Offline Signature Verification for Detecting Signature Forgery: A Comparative Study

Anisha Soni [#1], Dharmendra Kumar Roy [*2]

[#]M.Tech Scholar, [*]Reader,
[#] *Computer Science & Engineering Department,
Rungta College of Engineering and Technology,
Kohka Kurud Road, Bhilai,
Chhattisgarh Swami Vivekanand Technical University, Bhilai, Chhattisgarh, India

***Abstract*— As signature is generally used as a means of individual verification, there is a need for an automatic verification system. Signatures provide a safe means of verification and authorization in authorized documents. However one of the key challenges is the ability of the system to detect skilled and unskilled forgery. Many cases of bank cheque forgeries have been reported. Most of the offline signature verification system adopts recognition based technique where the system classifies a given signature sample as one of the samples from the database. However detection of a forgery in a given sample is challenging as the input sample looks alike to one of the samples in the database. A simple and a consistent system has to be designed which should identify various types of forgeries. Various approaches have been used to implement biometric signature verification some of which are dynamic time warping (DTW), Bayesian Learning, Template Matching Technique, Hidden Markov Model (HMM), Support Vector Machine (SVM) etc. This paper presents a comparative and qualitative study of these methods used for offline signature verification.**

***Keywords*—Skilled and Unskilled Forgery, Signature Verification, Forgery detection, Dynamic Time Warping, Bayesian Learning, Template Matching Technique, Hidden Markov Model (HMM), Support Vector Machine (SVM).**

I. Introduction

Biometrics is widely implemented in today's world to deal with the security requirement issues. A biometric system can either do identification or verification task. In identification, the system can establish identity of a person whereas verification authenticates the person's claimed identity from the sample stored in the database [1]. Handwritten signatures are socially and legally accepted as a convenient means of writer verification. Signature verification offers a quick, simple and cost effective means for validating the authenticity of a document by determining the difference between an original signature and a forgery. Approaches to signature verification fall into two categories according to the acquisition of the data: On-line and Off-line. Online systems use dynamic information of a signature captured at the time the signature is made. Offline system means to verify a signature written on paper which is scanned to convert it into a digital image. As compared to online signature verification systems, off-line systems are difficult to design as many desirable characteristics such as the number of strokes, pressure applied, the speed of writing and other dynamic information are not available in the off-line case[5]. Our work focuses on the techniques of offline verification. In an off-line signature verification system, the task is to decide whether two signatures, given as scanned images are written by the same person or not. The system is initially trained using a database of signatures obtained from those individuals whose signatures have to be authenticated by the system. For each individual a mean signature is obtained integrating the features derived from a set of his/her genuine sample signatures. This mean signature acts as the template for verification against a claimed test signature. The objective of such a system is to distinguish between the original and forgery signatures. The forgeries involved in handwritten signatures have been categorized based on their characteristic features [2].

During verification two kinds of variation found in the signatures these are: Inter personal variability and Intra personal variability.

The intra personal is known as the variation among the signatures of the same signer it can be happen during illness, time and abnormal situations, whereas inter personal means the variation between originals and forgeries.

Forgery means someone attempt to copy someone else signature to steal properties of original signer[3][5]. The signature forgery can be classified into three categories:

1) Hit-or-miss Forgery: It is a very simple type of forgery and can be uncovered easily. The forger has no knowledge of the original signature and creates a signature in his own style. It is also known as Random Forgery.

2) Well-versed Forgery: In this type of forgery, the forger may be a master in imitating the original signature and may also have the knowledge about original signature that how it looks like. It is also known as Skilled Forgery.

3) Amateur Forgery: In Amateur forgery, the forger keeps an eye on the original signature and then tries to create a similar sign. Here, the forger is not an expert in forgery. It is also known as Simple Forgery.*[6]*

Signature authentication system involves two different but strongly associated tasks:

- Identify the owner of signature.

