing) them, where always the first inputted one is a genuine signature (reference signature) and the second one could be either another genuine or forged one (new signature), but the important issue we have to take into consideration is that all of these signatures must be colour coded before being overlapped. The generated check pattern consists of a range of different colours, some of those colours already exist in the final pre-processed signatures images(the colour coded ones), while the residual apparent colours are new ones, and they are not of a random emergence.

4.3 Check pattern analysis

The analysis of the generated check pattern gives a set of features, those features are represented in the number of the different colour codes appear in it, where subsequently each colour has its own meaning and its own impact on the verification process.

4.3.1 Check pattern algorithm

The colour codes that will appear in the generated check pattern are represented in (R G B) format, and listed in table 4.1. With regards to the emergence of any of the previously mentioned colours, one or several terms must be met, then whenever one of those terms is achieved the related colour appears. Those terms are illustrated in table 4.2.

The total number of the colours in the generated check pattern is ten, but despite that we do not use the count of all of them as features in the analysis stage, because only the white, the black, the original background, the red, the green, and the blue colours are the most important and meaningful ones. The emergence of each of the white, the black, and the background coloured pixels in the generated check pattern reflects the extent of match between the first coloured contour (the black contour) in the reference signature image, and all other parts of the other test signature image.

Table 4.1: Colour codes in the check pattern.

The Colour	(RGB) Format
White	(255, 255, 255)
Original background (dark teal)	(0, 100, 96)
Black	(0, 0, 0)
Red	(255, 0, 0)
Green	(0, 255, 0)
Blue	(0, 0, 255)
Colour 1 (Aqua)	(0, 252, 255)
Colour 2 (Purple)	(255, 8, 255)
Colour 3 (Yellow)	(255, 252, 0)
Test background (Rose)	(255, 156,168)

Table 4.2: Combinations for colour emergence in the check pattern.

XOR table		**Reference Signature**				
		Background	**Black**	**Red**	**Green**	**Blue**
Test Signature	**Background**	Test BG	white	Colour1	Colour2	Colour3
	Black	BG	Black	Red	Green	Blue
	Red	Test BG	white	Colour1	Colour2	Colour3
	Green	Test BG	white	Colour1	Colour2	Colour3
	Blue	Test BG	white	Colour1	Colour2	Colour3

More clearly, the number of the white coloured pixels in the check pattern will increase whenever the first coloured contour of the reference signature image touches any of the other coloured contours (except the first one) or the background of the test signature image during the overlapping process because of its deviation in the test image which causes a different in the shape of the other contours, while the deviation of the test signature's black contour will lead into an increment in the count of the original background pixels. Besides that the number of the black coloured pixels will decrease accordingly, as the appearance of those black pixels directly depend on the match between the first coloured contour in the two overlapped signatures only.

On the other hand, the red, the green, and the blue coloured pixels in the generated check pattern appears whenever there is an overlapping between the other three coloured contours related to the reference signature image with only the first coloured contour of the test signature image, which is the black one.

4.3.2 Illustrative examples

For example figure 4.1 shows a generated check pattern resulting from the overlapping of two identical colour coded signatures images, as indicated before and because of the exact match between the first black coloured contour of the two signatures images, the generated check pattern will contain no white, no original background, no red, no green, and no blue coloured pixels, while the number of the black pixels will be in its highest possible value as illustrated in table 4.3.

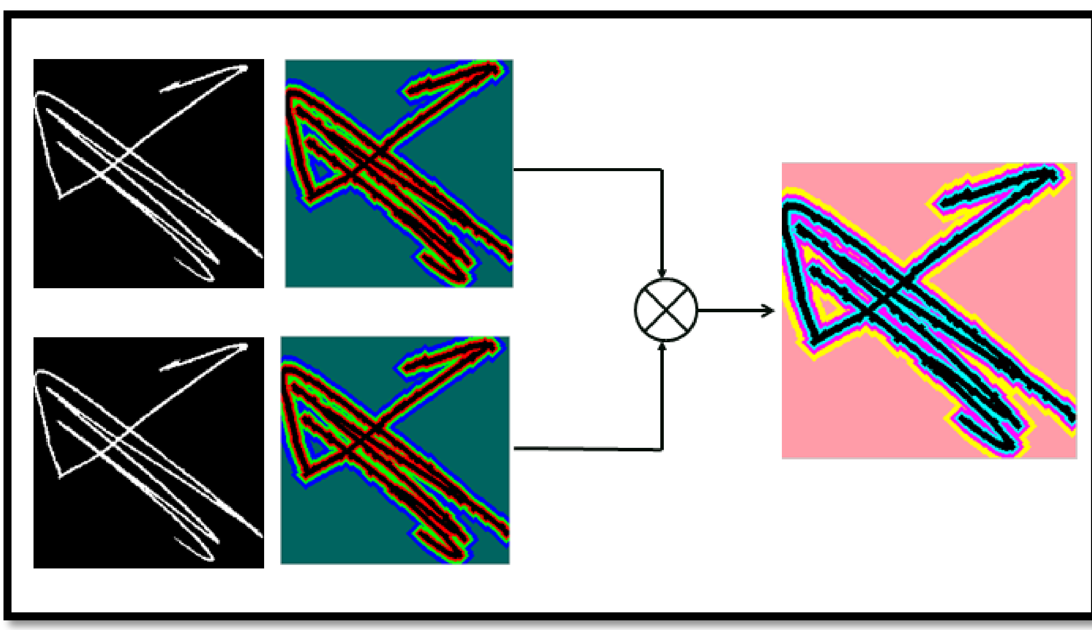


Figure 4.1: Overlapping two identical signatures.

Table 4.3: Colour counts of the generated check pattern in figure 4.1.

The Colour	Value
White	0.0
Original background (dark teal)	0.0
Black	9640.0
Red	0.0
Green	0.0
Blue	0.0

But whenever there is a change between the two overlapped signatures images the resulting coloured pixels counts will alter each according to the degree of difference between the two overlapped ones, as can be seen in both figure 4.2 and figure 4.3. The counts of the colours in figure 4.2 are illustrated in table 4.4, while the colours counts of figure 4.3 are illustrated in table 4.5.

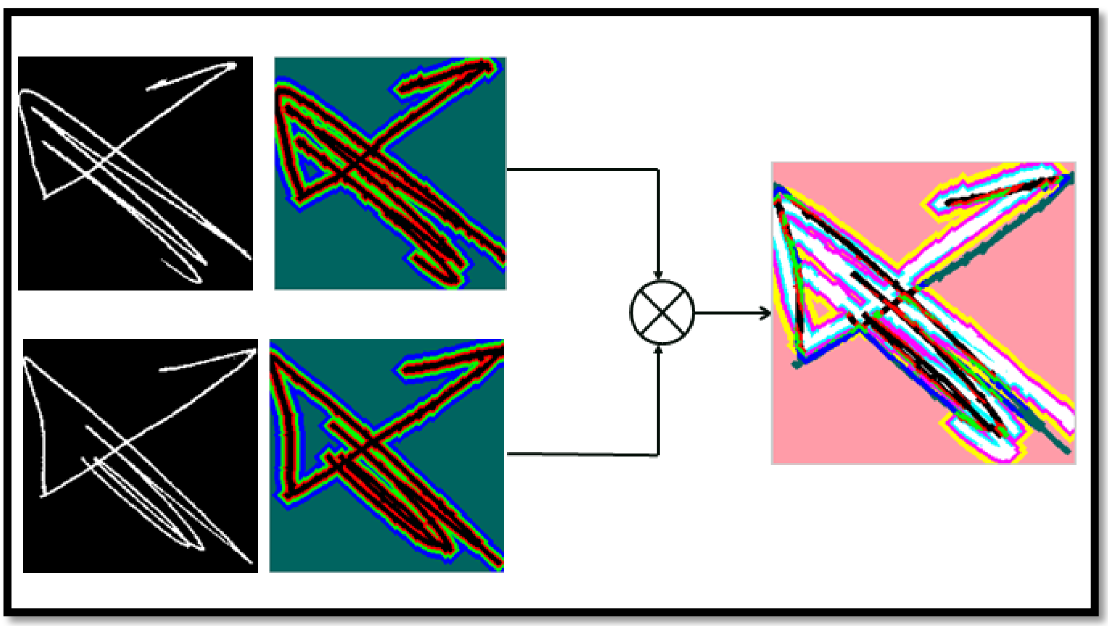


Figure 4.2: Overlapping two somewhat similar signatures.

Table 4.4: Colour counts of the generated check pattern in figure 4.2.

The Colour	Value
White	7886.0
Original background (dark teal)	1567.0
Black	2933.0
Red	1990.0
Green	1490.0
Blue	1314.0

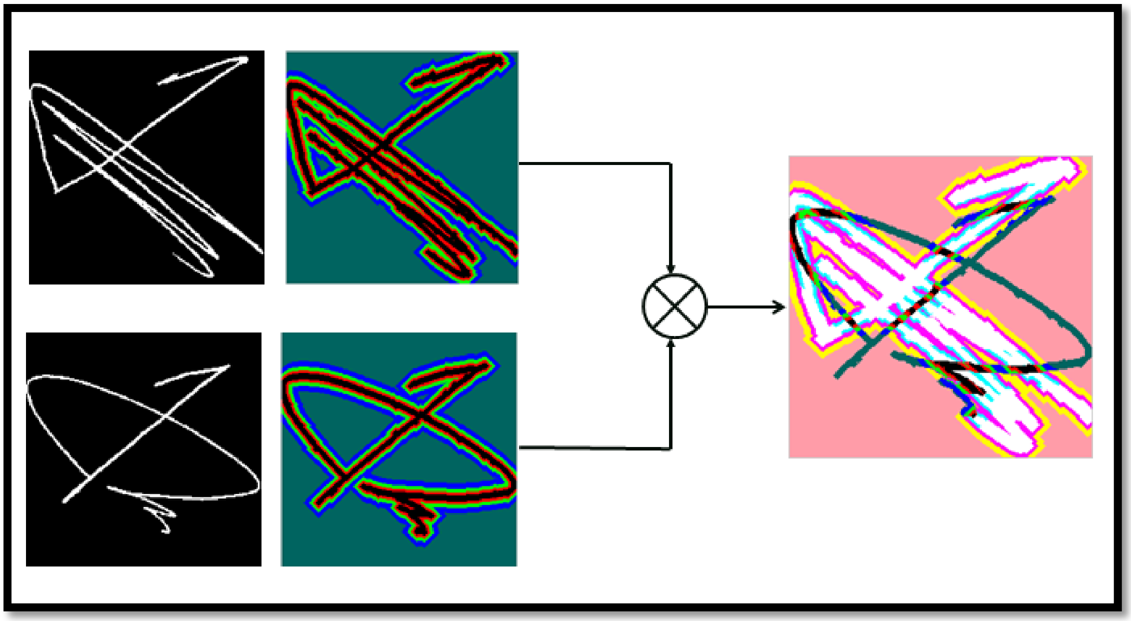


Figure 4.3: Overlapping two different signatures.

Table 4.5: Colour counts of the generated check pattern in figure 4.3.

The Colour	Value
White	12051.0
Original background (dark teal)	2624.0
Black	1150.0
Red	768.0
Green	747.0
Blue	1030.0

CHAPTER 5

SIGNATURE VERIFICATION IN A COLOUR CODE 3D EUCLIDEAN SPACE

5.1 Introduction

The Database of Signatures [20] is composed of groups of genuine and forged versions of twenty persons' signatures; for each person there are ten genuine and nine forged signatures. The first step in the processing of these signatures is the extraction of a certain number of features; these features are colour codes that represent different degrees of overlapping between any two of the signatures. The overlapping is found using the XOR logic operation.

For each person, an arbitrarily chosen genuine signature is XORed with five other genuine signatures and five forged ones. In the resulting image, the XORing reveals the different degrees of overlapping as coloured regions in the field of the image. For each signature we extract the number of pixels of each coloured region: number of Red (R) pixels, number of Green (G) pixels, number of Blue (B) pixels, number of White (W) pixels, number of Black (B_K) pixels, and number of Background (B_G) pixels. Accordingly we have two data sets of features (number of pixels per colour), one that describes the genuine signatures, and a second that describes the forged signatures. The six colours of each description are separated into two groups: the Red/Green/Blue (RGB) group and the White/Background/Black (WB_GB_K) group. This separation is done according to what each group represents in the resulting image of the XORing process (see chapter 4): WB_GB_K represents the level of similarity to the reference signature, while RGB represents the dissimilarity as they result from the intersection between the test signature with the dilations regions (see chapter 3) of the reference signature. Hence we define two three dimensional spaces where we will study the separability of two clusters of signatures per person, the genuine cluster and the forged cluster. It is suggested to study the separability of the two clusters in each space using different geometric and algebraic methods. According to the degree of success we might decide if necessary to go to a higher dimensional hyper-space or not; where the new axes will be represented by new extracted features such as the height of the signature, the width of the signature, the height to width ratio and others.

Let $\boldsymbol{G}$ be the matrix of features of **G**enuine signatures where each column represents the colour code of the XORing operation. Let $\boldsymbol{F}$ be the matrix of features of **F**orged signatures built in the same way. The matrix $\boldsymbol{G}$ is formed using five genuine signatures of a specific person from the database. The matrix $\boldsymbol{F}$ is formed using five forged signatures of a specific person from the database. The first step is to find the vectors $\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3$ that span the column space of $\boldsymbol{G}$, and for this three-dimensional space we calculate an orthogonal basis, $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$. The columns of the matrices $\boldsymbol{G}$ and $\boldsymbol{F}$ are then transformed into the space spanned by the orthogonal basis $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$. For each, having all the points in this space, we search for a plane that separates between the cluster of points that represent the genuine and the forged signatures. This plane is found at the intersection between a sphere that encloses all the points that represent genuine signatures $\boldsymbol{G}$, and a second sphere that encloses all the points that represent forged signatures $\boldsymbol{F}$.

