CHAPTER 1

INTRODUCTION

The security requirements of the today's society have placed biometrics at the centre of a large debate, as it becomes a key aspect in multitude of applications. Biometrics measure individuals' unique physical or behavioural characteristics to recognize or authenticate their identity. Common physical biometrics include fingerprints; hand or palm geometry; and retina, iris, or facial characteristics. Behavioural characters include signature, voice (which also has a physical component), keystroke pattern, and gait. Of this class of biometrics, technologies for signature and voice are the most developed.

Signature verification is an important examination area in the field of person endorsement. The detection of human script is important concerning about the progress of the interface between human-beings and computers. If the computer is sharp enough to be aware of human handwriting it will provide a more eye-catching and economic man computer interface. In this area signature is a special case that provides secure means for authentication, attestation consent in many high security environment. The goal of the signature verification system is to differentiate between two classes: the original and the forgery, which are related to intra and interpersonal variability. The disparity among signatures of same person is called Intra Personal variation. The variation between originals and forgeries is called Inter Personal Variation.

1.1 Types of Forgeries

There are three different type of forgeries categorization:

Random Forgery- Here name of the victim is used by the signer in his own style to create a forgery known as simple or random forgery. The shape and structure of the original signature is not known by the person who creates the forgery. It is very easy to detect.

Casual Forgery- In casual forgery person didn’t have clear information about the actual signature, but is signing without much practice.

- Skilled Forgery- The last type is the skilled forgery, is signed by a person who has had practiced to sign a genuine signature of other person.

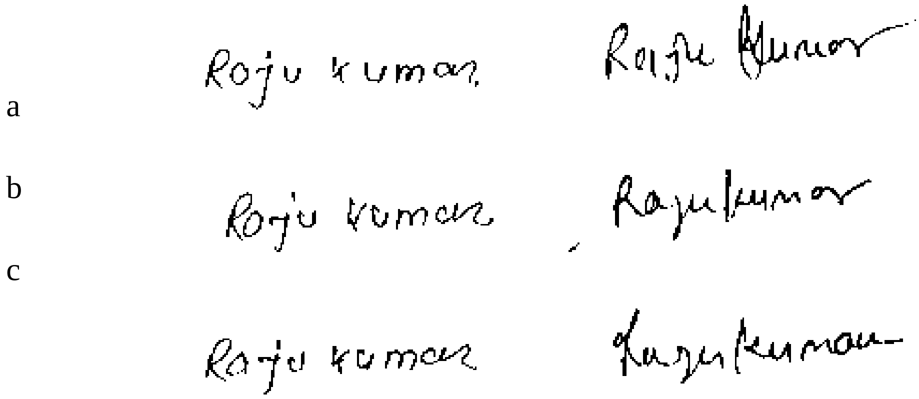


Fig 1.1: Types of forgeries

a) genuine and corresponding random forgery b) genuine and corresponding skilled forgery c) genuine and corresponding casual forgery

In this forgery, person requires good knowledge about the original signature and proper practice of signing. Naturally it is more difficult to detect skilled forgeries than other forgeries.

1.2 Motivation of Work

Signature of a person may change over time and it is not always unique and difficult to forge as iris patterns or fingerprints. One recent survey finds that “27% of cardholders (debit, credit and prepaid) around the world have experienced fraud in the past five years. A proper solution to reduce such losses could be Reliable signature verification.

We can reduce the security threats and risks by using the signature verification techniques in optimized way. Doing authorization and identification manually is a big problem when data collection is huge.so by providing a new automatic verification model will be a next step. On line signature verification is more reliable than Offline signature verification. Accuracy % obtained in Online signature verification is close to 99% while for offline signature verification is much low comparatively. In real time use of online signature verification is still low as compare to offline signature verification as we can see in banking systems. Hence development of a new verification model with a reasonable high accuracy would be useful, which is the main motivation of this research work.

1.3 Goal of the Thesis

Primary goal of this research work is to study the various techniques of Offline Signature Verification System, survey all these techniques and propose a new method for verification which can give optimized results. The goals of this thesis are:

- To survey the maximum existing methodologies which are being used by many researchers for verification.
- To propose a new technique using features point Extraction and Euclidean Distance.
- Apply proposed technique on self-created data set and validate it.

1.4 Research Objective

With the motivation explained in the previous section, the objective of our research work can be identified as:

- Identification of fraud signature along with the original signature mechanism when there are certain number signatures with number of parameter associated with it, provide a suitable mechanism to find original signature.
- In this research we consider two factor of signature, i.e. feature extraction and Euclidean distance.
- To improvise the signature forgery we provide set of requirement percentage of priority, person can choose number of features which want to extract.
- Let there be N training signatures (1, 2….n) that are waiting to be processed by a splitting and Euclidean distance. Each signature has its features according to vertical and horizontal splitting.

1.5 Thesis Organization

We start this dissertation with introduction in chapter 1. A detailed description of background is presented in chapter 2 which includes Biometric & its applications, literature review of Optimization Algorithm. Chapter 3 explains about proposed problem statement and its proposed solution. Chapter 3 also gives a brief about the optimization technique we have used. Chapter 3 also explains in detail about our proposed algorithm Modified-Feature Extraction. We evaluate the performance of the proposed algorithm and technique with Signature in chapter 4. We conclude about the work done and observations in chapter 5.

CHAPTER 2

LITERATURE REVIEW

2.1 Biometric Applications

There are two ways to implement Biometric technologies based on physiological and behavioural features. Finger prints, iris, and handwriting, facial recognitions are in the category of Physiological features. Voice and handwritten signatures are in the category of Behavioural features[1].

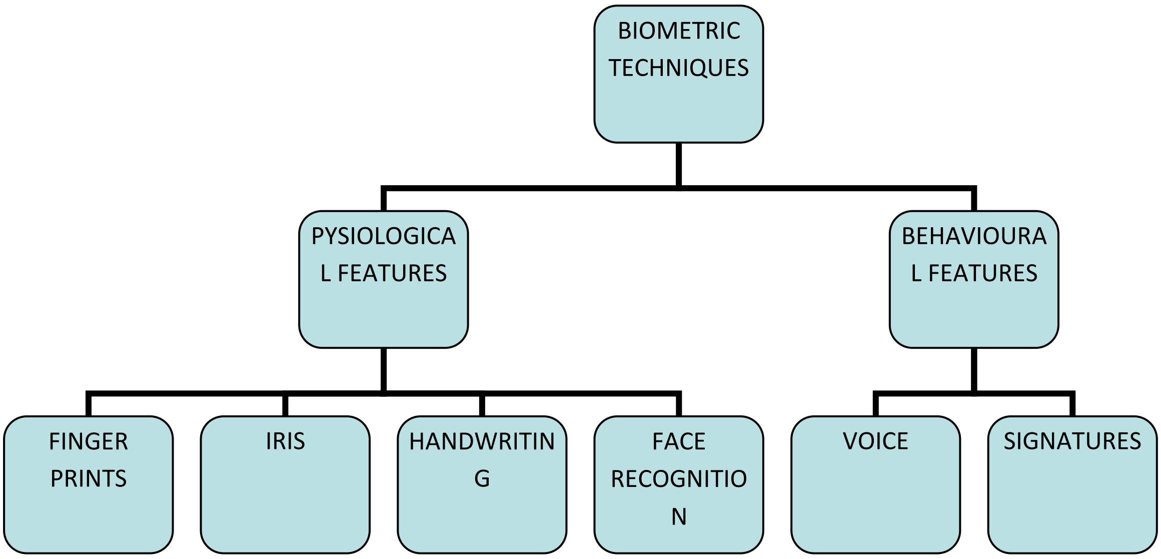


Fig 2.1: Hierarchy of biometrics techniques

Physiological feature are usually unalterable without causing trauma to individual. . On the other hand, behavioural biometric characteristics are traits that are learned or acquired, which later stabilize over a period of time .The selection of a particular biometric for use in a specific application involves a weighting of several factors. Jain et al. (1999) identified seven such factors to be used when assessing the suitability of any trait for use in biometric authentication. Universality means that every person using a system should possess the trait. Uniqueness means the trait should be sufficiently different for individuals in the relevant population such that they can be distinguished from one another. Permanence relates to the manner in which a trait varies over time. More specifically, a trait with 'good' permanence will be reasonably invariant over time with respect to the specific matching algorithm. Measurability (collectability) relates to the ease of acquisition or measurement of the trait. In addition, acquired data should be in a form that permits subsequent processing and extraction of the relevant feature sets. Performance relates to the accuracy, speed, and robustness of technology used. Acceptability relates to how well individuals in the relevant population accept the technology such that they are willing to have their biometric trait captured and assessed. Circumvention relates to the ease with which a trait might be imitated using an artefact[2].

Adaptive biometric Systems aim to auto-update the templates or model to the intra-class variation of the operational data. The two-fold advantages of these systems are solving the problem of limited training data and tracking the temporal variations of the input data through adaptation. Recently, adaptive biometrics has received a significant attention from the research community. This research direction is expected to gain momentum because of their key promulgated advantages. First, with an adaptive biometric system, one no longer needs to collect a large number of biometric samples during the enrolment process. Second, it is no longer necessary to re-enrol or retrain the system from scratch in order to cope with the changing environment. This convenience can significantly reduce the cost of maintaining a biometric system. Despite these advantages, there are several open issues involved with these systems. For miss-classification error (false acceptance) by the biometric system, cause adaptation using impostor sample. However, continuous research efforts are directed to resolve the open issues associated to the field of adaptive biometrics.

· From the past many years biometrics have been used widely in many domains and applications requiring authentication. It focuses on timely and attentive management. Besides more use of biometrics, it relates confusion and misconceptions and persist them. Confusion and false notation are easily removed when the authentication is done via biometrics.

· Parallelism is not shown in biometrics as able to quick and accurate real time, real data framework as they give a connectionless audit follow[1].

· There is capability of powerful research in biometrics and results are in our division when it on the correct way and attention .it works well and safe, correct and authenticate.

· There are variation in biometrics in terms of straight and non-straight time and processing advantages than other alternative methods as for security and validity.

· There are thousands of good agencies which are based upon smooth clock time and have good attendance system to automatically capture the employee presence and it will show significantly reduction on money and timely aspects.