Therefore, for each person we will have a geometric depiction of her/his signatures taken from the database. This geometric representation is to be used to verify the authenticity of a new appearance of the signature of that same person. The issue is now to study the separability of the clustering of signatures in the above mentioned three-dimensional space compared to the separability of different other algorithms which are the Artificial Neural Networks (ANNs) and the Adaptive Neuro-Fuzzy Inference System (ANFIS) that will be implemented using the same set of features which are used in the geometric clustering algorithm.

5.2 Finding the transformation T: $\mathbb{R}^3 \rightarrow \mathbb{R}^3$

For both matrices $\boldsymbol{G}$ and $\boldsymbol{F}$ formed of different appearances ($\boldsymbol{u}_k$ for $k = 1,2,3,\ldots$) of some genuine and forged signatures of an individual we want to study a new appearance $\mathrm{m_n}$ and decide whether it is genuine or forged. For this purpose we locate this appearance in the column space of $\boldsymbol{G}$ where a separation plane between the clusters of genuine and forged signatures is already defined.

Let:

$$\boldsymbol{G} = [\boldsymbol{u}_1 \quad \boldsymbol{u}_2 \quad \cdots \quad \boldsymbol{u}_k \quad \cdots \quad \boldsymbol{u}_K]$$

$$\boldsymbol{F} = [\boldsymbol{o}_1 \quad \boldsymbol{o}_2 \quad \cdots \quad \boldsymbol{o}_k \quad \cdots \quad \boldsymbol{o}_K]$$

We want to view the column vectors of both $\boldsymbol{G}$ and $\boldsymbol{F}$ in an N-dimensional space spanned by an orthogonal basis and then study the separability of the clusters of genuine and forged signatures of a given person using different methods.

5.3 The set of vectors that span the column space of $\boldsymbol{G}$

Suppose the column vectors of $\boldsymbol{G}$ are linearly independent, using row reduction we find three vectors that span the column space of $\boldsymbol{G}$:

$$\boldsymbol{S} = [\boldsymbol{u}_l \quad \boldsymbol{u}_m \quad \boldsymbol{u}_n]$$

To prove that these vectors span the column space $C(\boldsymbol{S})$ it is enough to show that $\det(\boldsymbol{S}) \neq 0$. The column space $C(\boldsymbol{S})$ consists of all linear combinations of the columns of $\boldsymbol{G}$. We can write the vectors in $\boldsymbol{G}$ as linear combinations of the input basis $\boldsymbol{B}_i = \{\boldsymbol{u}_l \quad \boldsymbol{u}_m \quad \boldsymbol{u}_n\}$. The coordinates of the vectors $\boldsymbol{u}_k$ in the basis B_i are as follows:

$$[\boldsymbol{u}_k]_{B_i} = \boldsymbol{S}^{-1}\boldsymbol{u}_k \tag{5.1}$$

The coordinates of the vectors $\boldsymbol{o}_k$ of the matrix $\boldsymbol{F}$ in the basis $\boldsymbol{B}_i$ are as follows:

$$[\boldsymbol{o}_k]_{B_i} = \boldsymbol{S}^{-1}\boldsymbol{o}_k \tag{5.2}$$

Assuming that $C(\mathrm{G})$ has the Euclidean inner product, we can show that $\langle \boldsymbol{u}_k, \boldsymbol{u}_l \rangle \neq 0$ for $\mathrm{l} \neq \mathrm{k}$. This means that the column vectors of $\boldsymbol{S}$ do not form an orthogonal basis. Using Gram-Schmidt process we find an orthogonal basis that spans the column space of $\boldsymbol{G}$:

$$\boldsymbol{OS} = [\boldsymbol{e}_1 \quad \cdots \quad \boldsymbol{e}_N] = \boldsymbol{E}$$

5.4 Change of basis

Coordinates come from a basis. If we change the basis, we change the coordinates. The coordinates of the vectors in $\boldsymbol{G}$ in the orthogonal (output) basis $\boldsymbol{B}_{o1} = \{\boldsymbol{e}_1 \quad \boldsymbol{e}_2 \quad \boldsymbol{e}_3\}$ are given by:

$$[\boldsymbol{u}_k]_{B_{o1}} = \boldsymbol{E}^{-1}\boldsymbol{u}_k \tag{5.3}$$

The coordinates of the vectors in $\boldsymbol{F}$ in the orthogonal (output) basis $\boldsymbol{B}_{o2} = \{\boldsymbol{n}_1 \quad \boldsymbol{n}_2 \quad \boldsymbol{n}_3\}$ are given by:

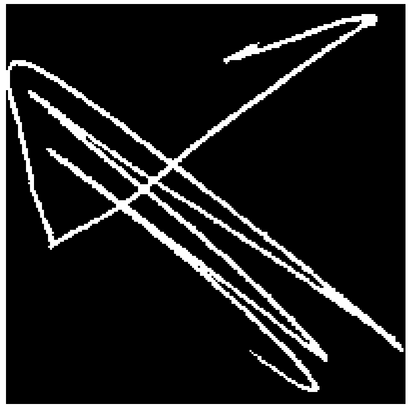
$$[\boldsymbol{o}_k]_{B_{o2}} = \boldsymbol{E}^{-1}\boldsymbol{o}_k \tag{5.4}$$

We locate the different appearances of the signatures to the same individual in the N-dimensional space of genuine signatures using these coordinates.

5.5 An illustrative example

5.5.1 Finding both $\boldsymbol{G}$ and $\boldsymbol{F}$ matrices.

Consider the first genuine signature of a person from the database (shown in the figure below). This genuine signature is considered to be the reference to which all other signatures of that person are compared to for verification.



Genuine Signature #1

Figure 5.1: Reference signature from the database of a given person.

At the beginning, this reference signature is XORed with another five genuine signatures of the same person (shown in figure 5.2), leaving the rest genuine ones as test signatures. Figure 5.3 shows the images resulting from the XORing process. It clearly appears that there is a difference in their manifestations meaning that there will be differences in the colours counts later. The matrix of WB_GB_K pixels of appearances of the five genuine signatures is as follows:

$$\boldsymbol{G} = \begin{bmatrix} 9818.0 & 7886.0 & 9275.0 & 10487.0 & 8280.0 \\ 1264.0 & 1567.0 & 1365.0 & 949.0 & 335.0 \\ 1786.0 & 2933.0 & 2362.0 & 1561.0 & 2225.0 \end{bmatrix}$$

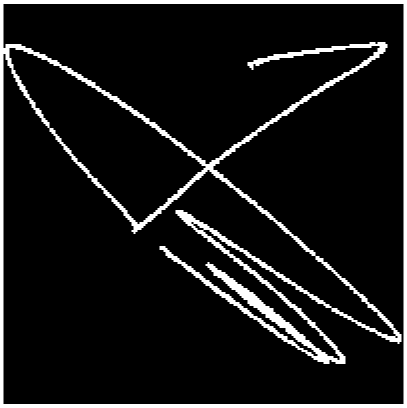
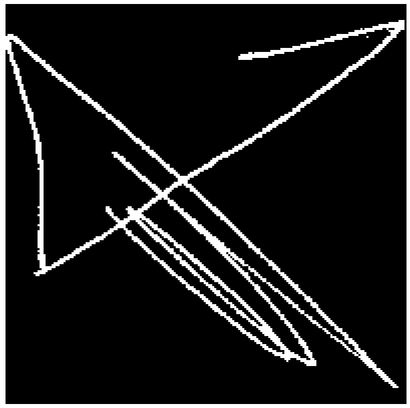
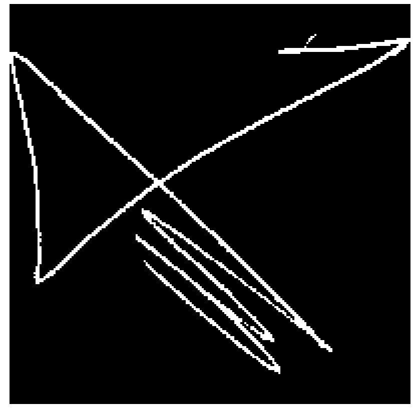
$$= [\boldsymbol{u}_1 \quad \boldsymbol{u}_2 \quad \boldsymbol{u}_3 \quad \boldsymbol{u}_4 \quad \boldsymbol{u}_5]$$

On the other hand, the same reference genuine signature is XORed with the first five forged signatures that mimic it (shown in figure 5.4). The resulting images to some extent differ from the first ones, where the emergence of both white and background colours is clearer as can be seen in figure 5.5. The matrix of WB_GB_K pixels of appearances of the five forged signatures is as follows:

$$F = \begin{bmatrix} 10986.0 & 12051.0 & 10910.0 & 12029.0 & 11209.0 \\ 1799.0 & 2624.0 & 1740.0 & 1701.0 & 3040.0 \\ 1698.0 & 1150.0 & 1752.0 & 1144.0 & 1491.0 \end{bmatrix}$$

$$= [\boldsymbol{o}_1 \quad \boldsymbol{o}_2 \quad \boldsymbol{o}_3 \quad \boldsymbol{o}_4 \quad \boldsymbol{o}_5]$$

The column space of $\boldsymbol{G}$ is the space of genuine signatures of a given person. We want to view the column vectors of both $\boldsymbol{G}$ and $\boldsymbol{F}$ in a three-dimensional space spanned by an orthogonal basis and then study the separability of the clusters of genuine and forged signatures of that person using different methods.

		
Genuine Signature #2	**Genuine Signature #3**	**Genuine Signature #4**

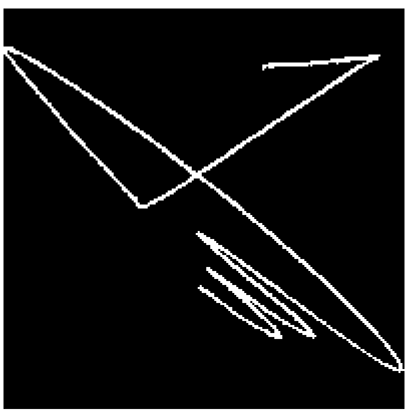
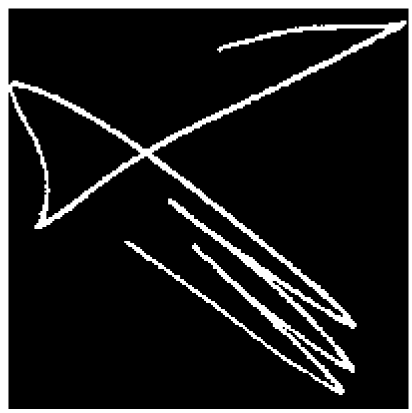
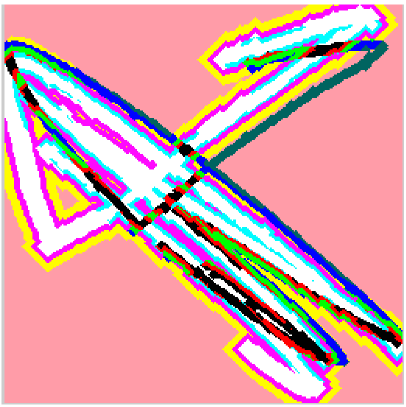
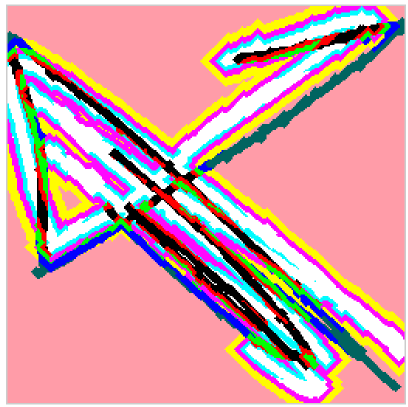
	
Genuine Signature #5	**Genuine Signature #6**

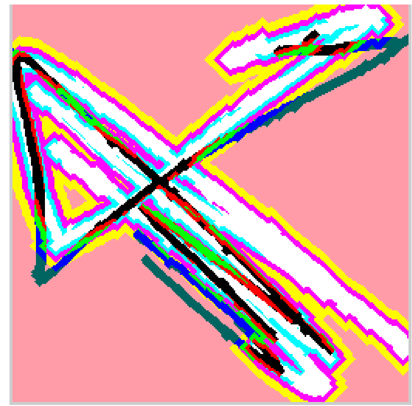
Figure 5.2: Five genuine signatures of the same person.