There are two base modes of a biometric process as shown in figure. The first one is verification; in this we match the person’s instant biometric feature with the stored database biometric features to verify the correct person authentication or his physical signature. Reference data for all individuals are taken into account and keep into the purposed database, in first phase. Some of the samples are compared with the reference data to generate the correct and fraud scores and threshold value is calculated in second phase. Testing is done is third phase. Smart card, username or unique number (e.g. PIN) are used for testing to confirm to which one is used for comparative analysis. Natural use of the verification phase is positive identification; the goal is to stop the same identity to using many people.

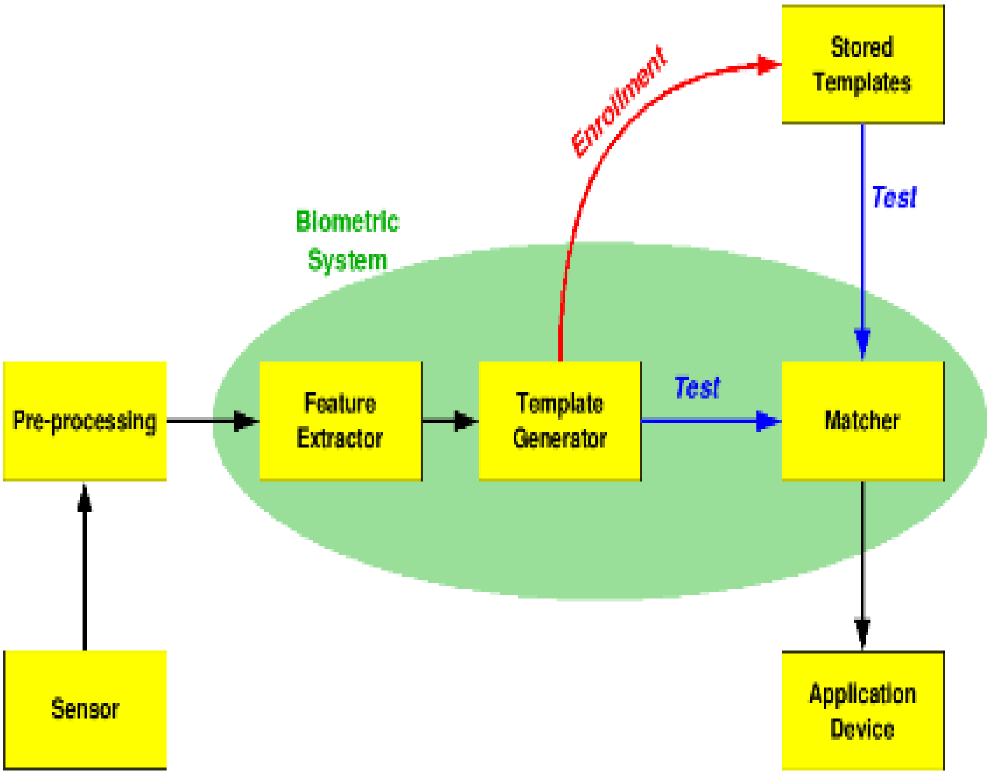


Fig 2.2: Biometrics Authentication System

Enrolment is the way for an individual which uses biometric system first time. Individual information is captured and stored during the enrolment[3]. In subsequent step; comparative information is obtained after matching the stored and captured information of the individual. Biometric authentication takes three processing into account named as feature extraction, template generator and matcher.

Note that it is crucial that storage and retrieval of such systems themselves be secure if the biometric system is to be robust. The first block (sensor) is the interface between the real world and the system; it has to acquire all the necessary data. Most of the times it is an image acquisition system, but it can change according to the characteristics desired. In pre-processing steps, it removes artefacts from the sensor, to enhance the input (e.g. removing background noise), to use some kind of normalization, etc. features extraction are done in third step. This step is an important step as the correct features need to be extracted in the optimal way. A vector of numbers or an image with particular properties is used to create a template. A template is a synthesis of the relevant characteristics extracted from the source. Elements of the biometric measurement that are not used in the comparison algorithm are discarded in the template to reduce the file size and to protect the identity of the enrolee. During the enrolment phase, the template is simply stored somewhere (on a card or within a database or both). During the matching phase, the obtained template is passed to a matcher that compares it with other existing templates, estimating the distance between them using any algorithm (e.g. Hamming distance). The matching program will analyze the template with the input. This will then be output for any specified use or purpose (e.g. entrance in a restricted area). Selection of biometrics for any practical application depends upon the characteristic measurements and user requirements[1]. We should consider Performance, Acceptability, Circumvention, Robustness, Population coverage, Size, Identity theft deterrence in selecting a particular biometric. Selection of biometric based on user requirement considers Sensor availability, Device availability, Computational time and reliability, Cost, Sensor area and power consumption

Signature Verification**:**

Signature verification is a common behavioural biometric to identify human beings for purposes of verifying their identity. Signatures are particularly useful for identification of a particular person because each person's signature is highly unique, especially if the dynamic properties of the signature are considered in addition to the static features of the signature. Even if skilled forgers can accurately reproduce the shape of signatures, but it is unlikely that they can simultaneously reproduce the dynamic properties as well.

2.2 Types of Signature verification

Signature verification is of two categories according to the available data available in input.

Offline (Static): here the image of a signature is input of offline signature verification system and is useful in automatic verification of signatures found on bank checks and documents. Some examples of offline signature shown in Figure 2.3[2]:

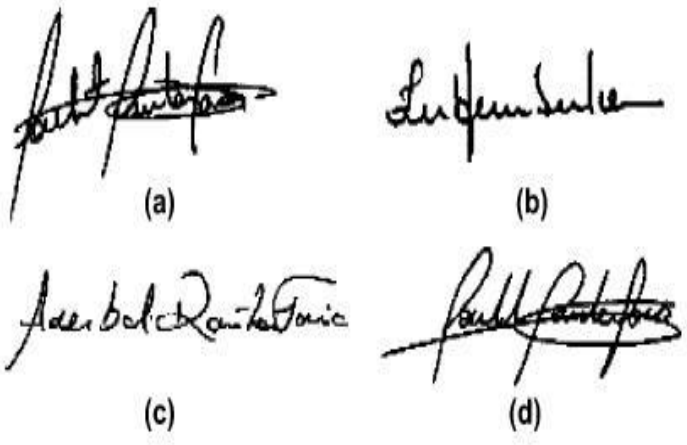


Fig 2.3: Offline Signatures

Online (Dynamic): Signatures that are captured by data acquisition devices like pressure-sensitive tablets (shown in Figure 1.3) and webcam that extract dynamic features of a signature in addition to its static shape features like curvature, length, width etc., and can be used in real time applications like credit card transactions, protection of small personal devices (e.g. PDA), authorization of computer users for accessing sensitive data or programs, and authentication of individuals for access to physical devices or buildings.

Fig 2. 4: Online Signatures

2.3 Steps involved in Offline Signature Verification

Offline signature Verification system requires various steps. There is a phase wise verification steps .Database creation, Pre-processing, Feature Extraction, Feature Comparison and Classification etc. are the phase of verification .Output of the first step or task is the input for next step. Block diagram shows the steps involved in Offline signature Verification:

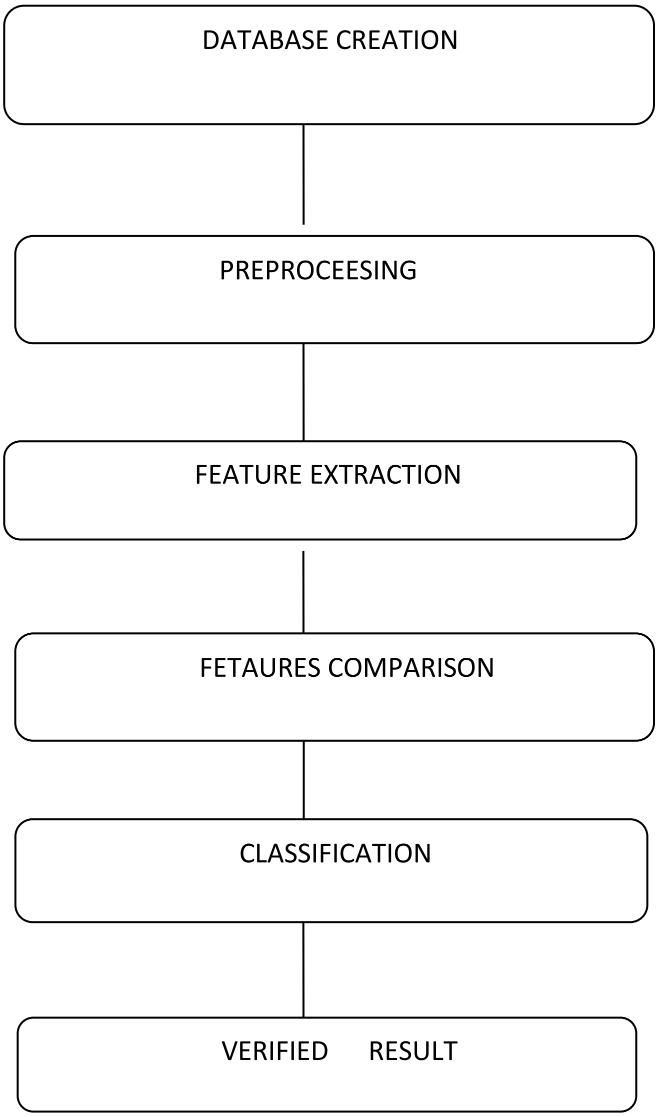


Figure 2.5: General System Overview

2.3.1 Database Creation

Signature samples from different users are collected in such a way that personal variance should be covered. If a person signs twice one can see small changes in between two signature samples, so this personal variability should be covered while collecting samples it is recommended that collect maximum no of samples from one user itself. If sufficient samples are taken from individual users it will be helpful at the time of comparison or matching in between the features and error rates can be minimized. Some research oriented organizations are providing datasets and one can also create datasets manually.

2.3.2 Pre-processing

Pre-processing is very important aspect in verification. Once you collect signature samples, samples need to be normalized because every person has its own style to do signature and they can highly differ in size i.e. length, width[4]. In pre-processing various steps are performed to remove noise. Some important steps are discussed below.

Scanning and Background Elimination:

Signature image is scanned through scanner and Data area cropping must be done for extracting features.