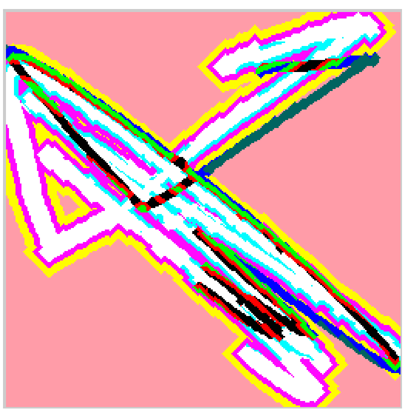
G.S # 1 xor G.S # 2



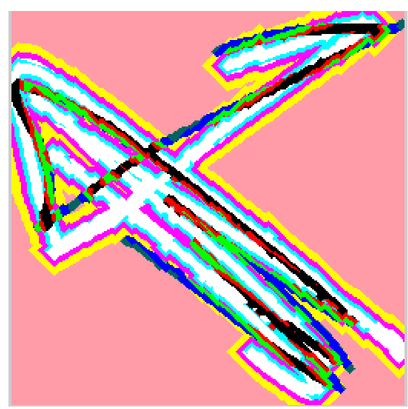
G.S # 1 xor G.S # 3



G.S # 1 xor G.S # 4

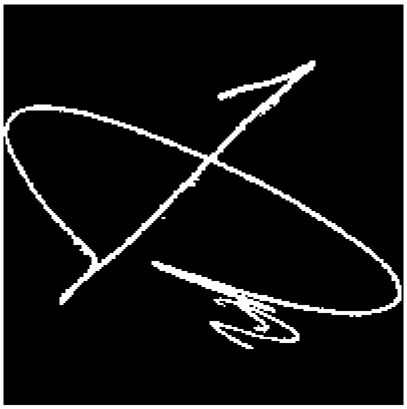


G.S # 1 xor G.S # 5

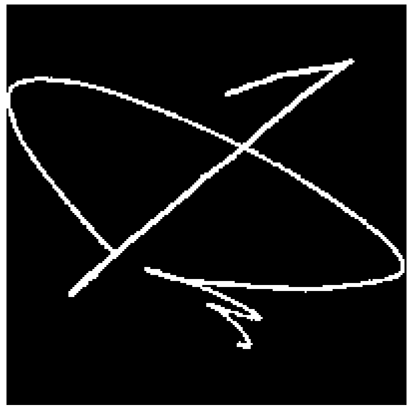


G.S # 1 xor G.S # 6

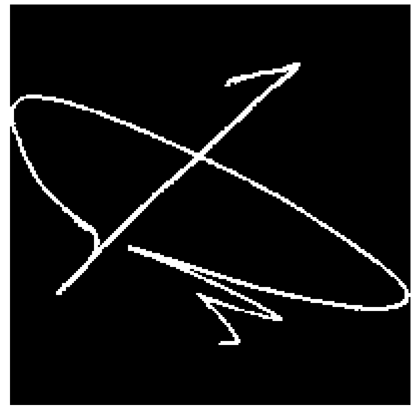
Figure 5.3: Images resulting from the XORing with the reference signature.



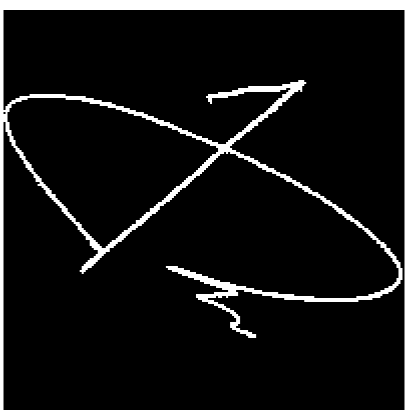
Forged Signature #1



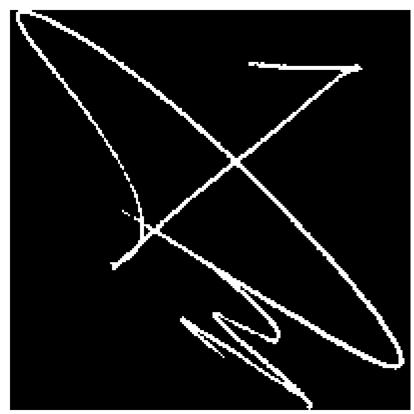
Forged Signature #2



Forged Signature #3

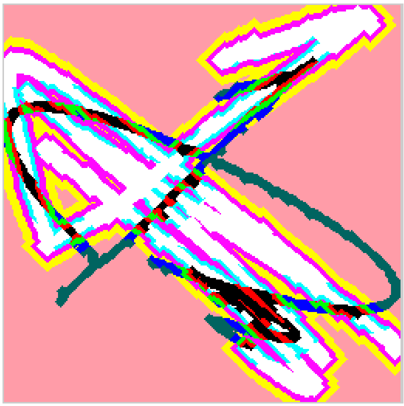
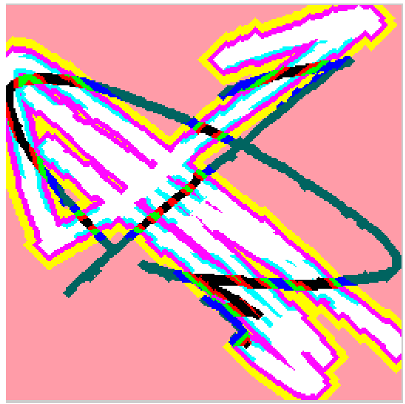
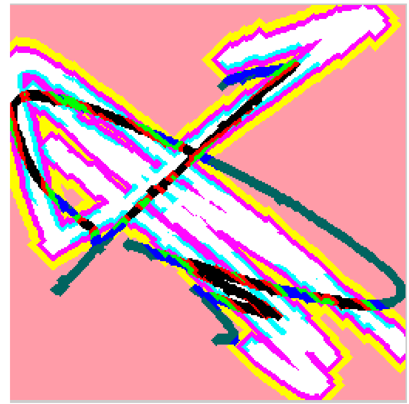


Forged Signature #4



Forged Signature #5

Figure 5.4: Five forged signatures mimicking the genuine ones.

		
G.S # 1 xor F.S # 1	**G.S # 1 xor F.S # 2**	**G.S # 1 xor F.S # 3**

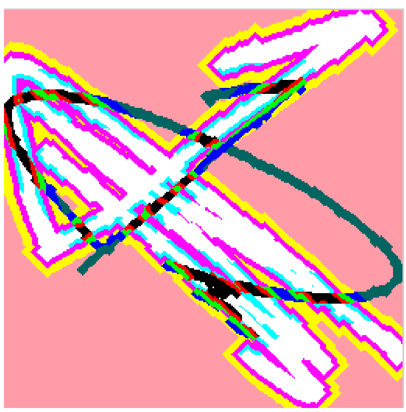
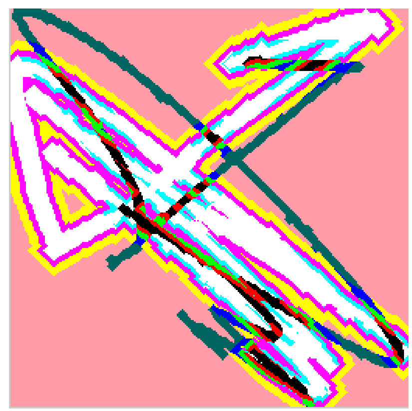
	
G.S # 1 xor F.S # 4	**G.S # 1 xor F.S # 5**

Figure 5.5: Images resulting from the XORing with the reference signature.

5.5.2 The set of vectors that span the column space of $\boldsymbol{G}$

After row reduction, the linearly independent vectors $\boldsymbol{u}_l \quad \boldsymbol{u}_m \quad \boldsymbol{u}_n$ are:

$$\boldsymbol{S} = \begin{bmatrix} 9818.0 & 7886.0 & 9275.0 \\ 1264.0 & 1567.0 & 1365.0 \\ 1786.0 & 2933.0 & 2362.0 \end{bmatrix}$$

$$= [\boldsymbol{u}_1 \quad \boldsymbol{u}_2 \quad \boldsymbol{u}_3]$$

Now:

$$\det(\boldsymbol{S}) = \begin{vmatrix} 9818.0 & 7886.0 & 9275.0 \\ 1264.0 & 1567.0 & 1365.0 \\ 1786.0 & 2933.0 & 2362.0 \end{vmatrix} = 1.14 \times 10^9 \neq 0$$

Which means that these vectors span the column space $\mathcal{C}(\boldsymbol{S})$.

Assuming that $\mathcal{C}(\boldsymbol{G})$ has the Euclidean inner product, we can find:

$$\langle \boldsymbol{u}_1, \boldsymbol{u}_2 \rangle = 8.5 \times 10^7 \neq 0$$

$$\langle \boldsymbol{u}_1, \boldsymbol{u}_3 \rangle = 9.7 \times 10^7 \neq 0$$

$$\langle \boldsymbol{u}_2, \boldsymbol{u}_3 \rangle = 8.2 \times 10^7 \neq 0$$

Therefore the column vectors of $\boldsymbol{S}$ do not form an orthogonal basis. Using Gram-Schmidt process we find from $\boldsymbol{u}_1 \quad \boldsymbol{u}_2 \quad \boldsymbol{u}_3$ an orthogonal basis $\boldsymbol{e}_1 \quad \boldsymbol{e}_2 \quad \boldsymbol{e}_3$ that spans the column space of $\boldsymbol{G}$:

$$\boldsymbol{OS} = \begin{bmatrix} 9818.0 & -327.4 & 4.2 \\ 1264.0 & 509.6 & -68.1 \\ 1786.0 & 1438.9 & 25.1 \end{bmatrix}$$

$$= [\boldsymbol{e}_1 \quad \boldsymbol{e}_2 \quad \boldsymbol{e}_3] = \boldsymbol{E}$$

5.5.3 Change of basis

The coordinates of the vectors in $\boldsymbol{G}$ in the orthogonal basis $\boldsymbol{B}_{o1} = \{\boldsymbol{e}_1 \quad \boldsymbol{e}_2 \quad \boldsymbol{e}_3\}$ are given by:

$$[\boldsymbol{u}_k]_{B_{o1}} = \boldsymbol{E}^{-1}\boldsymbol{u}_k = \begin{bmatrix} 0.0001 & 0.0000 & 0.0000 \\ -0.0001 & 0.0002 & 0.0006 \\ 0.0008 & -0.0129 & 0.0047 \end{bmatrix} \boldsymbol{u}_k$$

And the coordinates of the vectors in $\boldsymbol{F}$ in the same orthogonal basis are:

$$[\boldsymbol{o}_k]_{B_{o2}} = \boldsymbol{E}^{-1}\boldsymbol{o}_k$$

For the column vectors in the matrix $\boldsymbol{G}$ we get (see figure 5.6):

$$[\boldsymbol{u}_1]_{B_{o1}} = \boldsymbol{E}^{-1}\boldsymbol{u}_1 = \boldsymbol{E}^{-1}\begin{bmatrix} 9818.0 \\ 1264.0 \\ 1786.0 \end{bmatrix} = \begin{bmatrix} 1.0000 \\ 0.0000 \\ 0.0000 \end{bmatrix}$$

$$[\boldsymbol{u}_2]_{B_{o1}} = \boldsymbol{E}^{-1}\boldsymbol{u}_2 = \boldsymbol{E}^{-1}\begin{bmatrix} 7886.0 \\ 1567.0 \\ 2933.0 \end{bmatrix} = \begin{bmatrix} 0.8366 \\ 1.0000 \\ 0.0000 \end{bmatrix}$$

$$[\boldsymbol{u}_3]_{B_{o1}} = \boldsymbol{E}^{-1}\boldsymbol{u}_3 = \boldsymbol{E}^{-1}\begin{bmatrix} 9275.0 \\ 1365.0 \\ 2362.0 \end{bmatrix} = \begin{bmatrix} 0.9587 \\ 0.4341 \\ 1.0000 \end{bmatrix}$$

$$[\boldsymbol{u}_4]_{B_{o1}} = \boldsymbol{E}^{-1}\boldsymbol{u}_4 = \boldsymbol{E}^{-1}\begin{bmatrix} 10487.0 \\ 949.0 \\ 1561.0 \end{bmatrix} = \begin{bmatrix} 1.0570 \\ -0.2885 \\ 3.5266 \end{bmatrix}$$

$$[\boldsymbol{u}_5]_{B_{o1}} = \boldsymbol{E}^{-1}\boldsymbol{u}_5 = \boldsymbol{E}^{-1}\begin{bmatrix} 8280.0 \\ 335.0 \\ 2225.0 \end{bmatrix} = \begin{bmatrix} 0.8469 \\ 0.2715 \\ 12.8394 \end{bmatrix}$$

For the column vectors in the matrix $\boldsymbol{F}$ we get (see figure 5.7):

$$[\boldsymbol{o}_1]_{B_{o2}} = \boldsymbol{E}^{-\mathbf{1}}\boldsymbol{o}_1 = \boldsymbol{E}^{-1}\begin{bmatrix}10986.0\\1799.0\\1698.0\end{bmatrix} = \begin{bmatrix}1.1185\\-0.0970\\-6.3867\end{bmatrix}$$

$$[\boldsymbol{o}_2]_{B_{o2}} = \boldsymbol{E}^{-\mathbf{1}}\boldsymbol{o}_2 = \boldsymbol{E}^{-1}\begin{bmatrix}12051.0\\2624.0\\1150.0\end{bmatrix} = \begin{bmatrix}1.2224\\-0.3910\\-18.7792\end{bmatrix}$$

$$[\boldsymbol{o}_3]_{B_{o2}} = \boldsymbol{E}^{-\mathbf{1}}\boldsymbol{o}_3 = \boldsymbol{E}^{-1}\begin{bmatrix}10910.0\\1740.0\\1752.0\end{bmatrix} = \begin{bmatrix}1.1113\\-0.0672\\-5.4300\end{bmatrix}$$

$$[\boldsymbol{o}_4]_{B_{o2}} = \boldsymbol{E}^{-\mathbf{1}}\boldsymbol{o}_4 = \boldsymbol{E}^{-1}\begin{bmatrix}12029.0\\1701.0\\1144.0\end{bmatrix} = \begin{bmatrix}1.2087\\-0.5846\\-6.9226\end{bmatrix}$$

$$[\boldsymbol{o}_5]_{B_{o2}} = \boldsymbol{E}^{-\mathbf{1}}\boldsymbol{o}_5 = \boldsymbol{E}^{-1}\begin{bmatrix}11209.0\\3040.0\\1491.0\end{bmatrix} = \begin{bmatrix}1.1520\\0.0103\\-23.1952\end{bmatrix}$$

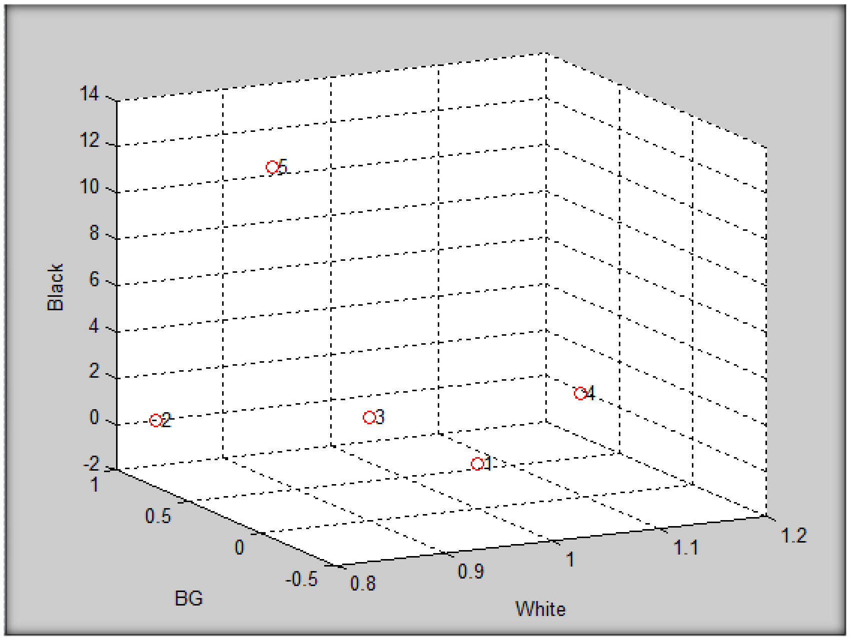


Figure 5.6: Representation of the points of the matrix $\boldsymbol{G}$.

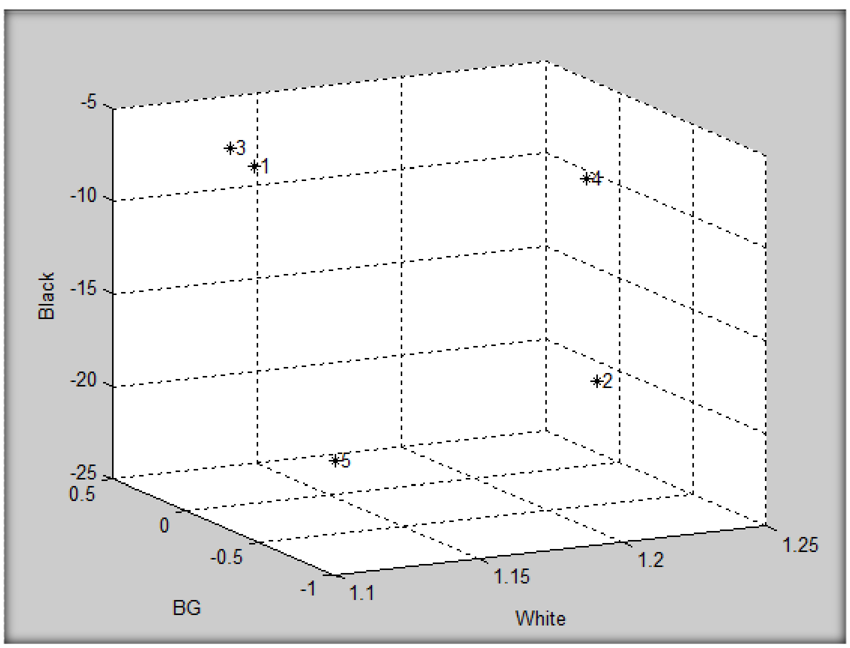


Figure 5.7: Representation of the points of the matrix $\boldsymbol{F}$.

The equation of the sphere (see figure 5.8) that encapsulates all the points related to the matrix of genuine signatures, $\boldsymbol{G}$, is:

$$(\boldsymbol{x}-\boldsymbol{x}_0)^2+(\boldsymbol{y}-\boldsymbol{y}_0)^2+(\boldsymbol{z}-\boldsymbol{z}_0)^2=\boldsymbol{r}^2 \quad (5.5)$$

$$(\mathbf{x}-\mathbf{0.9398})^2+(\mathbf{y}-\mathbf{0.2834})^2+(\mathbf{z}-\mathbf{3.4732})^2=(\mathbf{9.3667})^2 \quad (5.6)$$

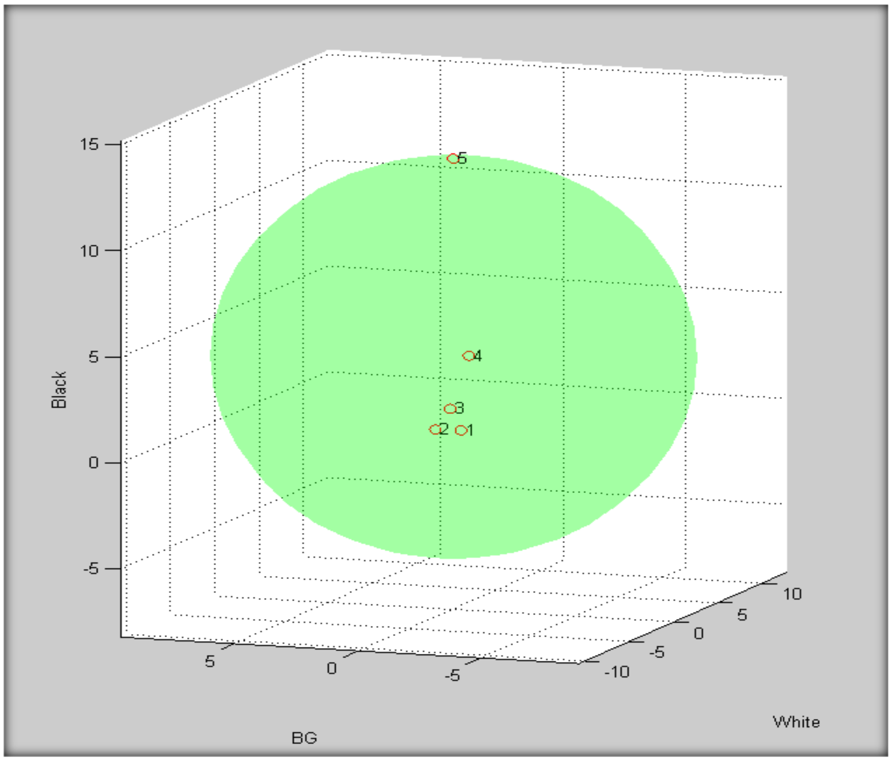


Figure 5.8: Representation of the sphere of Genuine Signatures.

While the equation of the sphere (see figure 5.9) that encapsulates all the points related to the matrix of forged signatures, $\boldsymbol{F}$, is:

$$(\mathbf{x}-\mathbf{1.1626})^2+(\mathbf{y}+\mathbf{0.2259})^2+(\mathbf{z}+\mathbf{12.1427})^2=(\mathbf{11.0550})^2 \quad (5.7)$$

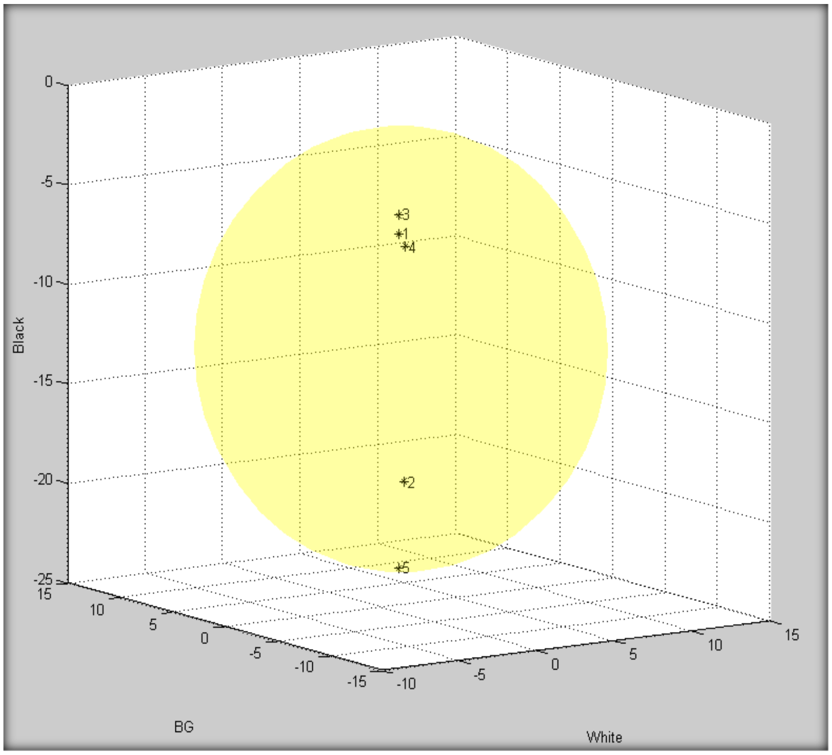


Figure 5.9: Representation of the sphere of Forged Signatures.

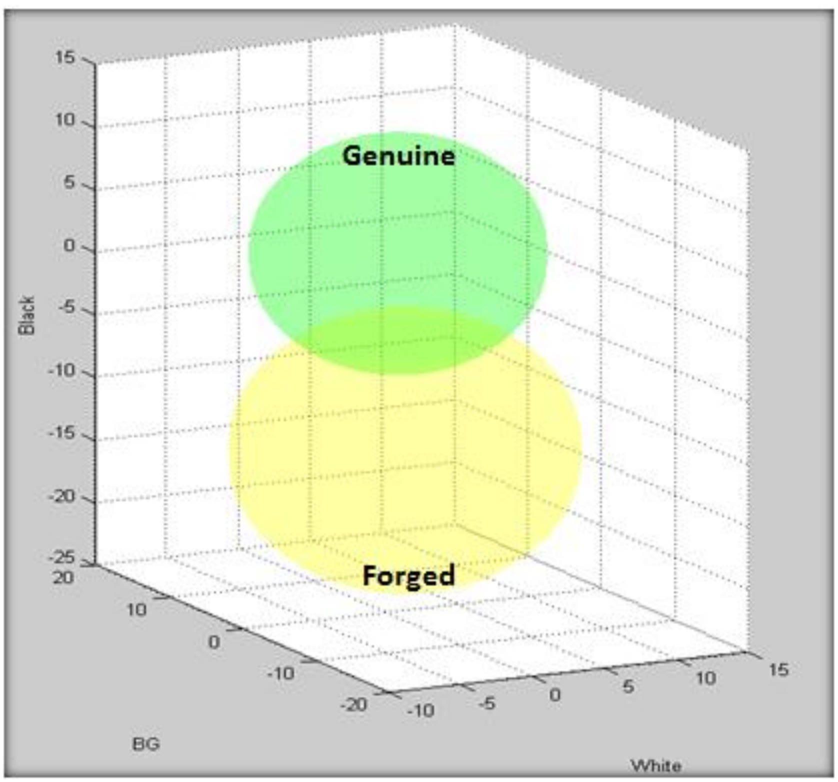


Figure 5.10: The two spheres representation.

5.6 The separation plane

As can be seen in figure 5.10 and after representing the two spheres together, a critical issue directly appears represented in the intersection region between them, which makes the verification process of a new appearance harder, especially if it is located in that area, because of the inability of making a decision whether it refers to the region of genuine signatures, or to the other region related to the forged ones. A separation plane between the two spheres will be calculated, but initially the intersection points between the two spheres must be found. In order to facilitate the calculations, a translation of the whole structure (the two intersecting spheres) is wanted, so that the centre of the sphere with the higher radius is at the origin of the vector space of the orthogonal basis in $\mathbb{R}^{\mathbf{3}}$, as can be seen in figure 5.11 hereafter.

Then the two spheres are rotated in $\mathbb{R}^3$, where the rotation is used to bring the centre of the other sphere (the sphere of the shorter radius) to the x-axis, as shown in figure 5.12. The condition for this is to have the distance D between the centers of the two spheres constant and non-zero.

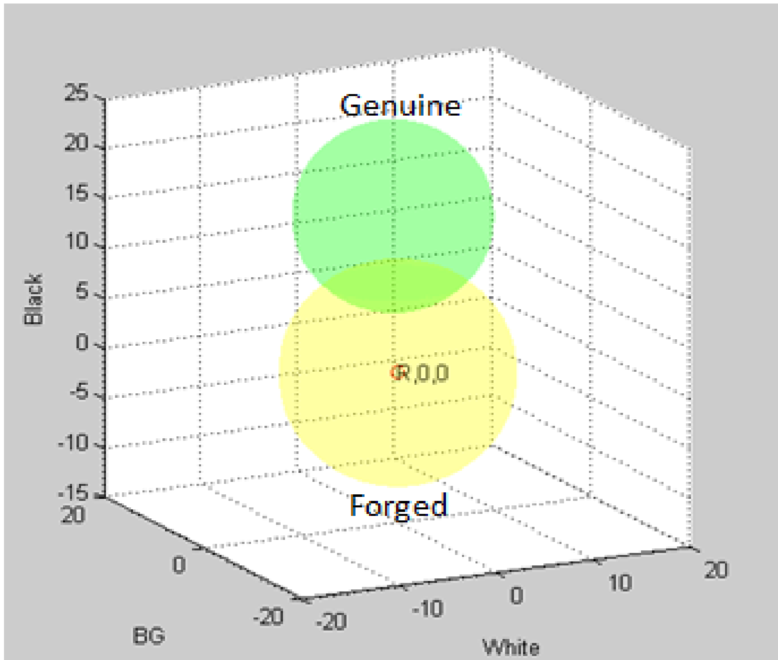


Figure 5.11: The two spheres after translation.