Noise Reduction:

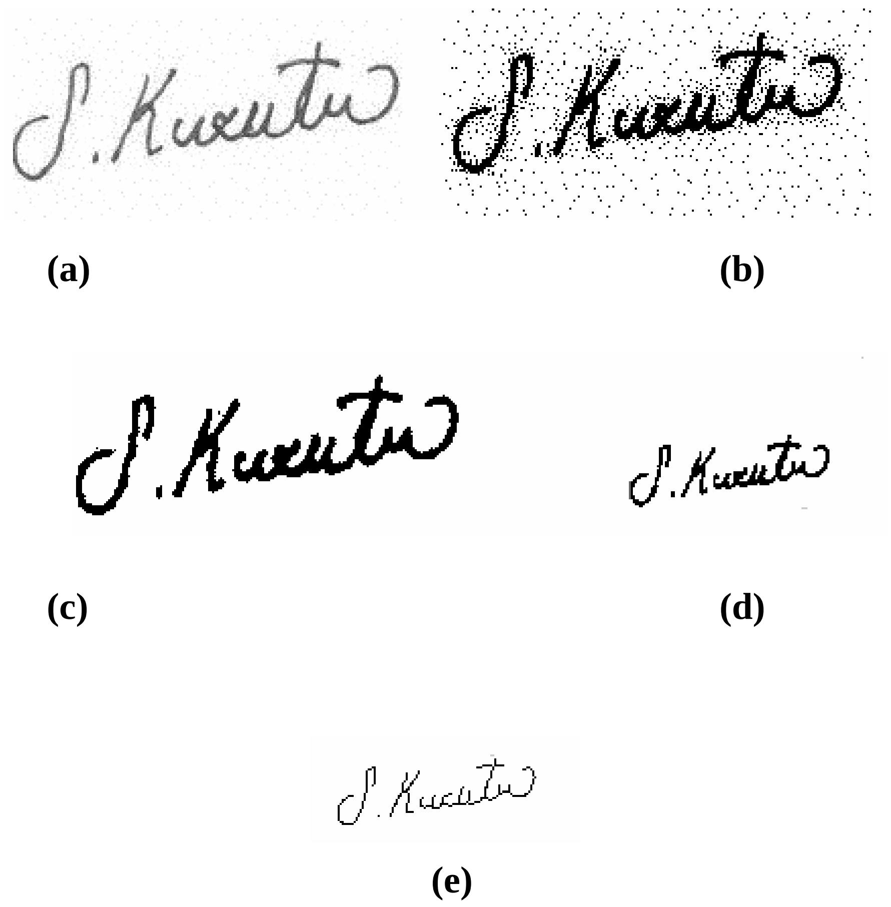
A noise reduction filter is applied to the binary image for eliminating single black pixels on white background. 8-neighbors of a chosen pixel are examined. If the number of black pixels is greater than number of white pixels, the chosen pixel will be black otherwise it will be white.

Width Normalization:

Signature dimensions may have intrapersonal and interpersonal differences. So the image width is adjusted to a default value and the height will change without any change on height-to-width ratio.

Thinning:

The goal of thinning is to eliminate the thickness differences of pen by making the image one pixel thick.



(a) (b) (c) (d) (e)

Fig 2.6: Pre-processing steps (a) scanned image (b) background elimination (c) noise removal (d) width normalization (e) thinning applied image

2.3.3 Feature Extraction

Various important features are extracted from the signature samples .Many researchers have computed different features like global features, local features, and wavelet features.

2.3.4 Classification

The major approaches to off-line signature verification systems are the Template Matching approach, Statistical approach, Structural or Syntactic approach, Spectrum Analysis approach and Neural Networks approach[17].

Template Matching Approach – The template matching is the simplest and earliest but rigid approach to pattern recognition. Because of its rigidness, in some domains, this approach has a number of disadvantages. It may fail if the patterns are distorted due to the imaging process, viewpoint change or large intra class variations among the patterns as in the case of signatures. It can detect casual forgeries from genuine signatures successfully. But it is not suitable for the verification between the genuine signature and skilled ones. The template matching method can be categorized into several forms such as graphics matching, stroke analysis and geometric feature extraction, depending on different features.

Statistical Approach – In the statistical approach, each pattern is represented in terms of d features and is viewed as a point in a d-dimensional space. Features should be chosen such a way that the pattern vectors belonging to different categories occupy compact and disjoint regions in a d-dimensional feature space. The effectiveness of the representation space (feature set) is determined by how well patterns from different classes can be separated. Hidden Markov Model (HMM), Bayesian these are some statistical approach commonly used in pattern recognition. They can detect causal forgeries as well as skilled and traced forgeries from the genuine ones[5].

Structural Approach - Structural approaches mainly related to string, graph, and tree matching techniques and are generally used in combination with other techniques [4]. When the signature image is considered as a whole entity, the structural approach is used for the signature verification. It shows good performance detecting genuine signatures and forgeries. But this approach may demand a large training set and very large computational efforts.

Spectrum Analysis Approach – To decompose a curvature-based signature into a multi-resolution format, spectrum analysis approach is introduced. This method can be applied to different languages, including English and Chinese. Moreover this approach may be useful especially for long signatures like some of the Indian scripted signature.

2.4 Types of features

2.4.1 Global Features

Signature Area is the number of pixels which belong to the signature. This feature provides information about the signature density.

Signature Height to Width Ratio is obtained by dividing signature height to signature width. Signature height and width can change. Height-to-width ratios of one person's signatures are approximately equal.

Maximum horizontal histogram and maximum vertical histogram The horizontal histograms are calculated for each row and the row which has the highest value is taken as maximum horizontal histogram. The vertical histograms are calculated for each column and the column which has the highest value is taken as maximum vertical histogram[3].

Local maxima numbers of the signature: The number of local maxima of the vertical and horizontal histogram is calculated.

Edge point numbers of the signature: Edge point is the pixel which has only one neighbour, which belongs to the signature.

2.4.2 Moments Features

Moments are used for the purpose of image analysis. Images are rotated in between 0^0 to 360^0 .Image feature are computed by normalizing central moments through order three, that are invariant to object scale, position, and orientation.

2.4.3 Grid Features

Image is divided in grids and pixel values either in horizontal or vertically or in both directions is computed and feature vectors are obtained. Grid segmentation is a technique that is used for signature detail analysis. A grid of 12 x 8 segments are depends on the pre-processed image and the features mentioned below are calculated for each of the segment.

i) Pixels Density. This pixel density gives the number of black pixels of each segment.

ii) Pixels Distribution. It gives the pixel geometric distribution in a cell (intersection of row and column).

iii) Predominant Axial Slant. It is a value representing the predominant inclination of each cell. For each cell of the grid the number of three pixels that are connected to each other is calculated against the given templates.

2.4.4 TriSurface Features

The surface area of two visually different signatures could be the same. For the purpose of increasing the accuracy of a feature describing the surface area of a signature, the ‘triSurface’ feature was investigated as an extension in which the signature was vertically separated into three equal parts. The surface area feature is the surface covered by the signature, including the holes contained in it. The total number of pixels in the surface was tallied, and the proportion of the signature’s surface over the total surface of the image was calculated[3]. This process was used for the three equal parts of the signature, giving three values between 0 and 1.

2.4.5 Six fold Surface features:

This feature is different from the Tri Surface feature mainly in two ways. Firstly, the number of feature values obtained is doubled to six with the Six fold Surface. Secondly, centres of gravity are determined to assist in the calculation of the Six fold Surface features.

The Signature image is first divided into three parts vertically. The centre of gravity is calculated for each of the three parts, and the signature surface above and below the horizontal line of the centre of gravity (giving two subsections for each part) was calculated. The result was a set of six feature values corresponding to the surface of the six sub-sections as illustrated in Figure 7.

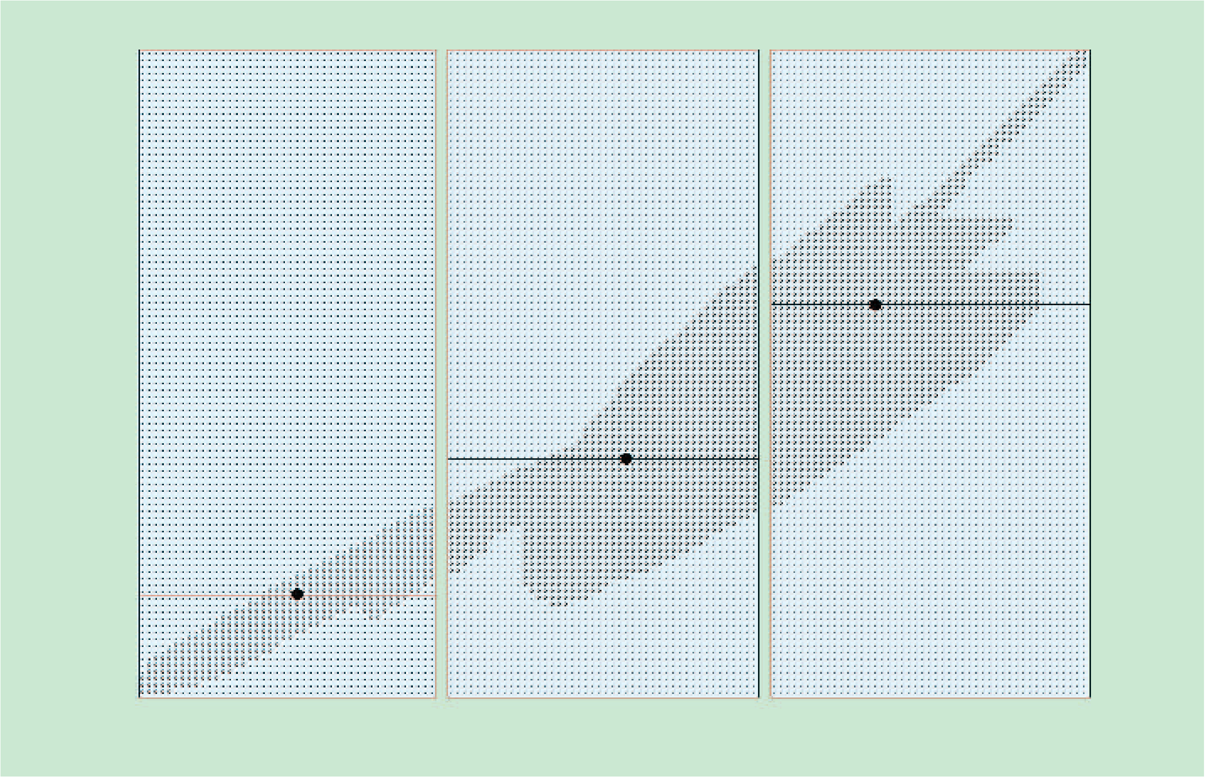


Fig 2.7: Six Fold Surface features. G1, G2 and G3 are the centres of gravity for the respective sections (a), (b) and (c).