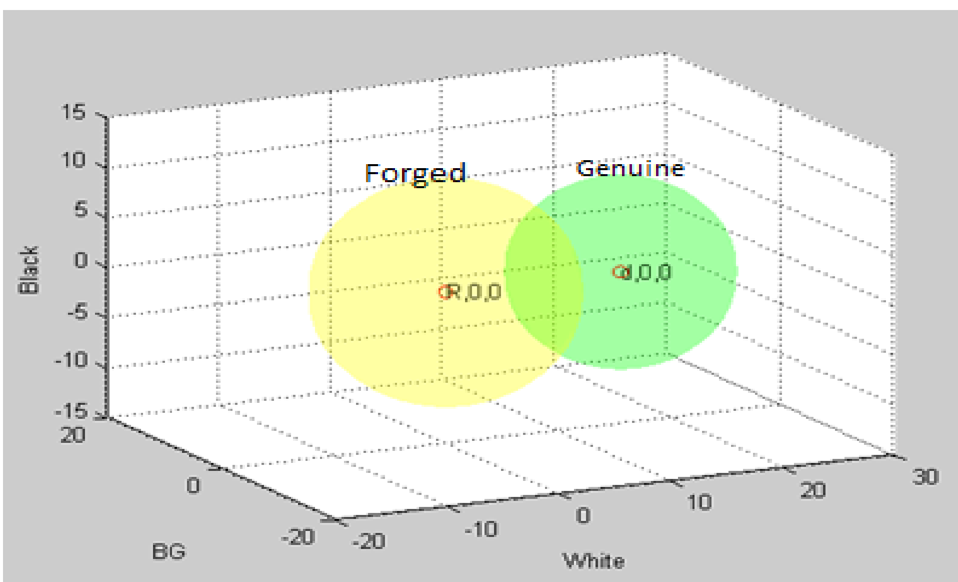


Figure 5.12: The two spheres after translation and rotation.

The two spheres are now in a suitable lay that facilitates the process of finding three intersection points between them, as shown in figure 5.13.

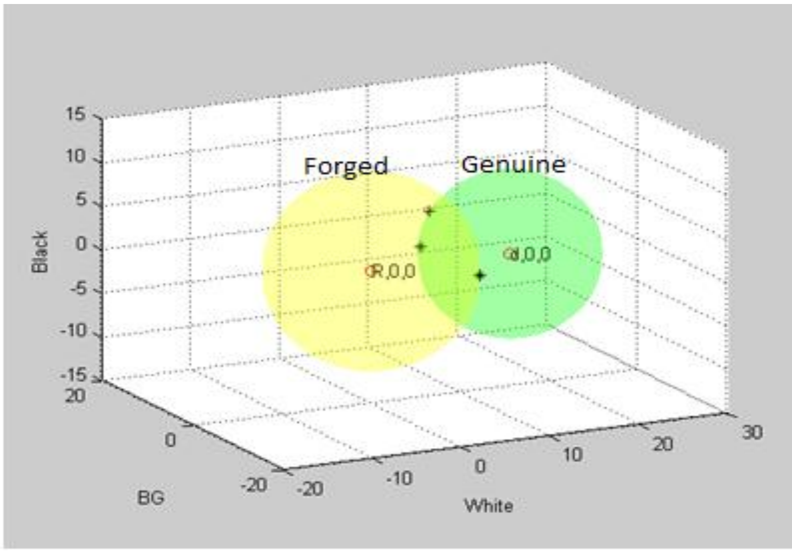


Figure 5.13: Three intersection points between the spheres.

Once those intersection points are found, a separation plane between them can be defined and represented as can be seen in figure 5.14.

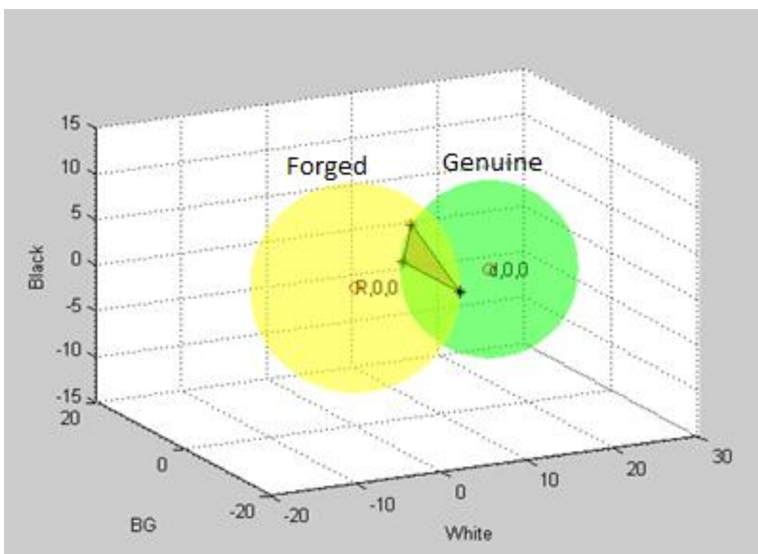


Figure 5.14: The separation plane between the spheres.

5.6.1 Testing a new genuine signature

If we choose a new genuine signature of the abandoned ones related to the same person in this example, and after applying all the pre-processing steps on this new appearance, followed by overlapping (XORing) it with the reference genuine signature, the resulting image is shown in figure 5.15.

After that, the vector coordinates of this new appearance are viewed in the WB_GB_K three dimensional space, using the orthogonal basis that spans the column space of $\boldsymbol{G}$. Figure 5.16 clarifies the location of the new appearance with respect to the separation plane between the two spheres, since it appears to the right side of it and in the area related to the genuine sphere, the algorithm verifies this new appearance as a genuine signature, which is a correct verification decision.

5.6.2 Testing a new forged signature

A new forged signature of the abandoned four ones related to the same person in this example are chosen, and after applying all the pre-processing steps on this new appearance, followed by overlapping (XORing) it with the first reference genuine signature as can be seen in figure 5.17. Then the vector coordinates of this new appearance are viewed in the WB_GB_K three dimensional space, using the orthogonal basis that spans the column space of $\boldsymbol{G}$.

This new appearance manifests to the left side of the separation plane and in the area related to the forged sphere, as can be seen in figure 5.18; promoting the algorithm to verify it as a forged signature, which is a correct verification decision.

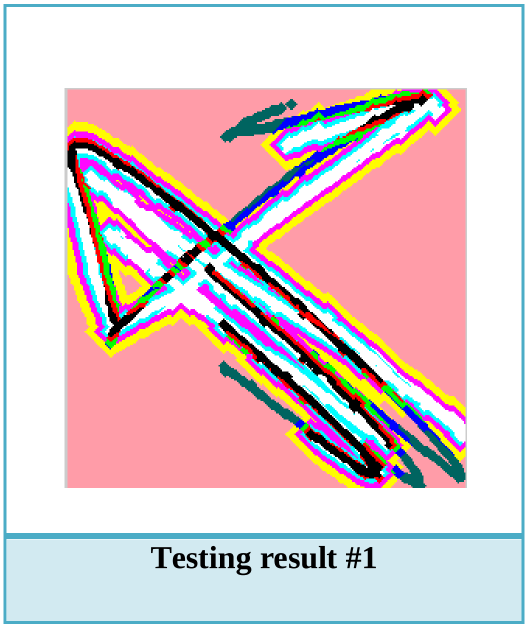


Figure 5.15: The resulting genuine testing image.

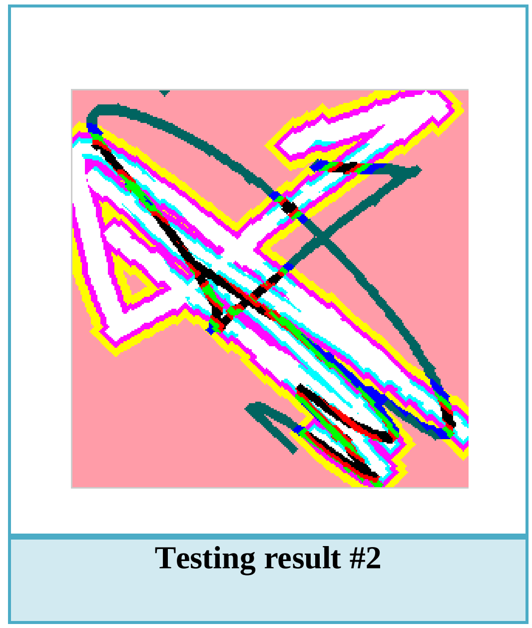


Figure 5.17: The resulting forged testing image.

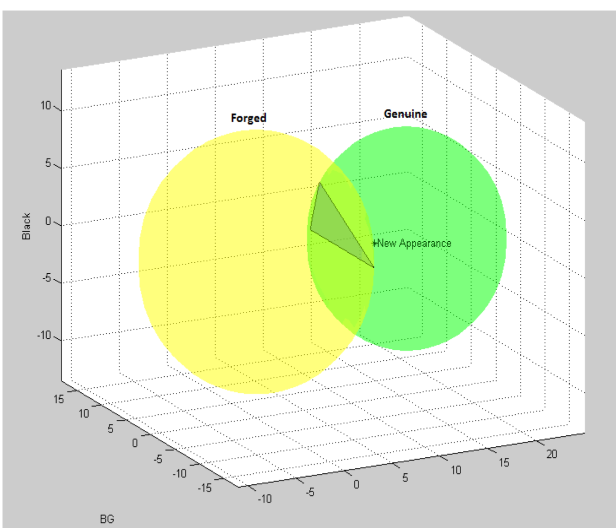


Figure 5.16: Genuine signature representation

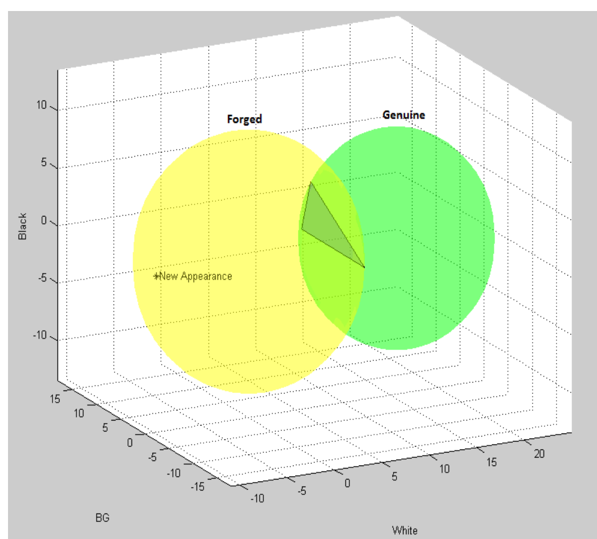


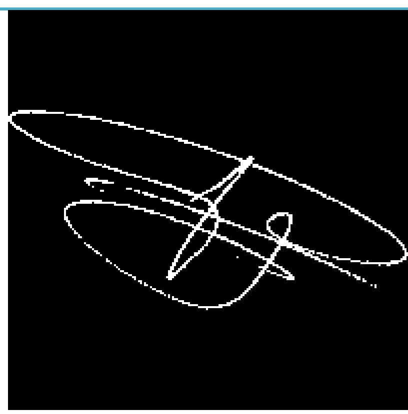
Figure 5.18: Forged signature representation.

5.7 RGB space

The other three dimensional space on which the separability of the signatures will be also studied is built using the group that is composed of the other three colour codes, which are the Red/Green/Blue (RGB). The RGB space is composed of two clusters of signatures per person, the genuine cluster and the forged cluster. The overall previously annotated process is repeated here, except that the resulting matrices of both genuine and forged signatures XORing processes are composed of RGB pixels of appearances.

5.7.1 An illustrative example

Consider the first genuine signature from the database shown in the figure below.



Genuine Signature #1

Figure 5.19: Somebody's first genuine signature in the database.

This genuine signature is considered as the reference to which all other signatures of that person are compared to, for verification. After XORing this genuine signature with another five genuine signatures related to the same person, and again XORing it with another five forged ones, both of the matrices $\boldsymbol{G}$ and $\boldsymbol{F}$ are then composed. After that a set of vectors that span the column space of $\boldsymbol{G}$ are founded, and an orthogonal basis of them is computed also. Then the column vectors of both $\boldsymbol{G}$ and $\boldsymbol{F}$ are transformed into the three-dimensional space spanned by that orthogonal basis. As can be seen in figure 5.20 and after representing the spheres related to the RGB colours for both of $\boldsymbol{G}$ and $\boldsymbol{F}$ matrices, we can notice that it will be impossible to find the intersection points between the two spheres while the genuine sphere here is totally encapsulated by the forged one, subsequently a separation plane could not be defined here.