2.4.6 The Best-Fit Features

The line of best fit usually attempts to represent a scatter of points in an area. In order to obtain an approximation for the signature's skew; the line of best fit was calculated using minima and maxima from the bottom edge of the signature. Similarly the line of best fit from the top of the signature was also calculated. The angles between each of these lines and the X-axis were calculated, giving two features. The surface area enclosed between the two lines became the third feature. Figure 7 depicts the concept.

2.5 Choice of Features

The choice of using global or local features depends mainly on style of the signature as well as the types of forgeries to be detected by the system. A suitable combination of global and local features has been found to improve a classifier's ability to recognize forgeries and to tolerate intrapersonal variances (cited by [13]).

The global features are extracted at a low computational cost, and they have good noise resilience. These features are less sensitive to noise and signature variations. So it does not give a high accuracy for skilled forgeries, but it is suitable for random forgeries and is better to be combined with other types of features [16].

On the other hand, even though the local features are dependent on the zoning process, still they are more suitable to identify skilled forgeries [15]. Local features describe only a small portion of signature and extract more detailed information from the image. Local features are more sensitive to noise and they are not affected by other regions of the signature. Although they are computationally expensive, they are much more accurate than global features [17].

The global features can deliver limited information for signature verification [14]. Small distortions in isolated regions of the signature do not cause a major impact on the global feature vector. They are, however, dependent upon the overall position alignment and therefore highly susceptible to distortion and style variations [15].

On the other hand, local features provide rich descriptions of writing shapes and are powerful for discriminating writers, but the extraction of reliable local features is still a hard problem [14].

The local features based approaches are more popular in online verification than in the offline. Because as compared to 2D images, it is much easier to calculate local shape features and to find their corresponding relations in 1D sequences [14]. In manual verification, global features are observed and it is seen that the intra personal variations with respect to the global aspect is very low.

2.6 Some approaches to Offline Signature verification

To improve the efficiency of the signature verification systems, researchers have tried different methods with various approaches. Some of them have employed two or three expert systems that evaluate the signature in two/three different ways and verify whether it is genuine or forgery.

- J. B. Fasquel and M. Bruynooghe give one offline signature verification method based on some statistical classifiers. These consisted of three steps – in first one we transform the original signature in four Gabor transforms, the second one is to inter correlate with the database and analysed original signature and then in the third step we analyse the original signature with fusing techniques. This technique gives the result of 62.4% for the fraud signature.

- Emre Özgündüz et al. [8] provide a method which was based on signature feature named as global, grid. In Signature area we used global feature, Aspect Ratio of the signature, Maximum horizontal histogram and maximum vertical histogram, Horizontal and vertical centre of the signature. Local maxima numbers of the signature and Edge point numbers of the signature. For classification purpose we use SVM.

- We can verify a signature based on two neural networks classifier and three features (global, texture and grid) was proposed by Mohammed A. Abdala & Noor Ayad Yousif. The first one is three Back Propagation NNs and the second one is Radial Basis Function NNs.

- V A Bharadi and H B Kekre [3] had designed a multi algorithmic signature recognition system considering the conventional features like Number of pixels, Picture Width, Picture Height, Horizontal max Projections, Vertical max Projections, Dominant Angle-normalized, Baseline Shift etc. For extracting information in pixel distribution of the Signature, they proposed Walsh Coefficients, Vector Histogram, Grid and Texture Feature as global as well as cluster based Features.

- H.N. Prakash and D. S. Guru [8] give a method which is based on an approach for offline signature verification rotation features of centroids. The proposed method works upon bi-interval valued feature vector. Distance and orientation features of centroids of offline signatures were used to form bi-interval valued symbolic feature vector for representing signatures. This method is more optimized then other method as result seen so far FRR of 27.77% and FAR of 26.11%, with 63 centroids and threshold = 977, FRR and FAR were 20.22% and 29.51% respectively.

- Madasu Hanmandlu et al. [5] give a method for offline signature verification and forgery detection approach based on fuzzy techniques. This model creates the Signature features as exponential function and a suitable box is designed for feature extraction. This method fixed some parameter as feature and gives the correct result as other method.

- Hai Rong Lv et al. [8] gives HMM approach to offline signature verification. Here the pixels of image are denoted as landmark point set consisting of turning, isolated, trifurcate, intersection and termination points on signature skeleton. Based on landmark point mixer, they made matching relations between planar regions to achieve deformable grids, and then extract grid features from them. To represent the grids of a signature image, they used features like pixels Density (numbers of pixels inside the cell), geometric centre (gravity centre distance in each cell), stroke curvature (curvature angle of the bigger stroke inside the cell), slant (predominant slant inside the cell) and grid area.

- J. F. Vargas et al. [11] gives a method for offline signature verification system which is relay on transformation to grey level picture using features extraction. They Treated picture as matrix form and convert into binary. Original and fraud signature are classified using SVM model. They were giving the result as EER of 12.82% for skilled forgery.

- Stephane Armand et al. [12] give a method for off-line signature verification and identification using combine feature of back propagation method and radial basis function. In this method ,they treated the image as binary .After that they classify the feature of image as physically and behaviourally and select that feature which are useful for signature processing ,After that a comparative result are given as comparative to other method.

- A method for signature verification using local Radon Transform was proposed by Vahid Kiani et al. [15]. In this method the writers used the approach for segmentation of an image as line segments .The classifier was SVM. Many of the advantages like noise reduction , size invariance and shift invariance are corrected in this method

- M. Taylan DAS and L. Canan DULGER [10] presented a technique for off-line signature verification based on machine learning technique which is involved in neural network. They used the particle swarm optimization technique for three types of forgery .40% of the signatures were detected correctly for skilled forgery.

2.7 Verification Techniques

2.7.1 Bayesian Learning

Bayesian reasoning estimates the posterior probability of a hypothesis given some initial knowledge or previously available data. Prior knowledge is combined in Bayesian learning along with the observed data to obtain posterior probability of the hypothesis. Bayesian method computes the posterior probability of the hypothesis according to Bayes' rule:

$$P(h \mid D) = \frac{P(D \mid h)P(h)}{P(D)}$$

It is a probabilistic approach, given prior probabilities of data and hypothesis, the most likely posterior hypothesis can be determined using this technique. This approach overcomes the limitation of having limited number of genuine samples [4]. Other techniques may require forgery samples as well, but this method overcomes this limitation as well. The most significant application of this method is that it just does not simply accept or reject a sample but it gives a probability as output of how likely the signature sample belongs to an individual, as a result a confidence value can be attached to all the probable choices. Bayesian method gives a probabilistic output for example this signature is 83% genuine or 90% forged. New instances can also be classified by combining the predictions of multiple hypotheses. Regarding signature verification, Bayesian learning can be implemented as follows: the hypothesis space can be defined as H = {genuine, forged}, and the data D can be the features of the signature samples such as velocity, pressure, no. of strokes etc. On the basis of the prior knowledge of these hypotheses and data, the posterior hypothesis can be estimated using Bayes' theorem.

2.7.2 Hidden Markov Model (HMM)

HMM is a strong and effective statistical tool for modelling generative sequences, characterized by an underlying process that generates an observable sequence. HMMs have been applied in many application areas such as signal processing, speech recognition, pattern recognition and can be effectively implied in signature verification as well. HMM is a generalization of Markov Model. It is a robust method to model the variability of discrete time random signals where time or context information is available [5]. It can manage time duration varying signals such as signatures speech etc. For this reason it is popular for speech and signature recognition applications [6]. The signing process is divided into several states that constitute the markov chain. Each of the signature segments corresponds to each state in the model. Sequences of probability distributions of the different features that are used in the verification task are taken and a matching is done on it [7]. The verification score in these systems is usually obtained as the signature log-likelihood. An important part in generative model-based signature verification systems is the verification score normalization [8]. The verification score is a score that determines whether a particular signature is genuine or forged using a threshold value. These threshold values can be writer dependent or feature dependent. The disadvantages of using HMM in signature verification is that it requires huge number of features to be set, and the number of data to train the model is very large as a result of which its time complexity is very high.

2.7.3 Neural Network

Neural network is a mathematical model that can learn from examples and based on this knowledge can solve many problems such as pattern recognition. A number of genuine and forged samples are stored in the database which is used for learning and thus judging whether a given test signature is genuine or forged. An artificial Neural Network is trained to recognize the variation that exists in the target signature with respect to the sample signature. Handwritten signature samples are considered input for the artificial neural network model and typically weights are learned during training a NN. The major factors of using ANN are Expressiveness, ability to generalize, sensitivity to noise, and graceful degradation. The major drawback of using ANN model is that it takes a lot of time for training.

In modelling of a signature verification system Neural Network can be used as follows: As training data, a vector of n number of sensors can be used where n is the number of features of the signature considered for verification. Here each of these vectors would estimate the similarity of the target feature with respect to the features of genuine signature samples. The ANN used for this purpose is a multilayer feed forward network which consists of n number of input units, one output unit signalling genuine or not genuine, and some units in one or more hidden layer(s). Backpropagation algorithm is used for training.

2.7.4 Support Vector Machine (SVM)

In supervised learning models, support vector machine are placed whose achieves stem from statistical learning theory. A support vector machine converts the low dimensional data set into higher dimensional data set which is more readable and used for classification of non probalistic input data sets which are linearly separable as consider as binary separator. SVM has been considered a good choice for solving the signature verification problem as it is frequently used for pattern recognition applications, classification and regression problems [9]. An SVM maximally creates a hyper plane which is used for classification of higher dimensional data [10]. An SVM takes a set of input data and determines to which of the two classes the input data belongs.

Technique	**A**pproach	Basis	Type
Support vector Machine	Predictive Modelling	Principle of structural minimization	Statistical, unsupervised Learning
Neural Network	Machine Learning	Adaptive system changing its Structure during a learning phase	Supervised Learning

Bayesian Learning	Probabilistic	Use of priori information to obtain posterior information	Statistical
Hidden Markov Model	Probabilistic	The hidden variables control the mixture component to be selected for each observation	Statistical

Table 2.1: Comparative Study of Various methods of Offline Signature Verification

The researches done in the field of offline signature verification are described below:

Author	Publication	Year	Extracted Features	Verification Method
A. Rathi, D. Rathi,P. Astya	using Pixel based Method	Sept2012	Pixels	Fuzzy Neural Network
Julian FiEERez	Using Contour Features	2008	Length-based and direction -based	Euclidean Distance Method
Julian FiEERez	Fusion Of Static Image	2009	Contour Features	Euclidean Distance Method
Eric Granger and Robert	A Multi-Hypothesis Approach	2009	States of signature	Hidden Markov Model

Sargur N. Srihari	Learning Strategies and Classification	2004	Combination of features	Distance Statistics
DakshinacRanjan Kisku, Phalguni Gupta	Fusion of Multiple Classifiers	2010	Global and Local Features	Support Vector Machine
S. Daramola	Offline Signature Recognition	2010	DCT Features	Hidden Markov Model
Sargur Srihari	Using Distance Statistics	2004	Gradient, 2004 Structural Concavity	Bayes Classifier

Table 2.2: Researches done in the field of offline signature verification

2.8 Characteristics of forgeries

In offline signature verification, some general characteristics of genuine signatures and forgeries need to be understood. Knowledge of these characteristics is important for determining those aspects or features of the signatures that are most important for automatic signature verification. In [7] Vamsi Krishna Madasu and Brian C. Lovell have mentioned few such characteristics outlined by several document examiners in the past in various literatures:

1. Enlargement of characters: Original signature is smaller than fraud signature. A fraud person takes much time to write a signature in comparison to right person. Size of alphabets and the area of whole signature are larger than the original in comparison to fraud.

2. Tendency of curves to become angles: In the forgery Observation is based on curved letters because they are more angular. In slower speed the fraud person try to capture the original shape of the right signature to produce the curve accurately. To making of the curves

very much time are elapses to create the angular shapes. In the same way, angled letters in the original signature can become smooth curves.

3. Retouching: After the imitation the forger try to make correction at later stage. Due to this reason signature is destroyed as somewhere lines are thicker and somewhere thinner at points, and they don't follow the sequential flow of the pen as in the original signature.

4. Poor line quality: There is difference of the pressure in case of forged and original signature. Fraud signature requires harder pressure on paper as comparative to original which requires smooth. The ink reveals variation in light and shade, pressure and speed, with either more or less ink appearing on the page. Sometimes, Lighter pressure can also be detected by the fraud signature this may be done due to the physical perspective of the fraud person. [23].

5. Hesitation: In the process of creating a forgery, the forger may pause to consult the genuine signature and then continue duplicating it. This can often create blobs.

6. Punctuation: In forgery full stops, dots on small letter „i" are found to be in the wrong place, missing or added.

7. Differing pressure: It is very hard to verify the pen pressure of the both of the signature .Forger cannot copy the original pen pressure profile as like as the correct person. The pen pressure may be differing as too heavy or too light, depending on the writing style of the forger. Pressure differences may be occurring at different location in the picture.

8. Sudden endings: It is character which involved as the feature of a forgery. The forged signature just stops as soon as the signature ends because of the attentive state of mind of forger, it does not have the free flowing trail which presents when signed naturally. It is very difficult to trail off in the same way as the genuine.

9. Forger's characteristics: Everyone has his/her own characteristics of handwriting. The forger unconsciously exposes his/her own handwriting characteristics when doing the forgery. It is observed that forger cannot avoid revealing some of his/her writing characteristics like the basic letter shapes, spacing and position of letters in relation to base line even in a forgery.

10. Baseline error: The line crossing boundary of the base signature that runs across the base of the signature is not same as the forged signature and the correct signature. The baseline in a signature is not horizontal and any notable variances in the baseline indicate forgery.

11. Spacing: Difficulty often occurs in spacing like as individual letters complete words, punctuation marks to remove in original signature copy. These spacing cannot be copied by tracing a signature, as they are smaller or larger.

12. Bad line quality: Bad line quality may also occur due to the hesitation of a man.

13. Forming characters not appearing in signatures: Sometimes a forger makes such type of signature which is not as containing alphabets as in original signature because forger knows the name of the correct person. Sometimes a forger makes such type of signature which are containing incorrect alphabets which are not containing in original signature because forger unknowns the name of the correct person. This is the very difficult task in computerised system.

CHAPTER 3

PROPOSED WORK

3 .1 Proposed Method

The main objective of this project is to design a robust offline signature verification system. New proposed method is based on feature point extraction and Euclidean distance method used for training and testing. This methodology consists of some crucial steps which need to follow in sequential order. All steps are discussed below:

3.1.1 Dataset Creation:

Datasets are created manually by collecting samples of signature from each person. Every individual person has signed on a paper. No of samples from a single person has taken in a way such that personal variance should be covered so 10 signature samples are take from single person.

3.1.2 Pre-processing

To normalize the scanned signature images, some pre-processing steps have to be applied. The purpose in this phase is to make signatures to be of standard size and ready for feature extraction .Pre-processing steps are very necessary to remove noise from the signature image and strictly need to be follow. Some crucial steps are listed below:

(a) Gray Scale Conversion: Since the scanned images are stored in database as a colour image, a three dimensional image (MXNX3) is not suitable for further processing, and should be converted into a grayscale image where each pixel is represented by a value in the range 0 to 255.’rgb2gray’ command is used for this purpose.

(b) Binary Conversion: It allows reducing image information by removing background so that the image is black & white type. This type of image is much easier for further processing.

(c) Resize: Every person signs in different way so all images need to kept in same dimension. So every image is resized into [10, 10] standard size.’imresize’ command is used for that purpose.

(d) Moving signature into the center of image: After resizing we calculate variance (signature is considered to be binary and consists of only black and white pixels).If a square block has a zero variance we remove that square, otherwise restore.as shown in figure below

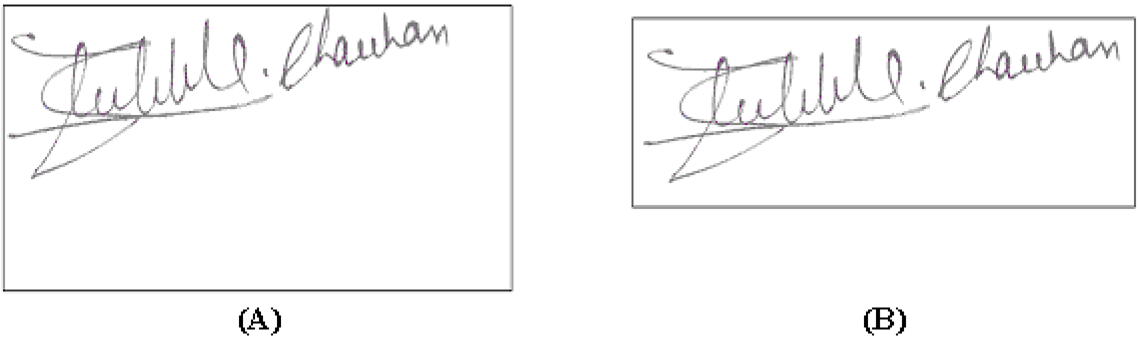


Fig 3.1: Captured signature (A) before adjustment and (B) after adjustment

3.1.3 Feature Extraction

The geometric features are based on two sets of points in 2-dimensional plane [7]. Considering a scenario where sixty feature points are required for classification of signature image, vertical splitting of the image results in thirty feature points (v1, v2, v3... v30) and the horizontal splitting results in thirty feature points(h1,h2,h3,…….,h30). These feature points are obtained with relative to a central geometric point of the image. Here the centered image is scanned from left to right and calculate the total number of black pixels. Then again, starting from top to bottom and calculate the total number of black pixels. Then divide the image into two halves w.r.t. the number of black pixels by two lines vertically and horizontally which intersects at a point called the geometric center. With reference to this point we extracted 60 feature points: 30 vertical and 30 horizontal feature points of each signature image.

Feature points based on vertical splitting:

In this method thirty feature points are obtained from vertical splitting w.r.t to geometric center. The steps for finding vertical feature points as below

Input: Static Signature Image after pre-processing

Output: Vertical Feature points v1, v2, v3……v30.

The steps are:

1. Divide the Image into two halves left part and right part by passing the vertical line passing through center of the image or Geometric center.

2. Again find the geometric centers of the image as v1 and v2 for left and right parts corresponding.

3. Divide the left and right part with horizontal lines through v1 and v2 to divide the two parts into four parts: Top-left, Bottom-left and Top-right, Bottom right parts from which we obtain v3, v4 and v5, v6.

4. We again split each part of the image through their geometric centers to obtain feature points v7, v8, v9… v13, v14.

5. We do one more split on each of the parts to obtain all thirty vertical features.

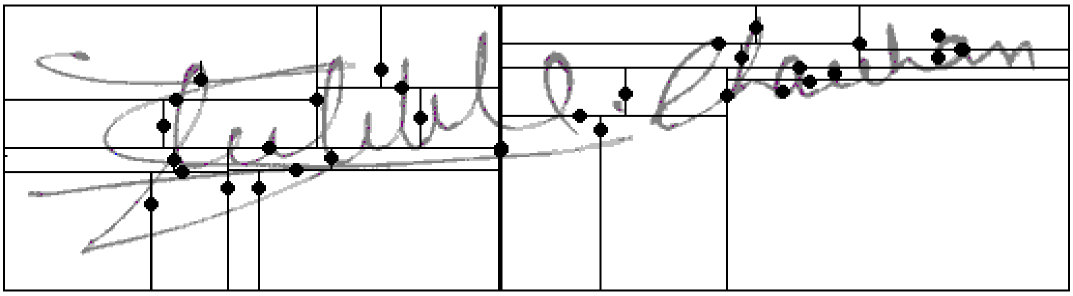


Fig 3.2: Vertical splitting of the signature image

Feature points based on horizontal splitting:

In this method thirty feature points are achieved through horizontal splitting w.r.t. the central feature point. The steps for finding horizontal feature points are given below:

Input: Static signature image after pre-processing
Output: Horizontal feature points: h1, h2, h3, h4… h29,
h30.
The steps are:
1) Split the image with a horizontal line passing through the geometric center (h0) which divides the
Image into two halves: Top part and Bottom part.

2) Find geometric centers h1 and h2 for top and bottom parts correspondingly.

3) Split the top and bottom part with vertical lines through h1 and h2 to divide the two parts into four parts: Left-top, Right-top and Left-bottom, Right bottom parts from which we obtain h3, h4 and h5, h6.

4) We again split each part of the image through their geometric centers to obtain feature points h7, h8, h9... h13, h14.

5) Then we split each of the parts once again to obtain all the thirty vertical feature points (as shown in Figure).