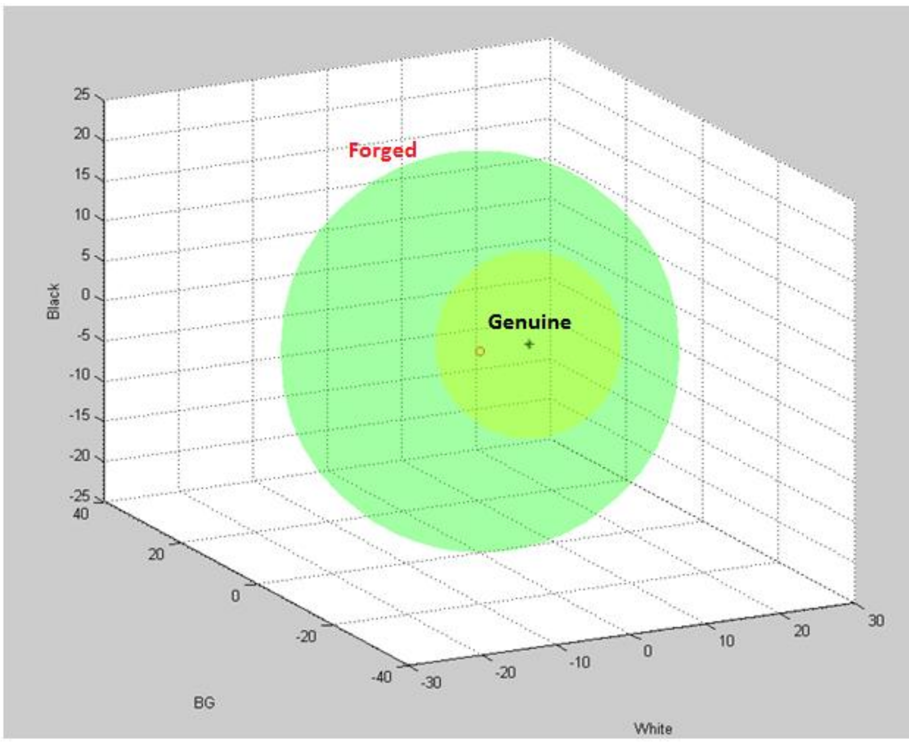


Figure 5.20: The two overlapped spheres after translation and rotation.

5.8 Artificial Neural Networks (ANNs)

One of the powerful data modelling tools that has the ability of capturing and representing complex input/output relationships is the Artificial Neural Networks (ANNs). The development of the ANNs technology was in response to the urgent desire of developing an artificial system that could perform intelligent tasks identical to those performed by the human brain. The actual benefit of the ANNs lies in their ability of learning different relationships from the data being modelled, and then representing those relationships either they are linear or non-linear ones. [21]

For more than half a century the ANNs have been used in a wide range of applications specially the computerised pattern recognition ones, according to their ease of use because of the existence of different training algorithms that are able to learn the underlying structure of any data set after collecting it. Therefore a new model for Signature Verification is presented using the ANNs architecture. [22]

5.9 Signature Verification using ANNs

5.9.1 Training the network

A two-layer feed-forward neural network model has been designed using the Neural Network Fitting Tool (*nftool*). The two hundred set of features which are the colour codes (White, Background, and Black pixels) that represent different degrees of overlapping between any two of the pre-processed signatures, are the inputs of the training phase of the network, which are used to classify the signatures into genuine and forged ones. The inserted set of features was then automatically and randomly divided into three different kinds of samples; training, validation and testing.

After launching the network we have to set the number of the hidden neurons in the fitting network's hidden layer; where the determination of the optimal number of those neurons is a crucial issue. If it is too small, the network cannot possess sufficient information, and thus yields inaccurate results. On the other hand, if it is too large, the training process will be very long. [23]

According to that and after trying different values for the hidden neurons, a moderate value that equals twenty five was chosen, besides training the network by Levenberg-Marquardt back-propagation algorithm several times each time increasing the number of iterations until it reaches the minimum error with forty nine ones (shown in figure 5.21). The value of the Regression that measures the correlation between the outputs and the targets equals approximately 0.84 as can be seen in figure 5.22.

5.9.2 Evaluating the Accuracy of the Network

As can be seen in figure 5.23 the value of the mean square error that evaluates the average squared difference between the outputs and the targets equals roughly 0.05; which is to some extent the best value we get of training the network. Other measures were calculated to evaluate the network's accuracy. The first one is the Mean absolute percentage error (MAPE) that is a relative measure which expresses errors as a percentage of the actual data [24]. The second one which is a statistical measurement of the relationship between two outputs is the correlation coefficient (r) [25]. These two measurements are calculated by the equations given below, and their values are illustrated in table 5.1.

$$MAPE = \frac{\sum\left|\frac{\mathrm{X_t}-\mathrm{F_t}}{\mathrm{X_t}}\right|}{\mathrm{m}}\,(100) \tag{5.8}$$

Where X_t is the actual value and F_t is the predicted value.

$$r = \frac{\sum(x_i - x_m)(y_i - \mathrm{y_m})}{\sqrt{\sum_i (x_i - x_m)^2}\,\sqrt{\sum_i (y_i - y_m)^2}} \tag{5.9}$$

Where x_m is the mean value of x_i, and y_m is the mean value of y_i.

Table 5.1: Accuracy of the ANNs.

MAPE	48.51
Correlation	0.834

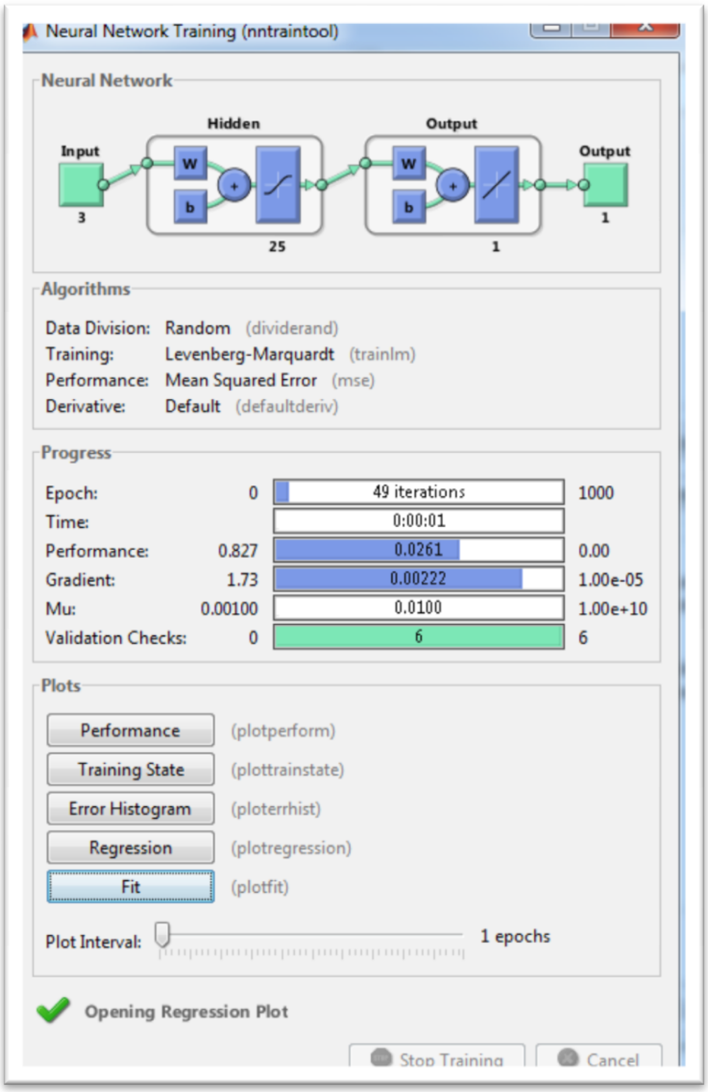


Figure 5.21: Neural Network training.

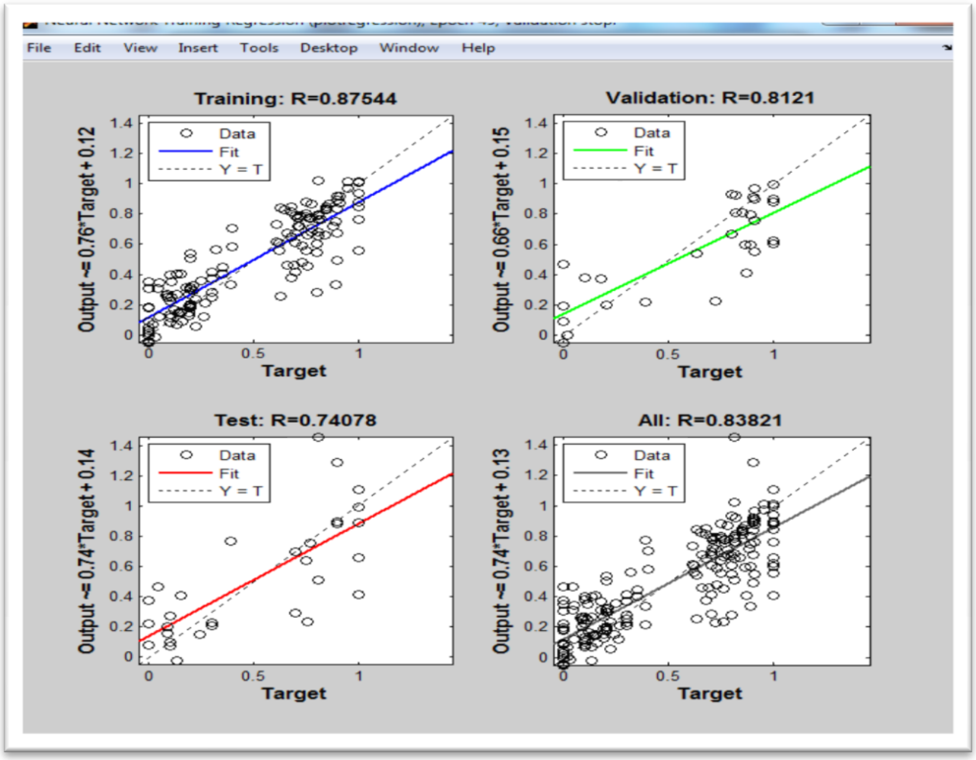


Figure 5.22: Template of the Regression plot.

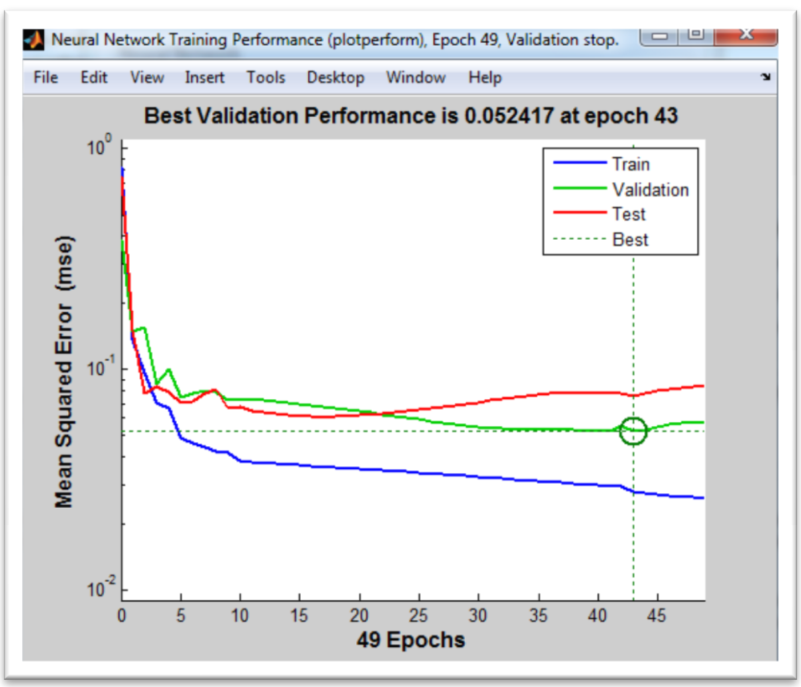


Figure 5.23: Performance graph of training the network.

5.10 Adaptive Neuro-Fuzzy Inference Systems (ANFIS)

One of the new Artificial Intelligence (AI) tools is the Adaptive Neuro-Fuzzy Inference Systems (ANFIS). The ANFIS are considered as a class derived from a general category of intelligent networks known as adaptive networks, which are consisting of a set of nodes that are connected by directional links. In addition to that, the ANFIS integrate the best features of both the ANNs and the fuzzy logic systems, and use a hybrid learning algorithm. [26]

5.11 Signature Verification using ANFIS

5.11.1 The ANFIS model

Again the same inputs that were used to train the ANNs are used to build the ANFIS model, setting the Membership Functions (MFs) for the input variables to be of type (Gauss), while the output is chosen to be (Linear). Figure 5.24 presents the structure of our developed ANFIS, where the number of MFs assigned to each input variable was seven.

5.11.2 Evaluating the Accuracy of the model

When the ANFIS model reaches an acceptable satisfactory level, the predictions from our model and the original target values for both the training and testing data set are presented in figures 5.25 and 5.26 respectively. The aim of increasing the number of epochs to one hundred was to minimize the value of the system's error to the lowest possible value, which can be seen in figure 5.27.

The correlation coefficient (r) and the Mean Absolute Percentage Error (MAPE) were calculated also to evaluate the model's accuracy. The values of these measures are shown in table 5.2.

Table 5.2: Accuracy of ANFIS.