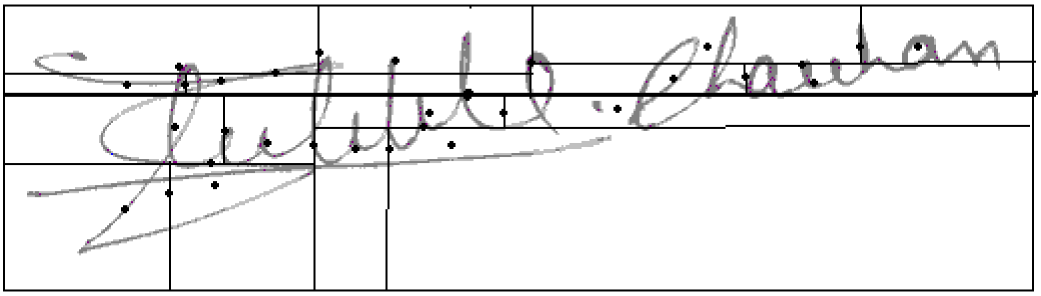


Fig 3.3: Horizontal splitting of the signature image

3.1.4 Classification

In this method features are based on geometric properties. So we use Euclidean distance model for classification .This is the simple distance between a pair of vectors of size n. Here vectors are nothing but feature points, so the size of vector is 2. How to calculate distance using Euclidean distance model is described in the following Section. In threshold calculation these distances are useful.

Euclidean distance model:

Let *A* (*a*1, *a2*…..*an*) and *B* (*b*1, *b2*….*bn*) are two vectors of size *n*. We calculate *distance* (*d*) by using equation as shown in below.

$$distance(d) = \sqrt{\sum_{i=1}^{n}(a_i - b_i)^2}$$

Vectors are Feature points on plane. D is the simple distance between two points.

Threshold:

After obtaining the vertical and horizontal splitting feature points, we have calculated the threshold for these feature points. These are the steps for calculating threshold:

1. We have taken n number of training signature. And x1, x2, x3………xn be the feature points of corresponding signatures. After this we calculate the median of all feature points as xmedian.

2. After this we calculate individual distance points of all the feature points from the median named as d1, d2 ….dn.

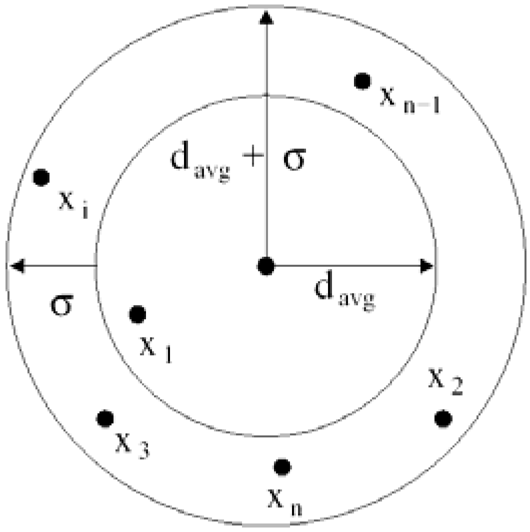


Fig 3.4: d_{avg} **(average distance) and** ***s*** **(standard deviation) derivation from distances**

Let d1, d2….dn are the distances
D1=dis (xmedian, x1)
D2=dis (xmedian, x2)
.
.
.
.
.

Dn=dis (xmedian, xn)

To calculating the threshold we need two parameters named as average of all distance points and standard deviation of all the distance points.

Davg = average (*d*1, *d2… dn*)
σ = *SD* (*d*1, *d2… dn*)

After obtaining the both of the parameter we calculate the threshold value of all feature points (vertical and horizontal feature points).

$$threshold(t) = \sqrt{\sum_{t=1}^{30} (d_{avg,t} + \sigma_t)^2}$$

CHAPTER 4

EXPERIMENTAL RESULTS

The proposed scheme is implemented on MatLab platform for binary images and tested using a standard database consisting of signatures of ten Dutch persons and 10 Chinese persons each providing 20 samples for training and 4 samples for testing. Eight different signatures are reproduced for each person where four amongst them constitute unskilled forgery and the rest of four constitute skilled forgery.

The results produced in terms of False Rejection Rate (FRR) and False Acceptance Rate (FAR) is studied for variations in training size and no. of features extracted from signatures.

Another term used in this section is the conversion level which denotes the threshold value of converting a gray scale image to binary image

4.1 Variation of FAR and FRR w.r.t. conversion level:

The conversion level is an important discriminator in determining the signature’s quality which will be further provided to the feature extractor to extract features. The variation studied in the proposed scheme id for conversion level selecting from these values [0.9, 0.92, 0.95 and 0.96].

4.1.1 Variation of FAR and FRR w.r.t. training size for conversion level of 0.95:

1. Training size = 5 and Conversion level = 0.95:

The below table displays the variation in features (extracted using Euclidean distance) w.r.t. FAR and FRR over a small training set of five samples.

The training dataset includes 5 signatures from 10 persons each and 12 testing signatures where 8 of them are forged and 4 of them are genuine.

No. Of features Extracted/Used	No. of Original Signatures Identified as Fraud	No. of Fraud signatures identified as Original	No. of Correctly classified signatures	FAR (%age)	FRR (%age)
10	25	9	86	7.50	20.83
20	26	8	86	6.67	21.67
30	25	9	86	7.50	20.83
40	24	8	88	6.67	20.00
50	24	10	86	8.33	20.00
60	23	9	88	7.50	19.16
70	23	8	89	6.67	19.16
80	22	8	90	6.67	18.33

Table 4.1: Variation of FAR and FRR when training size=5 and conversion level=0.95

2. Training size = 10 and Conversion level = 0.95:

The below table displays the variation in features (extracted using Euclidean distance) w.r.t. FAR and FRR over a medium sized training set of ten samples.
The training dataset includes 10 signatures from 10 persons each and 12 testing signatures where 8 of them are forged and 4 of them are genuine.

No. Of features Extracted/Used	No. of Original Signatures Identified as Fraud	No. of Fraud signatures identified as Original	No. of Correctly classified signatures	FAR (%age)	FRR (%age)
10	15	14	91	11.67	12.50
20	17	13	90	10.83	14.16
30	16	13	91	10.83	13.33
40	16	12	92	10.00	13.33
50	15	11	94	9.17	12.50
60	16	12	92	10.00	13.33
70	17	13	90	10.83	14.16
80	16	14	90	11.67	13.33

Table 4.2: Variation of FAR and FRR when training size=10 and conversion level=0.95

3. Training size = 15 and Conversion level = 0.95:

The below table displays the variation in features (extracted using Euclidean distance) w.r.t. FAR and FRR over a medium training set of fifteen samples.

The training dataset includes 15 signatures from 10 persons each and 12 testing signatures where 8 of them are forged and 4 of them are genuine.

No. Of features Extracted/Used	No. of Original Signatures Identified as Fraud	No. of Fraud signatures identified as Original	No. of Correctly classified signatures	FAR (%age)	FRR (%age)
10	14	17	89	14.16	11.67
20	14	14	92	11.67	11.67
30	13	16	91	13.33	10.83
40	12	14	94	11.67	10.00
50	12	17	91	14.16	10.00
60	12	16	92	13.33	10.00
70	11	15	94	12.50	9.17
80	10	16	94	13.33	8.33

Table 4.3: Variation of FAR and FRR when training size=15 and conversion level=0.90

4. Training size = 20 and Conversion level = 0.95:

The below table displays the variation in features (extracted using Euclidean distance) w.r.t. FAR and FRR over a large training set of twenty samples.
The training dataset includes 20 signatures from 10 persons each and 12 testing signatures.

No. Of features Extracted/Used	No. of Original Signatures Identified as Fraud	No. of Fraud signatures identified as Original	No. of Correctly classified signatures	FAR (%age)	FRR (%age)
10	13	15	92	12.50	10.83
20	10	14	96	11.67	8.33
30	13	14	93	11.67	10.83
40	11	15	94	12.50	9.17
50	10	15	95	12.50	8.33
60	9	16	95	13.33	7.50
70	8	14	98	11.67	6.67
80	7	15	98	12.50	5.83

Table 4.4: Variation of FAR and FRR when training size=20 and conversion level=0.95

4.1.2 Variation of FAR and FRR w.r.t. training size for conversion level of 0.90:

1. Training size = 5 and Conversion level = 0.90:

The below table displays the variation in features (extracted using Euclidean distance) w.r.t. FAR and FRR over a small training set of five samples.
The training dataset includes 5 signatures from 10 persons each and 12 testing signatures where 8 of them are forged and 4 of them are genuine.

No. Of features Extracted/Used	No. of Original Signatures Identified as Fraud	No. of Fraud signatures identified as Original	No. of Correctly classified signatures	FAR (%age)	FRR (%age)
10	25	11	84	9.17	20.83
20	26	9	85	7.50	21.67
30	27	8	85	6.67	22.50
40	23	8	89	6.67	19.16
50	24	9	87	7.50	20.00
60	23	8	89	6.67	19.16
70	24	7	89	5.83	20.00
80	22	6	92	5.00	18.33

Table 4.5: Variation of FAR and FRR when training size=5 and conversion level=0.90

2. Training size = 10 and Conversion level = 0.90:

The below table displays the variation in features (extracted using Euclidean distance) w.r.t. FAR and FRR over a medium sized training set of ten samples. The training dataset includes 10 signatures from 10 persons each and 12 testing signatures.

No. Of features Extracted/Used	No. of Original Signatures Identified as Fraud	No. of Fraud signatures identified as Original	No. of Correctly classified signatures	FAR (%age)	FRR (%age)
10	15	13	92	10.83	12.50
20	15	12	93	10.00	12.50
30	16	12	92	10.00	13.33
40	17	12	91	10.00	14.16
50	16	11	93	9.17	13.33
60	12	12	96	10.00	10.00
70	13	12	95	10.00	10.83
80	14	13	93	10.83	11.67

Table 4.6: Variation of FAR and FRR when training size=10 and conversion level=0.90

3. Training size = 15 and Conversion level = 0.90:

The below table displays the variation in features (extracted using Euclidean distance) w.r.t. FAR and FRR over a medium sized training set of fifteen samples. The training dataset includes 15 signatures from 10 persons each and 12 testing signatures.