MAPE	47.598
Correlation	0.8597

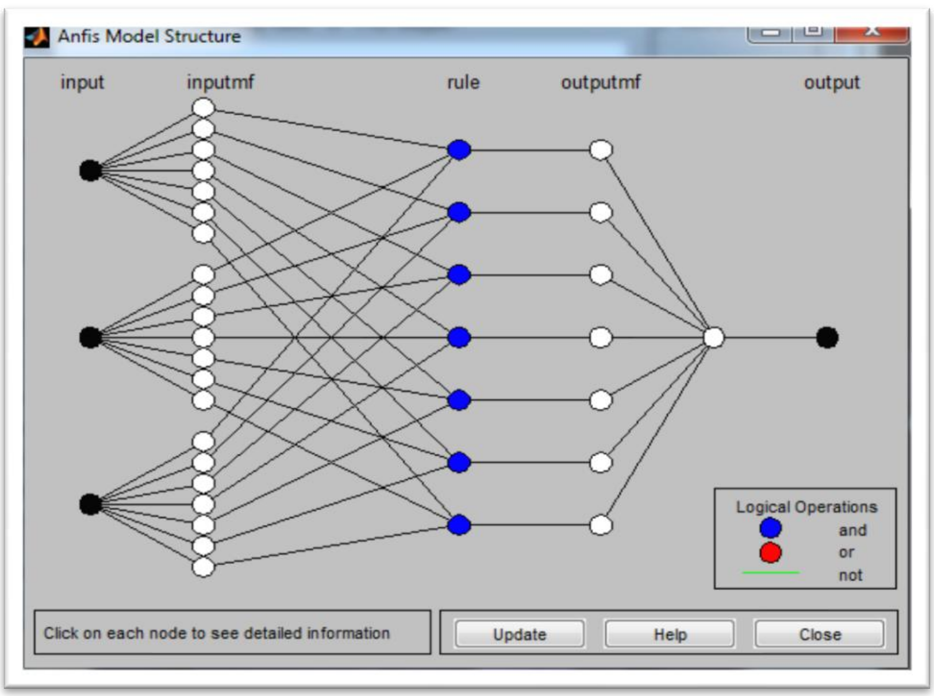


Figure 5.24: Structure of the proposed ANFIS model.

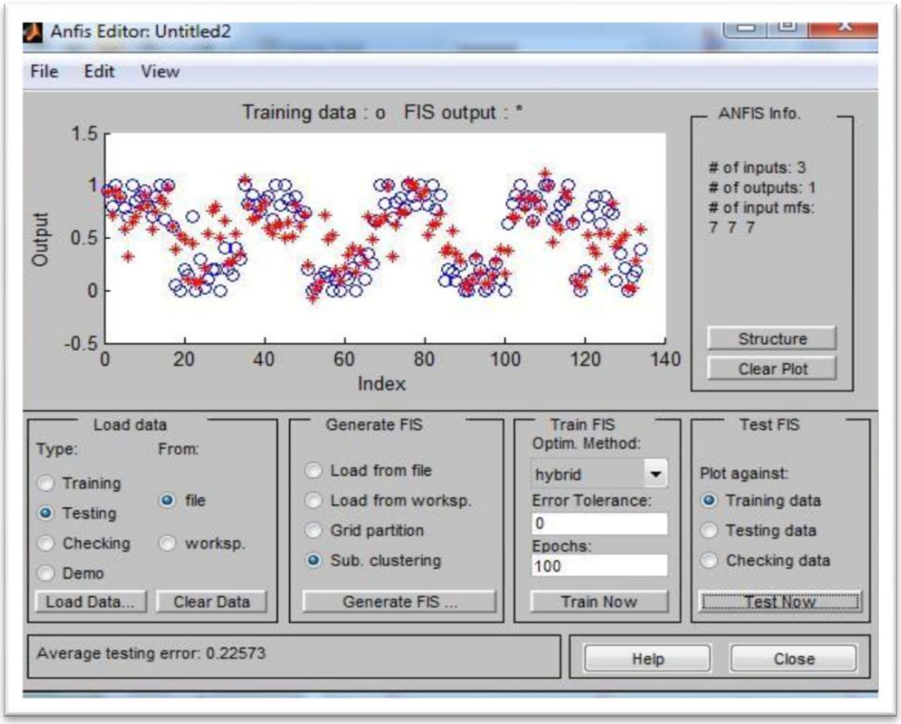


Figure 5.25: Testing the ANFIS against training data set.

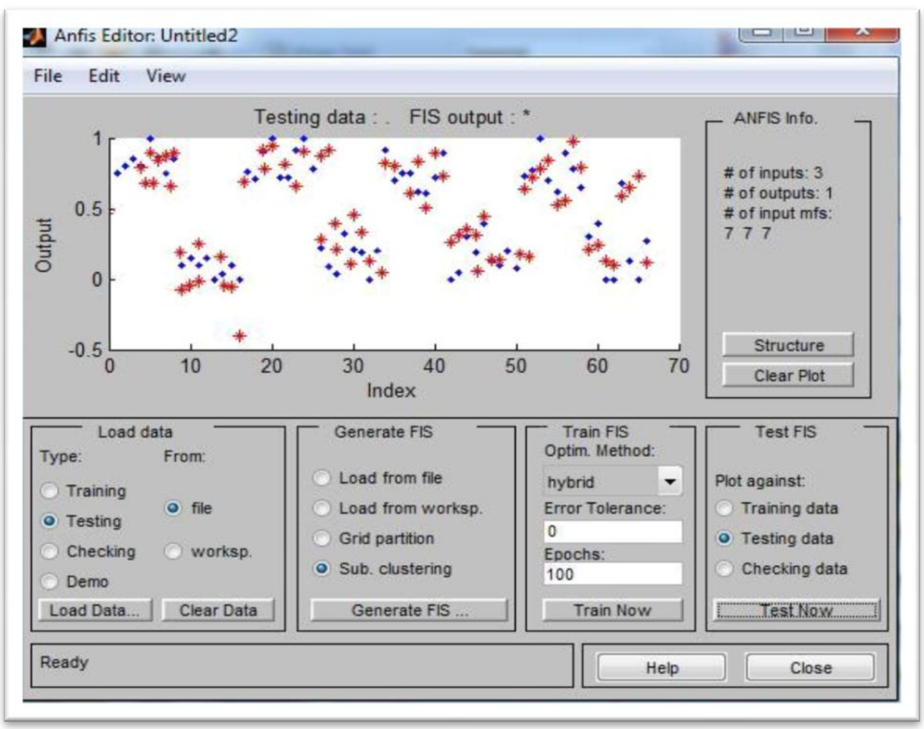


Figure 5.26: Testing the ANFIS against testing data set.

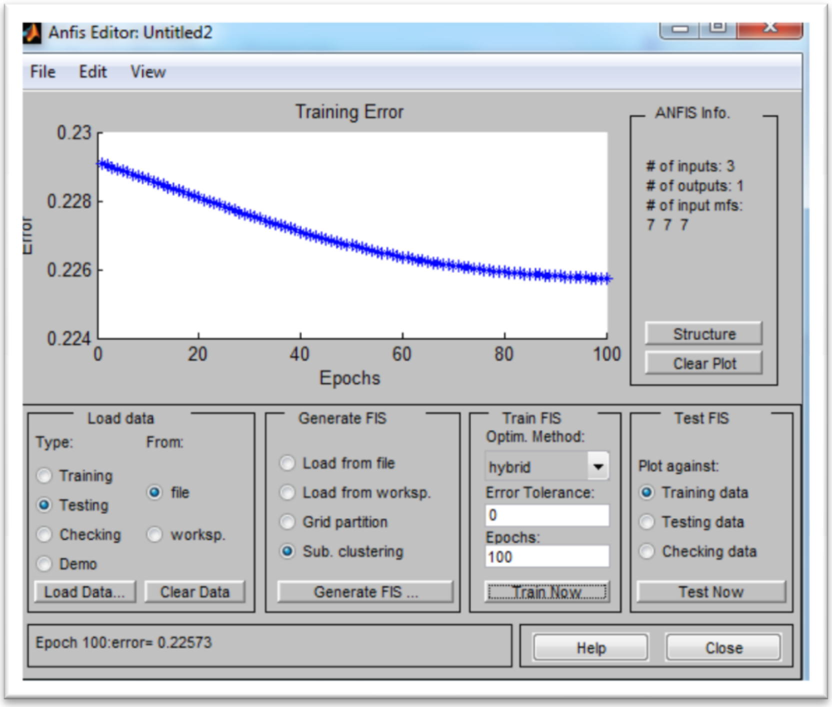


Figure 5.27: The training error of the ANFIS.

5.12 Accuracy analysis of both ANNs and ANFIS

Figure 5.28 and figure 5.29 illustrate that the value of MAPE is lower for the ANFIS model in comparison with the ANNs model, at the same time the value of the correlation coefficient related to the ANNs is lower than that of the ANFIS. Those values ensure that the predictions of the ANFIS model are closer to our eventual outcomes, which means that the accuracy of the ANFIS model is higher than the accuracy of other one.

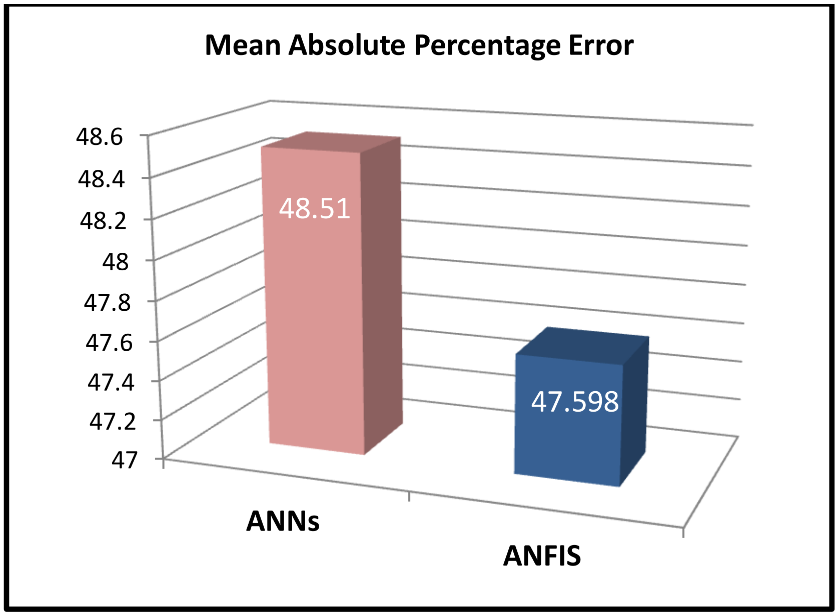


Figure 5.28: The MAPE value of the two models.

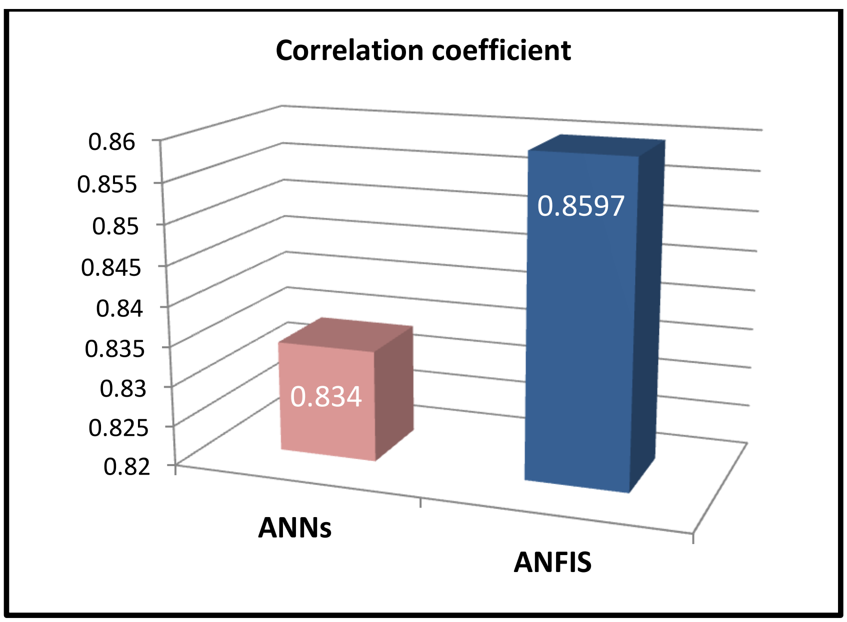


Figure 5.29: The Correlation coefficient value of the two models.

CHAPTER 6

RESULTS AND DISCUSSION

6.1 Introduction

The aim of a signature verification system is to accurately discriminate between two categories of signatures, namely genuine and forged signatures. False Acceptance Rate (FAR), False Rejection Rate (FRR) and Correct Classification Rate are the used parameters to evaluate the performance of any signature verification system [27][28]. FAR, FRR, and CCR are calculated by the equations given below.

$$FAR = \frac{Number\ of\ forgeries\ accepted}{Number\ of\ forgeries\ tested} \times 100 \qquad (6.1)$$

$$FRR = \frac{Number\ of\ genuines\ rejected}{Number\ of\ genuines\ tested} \times 100 \qquad (6.2)$$

$$CCR = \frac{Number\ of\ samples\ correctly\ recognized}{Number\ of\ samples\ tested} \times 100 \qquad (6.3)$$

For each of the twenty persons in the database an arbitrarily genuine signature (reference signature) is being selected, pre-processed and then XORed with five other genuine signatures and five forged ones regarding to the same person, this will result in having 10 sets of features per person, each set contains three different colours counts (White, Background and Black) that reveal the different degrees of similarity between them. Thereafter the same process is repeated for the rest, and then a training set consists of 200 sets of features will be used in the three different approaches. Another 160 sets of features will be used to test the systems, which are the overcome of overlapping the same pre-processed reference signature with the remaining four genuine and four forged signatures per person.

We have to note that the verification of the signatures by the clustering algorithm has been checked also using the other group of colour counts which are the (Red, Green and Blue), but unfortunately and after representing the spheres of the twenty signatures, we found that we could not define the separation plane for seven of them; because of the overlapping of the spheres with each other, the same as we have seen in chapter 3. According to that, only the WBgBk colour counts will be used for training and testing both ANNs and ANFIS models.

6.2 Results of testing Clustering algorithm

The performance and the effectiveness of the clustering algorithm have been tested using the WB_GB_k set of features extracted from the 160 pre-processed signatures samples, which are not used for the definition of the genuine and forged spheres per person. Those signatures (New Appearances) are separated into 80 genuine signatures, and another 80 forged ones. The location of each one of them according to the separation plane is checked, and accordingly a decision of whether it's been a genuine or forged signature is taken. After finishing the testing process, we found that only three of the eighty forged signatures were accepted, besides rejecting five signatures of the eighty genuine ones. Table 6.1 shows the values of FAR, FRR, and CCR for the clustering algorithm.

Table 6.1: Results of testing Clustering algorithm with new signatures samples

FAR	3.75%
FRR	6.25%
CCR	95%

6.3 Results of testing ANNs

The 200 sets of features were used in the training phase of the network, and the other 160 were used in the testing phase. Twenty one of the eighty forged signatures were classified as genuine ones; while twenty four of the genuine signatures were classified as forged. Table 6.2 shows the values of FAR, FRR, and CCR for the ANNs.

Table 6.2: Results of testing ANNs with new signatures samples

FAR	26.25%
FRR	30%
CCR	71.9%

6.4 Results of testing ANFIS

Once again the set of features that are extracted from the 160 pre-processed signatures samples, are separated into 80 genuine signatures, and another 80 forged ones, and used for testing the performance of the ANFIS. 19 of the 80 forged signatures were classified as genuine ones, while 10 of the genuine signatures were classified as forged. Table 6.3 shows the values of FAR, FRR, and CCR for the ANFIS.

Table 6.3: Results of testing the ANFIS with new signatures samples

FAR	23.75%
FRR	12.5%
CCR	81.875%

6.5 Performance analysis of the signature verification techniques

Figure 6.1 illustrates that the signature verification system which was designed based on the algebraic clustering algorithm presents the best performance in comparison to the others, and has a higher precision in classification, as the FAR and FRR values related to it are the lowest, and its CCR value is the highest. On the other hand, the figure also clarifies that the ANFIS comes after the clustering algorithm in the verification process, while the ANNs is the least accurate technique in verification.

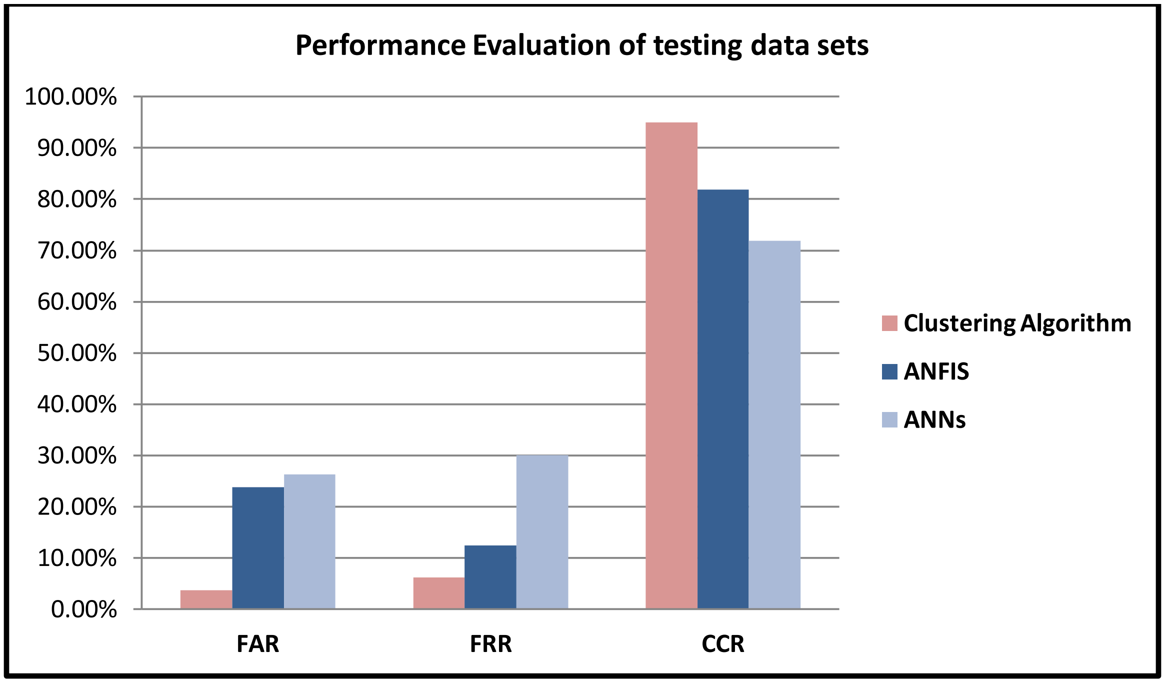


Figure 6.1: Comparing three signature verification methods in terms of FAR, FRR, and CCR.

CHAPTER 7

CONCLUSION AND FUTURE WORK

7.1 Conclusion

In this research we have devised an offline signature verification system; the proposed system uses different features represented in the count of specific coloured pixels in a generated check pattern, which results from the overlapping (XORing) process between a reference signature and another genuine or forged one. The verification process is then performed using three different techniques that are the spatial clustering in a three dimensional Euclidean space, the Artificial Neural Networks and the Adaptive Neuro-Fuzzy Inference System.

It is observed from the results of the described experiments that the colour code three dimensional Euclidian space classifies most of the signatures in a correct way. When compared to the other methods in terms of FAR, FRR and CCR it was found that it gave better results. A detailed comparison of the results achieved for the systems using the same database is provided in chapter 6. Since no standardized signature database is available on which all systems are evaluated, a direct comparison of performances achieved is generally not possible.

When used for the front office verification, the proposed system requires 0.415402 seconds per signature; this is real time verification compared to the time a person in the front office takes to undertake the same process. To this time we need to add other counts related to scanning, rotation and skewing. To extract the set of features for 11 signatures per person in a database, 11.66421 seconds are necessary. For the database used for this thesis, building-up the space and identifying the separation plane require 2.924384 seconds. This is obviously an acceptable time value, noting that these processes are done only once for each person, and their duration is not a critical issue even if we later include other pre-processing operations such as scanning, rotation and skewing; these are done in the back office.

7.2 Future Work

A first step is to use a larger data set which contains signatures for more reference writers under various different psychological and physical conditions. This will allow more variations in the training and testing phases for the different techniques of verification, which will probably lead to more interesting results. Furthermore we can extract more features in addition to the counts of coloured pixels, which means that a higher dimensional Euclidean space can be obtained. This means we need to define new features (static properties). Whatever dimension we decide to work with for the set of features Euclidean space, there is an intersection between the hyper-spheres of genuine and forged signatures. In order to deal with this issue we suggest to use the simplex method in linear programming.

Eventually, we can use the wavelets in order to transfer the data (set of features) from its original domain into another one making them smoother by eliminating noise that results from digitization. This should be done in the pre-processing phase of the algorithm.

In literature Hidden Markov Models (HMM) seems to outperform every other tool as it achieves a recognition performance of 99.2%. For all what have been mentioned here above, the results should be compared to those obtained using HMM, which would require the design of a tool, based on HMM algorithms as described in published articles. [11]

REFERENCES

[1] Simon Liu and Mark Silverman, "A Practical Guide to Biometric Security Technology", *IEEE*, pp.27-32, January/February 2001.

[2] Alisher Anatolyevich Kholmatov, "Biometric Identity Verification Using On-Line & Off-Line Signature Verification", Master Thesis, Sabanci University, Spring 2003.

[3] Anil Jain, A., Lin Hong, and Sharath Pankanti, "Biometric Identification", *Communications of the ACM*, Vol. 43, No. 2, pp. 91-98, February 2000.

[4] Emre Özgündüz, Tülin Şentürk and M. Elif Karslıgil, "Off-Line Signature Verification and Recognition By Support Vector Machine", *CAIP 2005*, pp.799-805, 2005.

[5] Anil K. Jain , Friederike D. Griess and Scott D. Connell, "On-line Signature Verification", Pattern Recognition, Vol. 35, pp. 2963–2972, 2002.

[6] Zhaoxiang Zhang, Kaiyue Wang and Yunhong Wang, "A Survey of On-line Signature Verification", Biometric Recognition, Vol 7098, pp 141-149, 2011.

[7] Donato Impedovo and Giuseppe Pirlo, “Automatic Signature Verification: The State of the Art”, *IEEE Transactions on systems*, Vol. 38, No. 5, pp 609-635, September 2008.

[8] Dr. H B Kekre, V A Bharadi, and A A Ambardekar, “Signature Recognition by a Novel and Simple Contour Technique”, 2007.

[9] Tim Morris, Computer Vision and Image Processing, Palgrave Macmillan, 2004.

[10] Manasjyoti Bhuyan, Kandarpa Kumar Sarma, and Hirendra Das, “Signature Recognition and Verification using Hybrid Features and Clustered Artificial Neural Network (ANN)s”, *International Journal of Electrical and Computer Engineering*, Vol.5:2, pp.128-133, 2010.

[11] S. Adebayo Daramola and T. Samuel Ibiyemi, “Offline Signature Recognition using Hidden Markov Model (HMM)”, *International Journal of Computer Applications* (0975 – 8887), Vol.10, No.2, November 2010.

[12] Kai Huang and Hong Yan,” Off-Line Signature Verification Based On Geometric Feature Extraction and Neural Network Classification”, Pattern Recognition, Vol.30, No.1, pp. 9-17, 1997.

[13] Mehdi Radmehr, Seyed Mahmoud Anisheh, Mohsen Nikpour and Abbas Yaseri, “Designing an Offline Method for Signature Recognition”, *World Applied Sciences Journal*, Vol.13, No.3, pp. 438-443, 2011.

[14] Sepideh Afsardoost , Siamak Yousefi, and Mohammad Ali Khorshidi , “Offline Signature Verification Using Geometric Center Features”, *ICSP*, pp. 1491-1494, 2008.

[15] Ramachandra C, Jyothi Srinivasa Rao, K B Raja, K R Venugopla, and L M Patnaik, “Robust Offline Signature Verification Based On Global Features”, *International Advance Computing Conference (IACC 2009)*, IEEE, March 2009.

[16] Ali Karouni, Bassam Daya, and Samia Bahlak, “Offline Signature Recognition Using Neural Networks Approach”, *Procedia Computer Science* , pp 155–161, 2011.

[17] Gisela Klette, “Skeletons in Digital Image Processing”, *Computer Science Department of the University of Auckland* CITR *at Tamaki Campus*, July 2002.

[18] Vahid Malekian, Alireza Aghaei, Mahdie Rezaeian and Mahmood Alian, “Rapid Off-line Signature Verification Based on Signature Envelope and Adaptive Density Partitioning”, *IEEE*, 2013.

[19] Vinayak Balkrishana Kulkarni, "A colour Code Algorithm for Signature Recognition," *Electronic Letters on Computer Vision and Image Analysis*, Vol.6, No.1, pp.1-12, 2007.

[20] Persian Handwritten Signature database, with permission of Mohamad Hoseyn Sigari, Ph.D. student, Ferdowsi University of Mashhad (FUM), 6th Feb 2012.

[21] Bradley Schafer,and Serestina Viriry, “An Offline Signature Verification System”, *IEEE International Conference on Signals and Image Processing Application,* 2009.

[22] Alan McCabe, Jarrod Trevathan, and Wayne Read, “Neural Network-based Handwritten Signature Verification”, *Journal of Computers,* Vol.3, No.8, pp.9-21, August 2008.

[23] O.C Abikoye, M.A Mabayoje, and R.Ajibade, “Offline Signature Recognition & Verification using Neural Network”, *International Journal of Computer Applications (0975–8887),* Vol.35, No.2, pp.44-51, December 2011.

[24] S. Makridakis and M. Hibon, “Evaluating Accuracy (Or Error) Measures”, Printed at INSEAD.

[25] Eugene K. Yen and Roger G. Johnston, “The Ineffectiveness of the Correlation Coefficient for Image Comparisons”, *Los Alamos National Laboratory*.

[26] M.N. SYED-AHMAD, Ahmed Bensenouci, Saleh A. Alghamdi and A. M. Abdel Ghany, “Short Term Load Forecasting using Adaptive Neuro-Fuzzy Inference System (ANFIS) Application to Aleppo Load Demand”, *Electrical Technology Department, College of Technology at Al-Baha*, pp. 1-5.

[27] Madhavi D.Malekar and Prof.Sachen Patel, “Off-line Signature Verification Using Artificial Neural Network”, *International Journal of Emerging Technology and Advanced Engineering*, Vol.3, September 2013.

[28] Ashwinin Pansare and Shalini Bhartia, “Off-line Signature Verification Using Neural Network”, *International Journal of Scientific and Engineering Research*, Vol.2, February 2012.