No. Of features Extracted/Used	No. of Original Signatures Identified as Fraud	No. of Fraud signatures identified as Original	No. of Correctly classified signatures	FAR (%age)	FRR (%age)
10	14	17	89	14.16	11.67
20	13	14	93	11.67	10.83
30	13	15	92	12.50	10.83
40	13	15	92	12.50	10.83
50	14	16	90	13.33	11.67
60	10	16	94	13.33	8.33
70	10	14	96	11.67	8.33
80	9	16	95	13.33	7.50

Table 4.7: Variation of FAR and FRR when training size=15 and conversion level=0.90

4. Training size = 20 and Conversion level = 0.90:

The below table displays the variation in features (extracted using Euclidean distance) w.r.t. FAR and FRR over a large training set of twenty samples. The training dataset includes 20 signatures from 10 persons each and 12 testing signatures.

No. Of features Extracted/Used	No. of Original Signatures Identified as Fraud	No. of Fraud signatures identified as Original	No. of Correctly classified signatures	FAR (%age)	FRR (%age)
10	10	14	96	11.67	8.33
20	10	15	95	12.50	8.33
30	10	14	96	11.67	8.33
40	10	16	94	13.33	8.33
50	9	13	98	10.83	7.50
60	7	14	99	11.67	5.83
70	7	13	100	10.83	5.83
80	5	17	98	14.16	4.17

Table 4.8: Variation of FAR and FRR when training size=20 and conversion level=0.90

The above experimental results prove that as training size increases, the value of FAR also increases but the value of FRR decreases. This variation can easily be understood as follows:

When the size of training data increases, there is more information available for the machine to learn from hence it reduces the False Rejection rate but it also increases the false acceptance rate since skilled forgeries would now be considered as a deviation from original signature but within increased range. Hence the training size of 5 will produce best results in terms of FAR whereas a training size of 20 will produce best results in terms of FRR.

4.1.3 Variation of FAR and FRR w.r.t. training size for conversion level of 0.96:

1. Training size=5 and conversion level=0.96

The below table displays the variation in features (extracted using Euclidean distance) w.r.t. FAR and FRR over a small training set of five samples.
The training dataset includes 5 signatures from 10 persons each and 12 testing signatures where 8 of them are forged and 4 of them are genuine.

No. Of features Extracted/Used	No. of Original Signatures Identified as Fraud	No. of Fraud signatures identified as Original	No. of Correctly classified signatures	FAR (%age)	FRR (%age)
10	25	9	86	7.50	20.83
20	24	9	87	7.50	20.00
30	24	9	87	7.50	20.00
40	23	8	89	6.67	19.16
50	23	10	87	8.33	19.16
60	23	9	88	7.50	19.16
70	23	8	89	6.67	19.16
80	22	8	90	6.67	18.33

Table 4.9: Variation of FAR and FRR when training size=5 and conversion level=0.96

2. Training size=20 and conversion level=0.96

The below table displays the variation in features (extracted using Euclidean distance) w.r.t. FAR and FRR over a large training set of twenty samples.
The training dataset includes 20 signatures from 10 persons each and 12 testing signatures where 8 of them are forged and 4 of them are genuine.

No. Of features Extracted/Used	No. of Original Signatures Identified as Fraud	No. of Fraud signatures identified as Original	No. of Correctly classified signatures	FAR (%age)	FRR (%age)
10	15	15	90	12.50	12.50
20	12	14	94	11.67	10.00
30	13	15	92	12.50	10.83
40	11	15	94	12.50	9.17
50	10	16	94	13.33	8.33
60	9	17	94	14.16	7.50
70	8	14	98	11.67	6.67
80	10	17	93	14.16	8.33

Table 4.10: Variation of FAR and FRR when training size=20 and conversion level=0.96

4.1.4 Variation of FAR and FRR w.r.t. training size for conversion level of 0.92:

1. Training size=5 and conversion level=0.92

The below table displays the variation in features (extracted using Euclidean distance) w.r.t. FAR and FRR over a small training set of five samples.
The training dataset includes 5 signatures from 10 persons each and 12 testing signatures.

No. Of features Extracted/Used	No. of Original Signatures Identified as Fraud	No. of Fraud signatures identified as Original	No. of Correctly classified signatures	FAR (%age)	FRR (%age)
10	19	10	91	8.33	15.83
20	20	8	92	6.67	16.67
30	21	7	92	5.83	17.50
40	22	7	91	5.83	18.33
50	21	6	93	5.00	17.50
60	18	7	95	5.83	15.00
70	19	7	94	5.83	15.83
80	19	7	94	5.83	15.83

Table 4.11: Variation of FAR and FRR when training size=5 and conversion level=0.92

2. Training size=20 and Conversion level=0.92

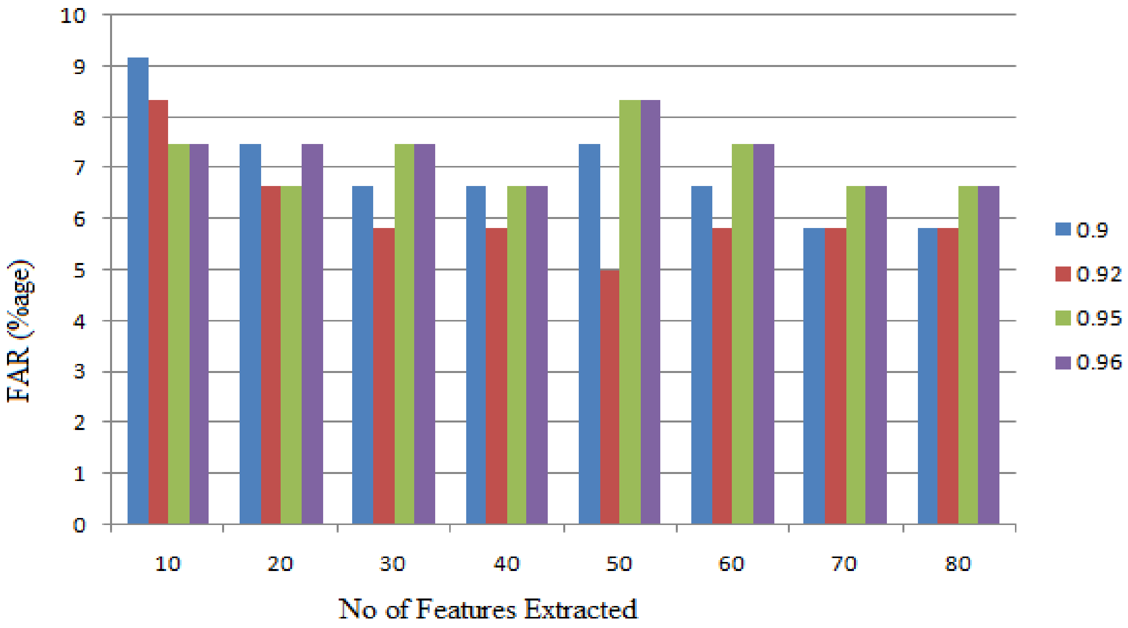
The below table displays the variation in features (extracted using Euclidean distance) w.r.t. FAR and FRR over a large training set of twenty samples.

The training dataset includes 20 signatures from 10 persons each and 12 testing signatures where 8 of them are forged and 4 of them are genuine.

No. Of features Extracted/Used	No. of Original Signatures Identified as Fraud	No. of Fraud signatures identified as Original	No. of Correctly classified signatures	FAR (%age)	FRR (%age)
10	12	14	94	11.67	10.00
20	10	15	95	12.50	8.33
30	11	15	94	12.50	7.50
40	10	15	95	12.50	8.33
50	10	12	98	10.00	8.33
60	7	14	99	11.67	5.83
70	7	13	100	10.83	5.83
80	5	17	98	14.16	4.17

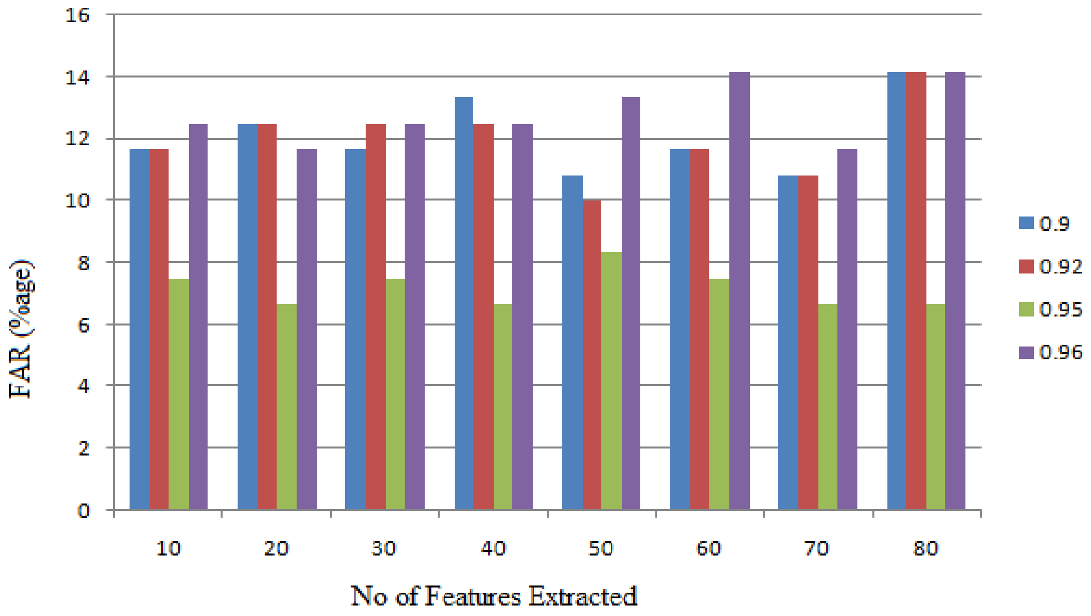
Table 4.12: Variation of FAR and FRR when training size=20 and conversion level=0.92

The below graph shows the variation of FAR w.r.t. number of features extracted for classification for 4 possible values of convert level when the training size is fixed to 5.



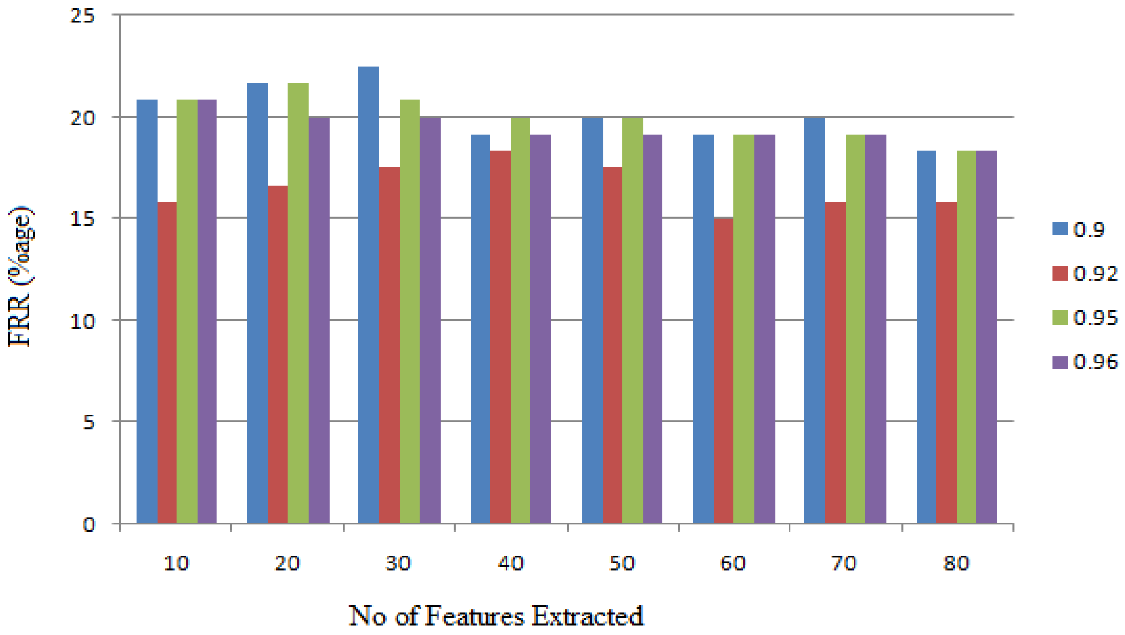
Graph 4.1: Variation of FAR (false acceptance rate) w.r.t. features for training size=5

The below graph shows the variation of FAR w.r.t. number of features extracted for classification for 4 possible values of convert level when the training size is fixed to 20 signatures



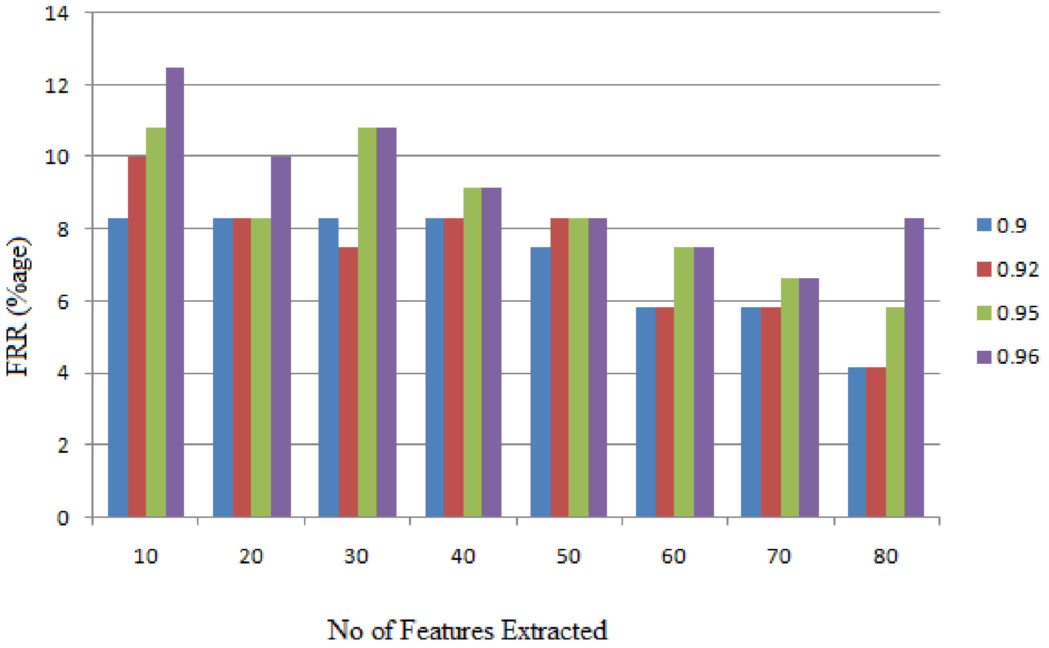
Graph 4.2: Variation of FAR (false acceptance rate) w.r.t. features for training size=20

The below graph shows the variation of FRR for training set of 5 signatures for variation in conversion levels w.r.t. No. of features:



Graph 4.3: Variation of FRR (false rejection rate) w.r.t. features for training size=5

The below graph shows the variation of FRR w.r.t. number of features extracted for classification for 4 possible values of convert level when the training size is fixed to 20 signatures:



Graph 4.4: Variation of FRR (false rejection rate) w.r.t. features for training size=20

From the above results it can be concluded that FAR is inversely proportional to the size of training set. The maximum efficiency achieved in terms of FAR is 5.00% for training size of 5 and conversion value of 0.9.

Also, it can also be deduced that FRR is inversely proportional to the size of training set. The maximum efficiency achieved in terms of FRR is 4.17% for training size of 20 and conversion value of 0.9 and 0.92 both.

The maximum correctness percentage achieved is 83.33% for training set size of 20 signatures for each person for a conversion level of 0.90 and no of extracted features being 70.
Average efficiency for a training size of 20 signatures for each person is computed to be 80.8% whereas average efficiency decreases when the size of training set decreases to 75% for a training size of 5.

CHAPTER 5

CONCLUSION AND FUTURE WORK

5.1 Conclusion

The above chapter of experimental results clearly demonstrates the variation of size of training set, conversion level and number of features w.r.t. FAR and FRR used to determine the efficiency of a signature verification system. Thus we may say that in a feature state space we have many variations and need to explore the entire space for finding the optimal value of parameters which define the method’s efficiency. The above study used a maximum of 20 training samples and 4 possible conversion levels which produced a peak correctness value of 83.33% along with minimum possible FAR and FRR values as 5.00% and 4.17%.

5.2 Future Work

Also other Machine Learning algorithms may be applied to compare the efficiency of the method w.r.t. other algorithms. A new learning mechanism called Deep Learning has been introduced in the field of Artificial Intelligence which has the capability to learn complex relations amongst data features when provided with enough training samples
A major drawback of deep learning is the availability of large training set in order to train the neural network. The method analyzed in this study works well for small training set and provides better results than other Machine Learning techniques mentioned in the above chapters.

The future work possible in this study would be to include new parameters determining the threshold values and also include a feature set based entirely on diagonal splitting of the signature image. Also studying the variation of the features set for large dataset is computationally expensive. Instead of using a centralized system for computation, parallel computing in MatLab may be introduced to speed up the computation and explore the parameter state space in less time and with ease. pMatlab is a library created by the M.I.T. to provide the feature of parallel computing and is widely used nowadays. The pMatlab library may be implemented for the same method and the results would be produced within a short interval of time.

REFERENCES

1. Debasish Jena, Banshidhar Majhi, Saroj Kumar Panigrahy, Sanjay Kumar Jena, “Improved Offline Signature Verification Scheme Using Feature Point Extraction Method”, IEEE, Dec 2008.
2. Banshidhar Majhi, Y Santhosh Reddy, D Prasanna Babu, “Novel Features for Off-line Signature Verification” International Journal of Computers, Communications & Control, Vol. I, No. 1, pp. 17-24, 2006.
3. J.J. Brault and R. Plamondon, "Segmanting Handwritten Signatures at Their Perceptually Important Points", IEEE Trans. Pattern Analysis and Machine Intelligence, Vol.15, No. 9, pp.953-957, Sept.1993.
4. B. Fang, C.H. Leung, Y.Y. Tang, K.W. Tse, P.C.K. Kwok and Y.K. Wong, "Off-line signature verification by the tracking of feature and stroke positions", Pattern Recognition 36, 2003, pp. 91–101.
5. Migual A. Ferrer, Jesus B. Alonso and Carlos M. Travieso, "Off-line Geometric Parameters for Automatic SignatureVerification Using Fixed-Point Arithmetic", IEEE Tran. on Pattern Analysis and Machine Intelligence, vol.27, no.6, June 2005.
6. R. Plamondon and S.N. Srihari, "Online and Offline Handwriting Recognition: A Comprehensive Survey", IEEE Tran. on Pattern Analysis and Machine Intelligence, vol.22 no.1, pp.63-84, Jan.2000.
7. A. Zimmer and L.L. Ling, "A Hybrid On/Off Line Handwritten Signature Verification System", Seventh International Conference on Document Analysis and Recognition, vol.1, pp.424-428, Aug.2003.
8. J Edson, R. Justino, F. Bortolozzi and R. Sabourin, "A comparison of SVM and HMM classifiers in the offlinesignature verification", Pattern Recognition Letters 26, 1377-1385, 2005.
9. J Edson, R. Justino, F. Bortolozzi and R. Sabourin, "An off-line signature verification using HMM for Random, Simple and Skilled Forgeries", Sixth International Conference on Document Analysis and Recognition, pp.1031-1034, Sept.2001.
10. J Edson, R. Justino, F. Bortolozzi and R. Sabourin, "The Interpersonal and Intrapersonal Variability Influences on Off-line Signature Verification Using HMM", Proc. XV Brazilian Symp. Computer Graphics and Image Processing, 2002, pp. 197-202, Oct.2002.
11. J Edson, R. Justino, A. El Yacoubi, F. Bortolozzi and R. Sabourin, "An off-line Signature Verification System Using HMM and Graphometric features", DAS 2000, pp. 211-222, Dec.2000.
12. B. Fang, C. Leung, Y. Tang, K. Tse, P. Kwok, and Y. Wong. Off-line signature verification by the tracking of feature and stroke positions. Pattern Recognition, 36(1):91–101, 2003.

13. A. Jain, F. Griess, and S. Connell. On-line signature verifi- cation. Pattern Recognition, 35(12):2963–2972, 2002.
14. M. Kalera, S. Srihari, and A. Xu. Offline signature verification and identification using distance statistics, 2004.
15. V. Nalwa. Automatic on-line signature verification. Proceedings of the IEEE, 85(2):215–239, 1997.
16. R. Plamondon and S. Srihari. Online and off-line handwriting recognition: a comprehensive survey. IEEE Trans. PAMI, 22(1):63–84, 2000.
17. Y. Qiao. Offline Signature Verification Using Online Handwriting Registration. Technical Report. Dept. of IE, The Chinese University of Hong Kong.