- Signature is authentic or not. [1][3][4]

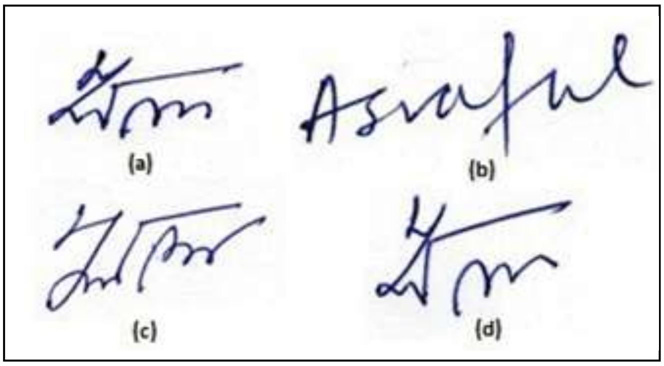

Figure 1.1 Signatures of Md. Asraful Haque:
(a) Original (b) Random forgery
(c) Simple forgery (d) Skilled forgery

Unlike other physiological biometrics, the characteristic of an individual's signature can only be established using an appropriate number of signature specimens. Since human signatures can vary over time, too few samples will increase the false rejection rate (FRR) of genuine signatures whilst too many samples will have the reverse effect of increasing the false acceptance rate (FAR). The collection of signatures from a large population for scientific research is not only labor intensive but also requires that the forgers are in possession of certain imitation skills [5].

There are two parameters by which we can define the performance of the signature verification:

- **FAR (False Acceptance Rate):** It is defined as the error rate when the signature is forged but the signature defines that it may be original .Actually, it is defined as the ratio of the no. of feature acceptances divided by the no. of identifications attempts.
- **FRR (False Rejection Rate):** It is defined to be forgery while it has been an original signature of a person. Actually it is defined as the ratio of the no. of false rejections identifications attempts [1][2][6][7].

II. RELATED METHODOLOGIES

A. *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. It can manage time duration varying signals such as signatures speech etc. For this reason it is popular for speech and signature recognition applications. 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 distribution of the different features that are used in the verification task are taken and a matching is done on it. 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. The verification score is a score that determine 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.

In an HMM model the states are hidden (i.e. it cannot be observed) and there are some other observations depending on the initial probabilities of these two terms the most likely state is determined using an algorithm like Baldi–Chauvin or Baum-Welch [1]. In signature verification the model can be represented as:

States= {genuine, forged}

Observations = {total time, velocity, pressure, no. of strokes}

B. *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 :

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

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. 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.[1]

C. *Dynamic Time Warping (DTW)*

DTW is the most popular technique for implementing signature verification. It is a method that determines the similarity between two time varying sequences. DTW can efficiently determine the most optimal distance between two sequences even if the accelerations of these time varying patterns are different. The most important feature of DTW is its ability to compute fast which makes it the most popular method in signature verification. It does not require huge data for training. It simply takes two sequences of time varying data or features and compares them and finds an optimal similarity between the two sample set. DTW uses a dynamic programming strategy that can manage the variability on the signatures length. In this method two signature samples are taken as sequences where points are taken in different discrete times. S={s1,s2,…,sn}, T={t1,t2,…,tm} are two time varying sequences that represents the value of the features at 1st,2nd and nth time. S is the sample signature stored in the database and T is the test signature sample. The time complexity of DTW is O(n2) where n is the number of points in the sequence. Although DTW is a fast technique but if the points taken on the sequence is very large then the time taken to compute the results in DTW becomes very high and therefore a variation of DTW i.e. VQ-DTW is used. VQ stand for vector Quantization. In this method clustering of some points that are in the same region are clustered together thus reducing the time complexity of algorithm [1].

D. *Template Matching Techniques*

Template matching approach is one of the simplest and earliest approaches to pattern recognition. Matching is a generic operation in pattern recognition, which is used to determine the similarity between two entities. Yoshimura et al. showed that a pattern matching method is able to achieve a good verification performance for Japanese signature. However, the similarity between two signatures obtained by a pattern matching method is affected by their stroke widths. The stroke widths vary with the pen used for signing.[2]

E. *Support Vector Machine (SVM)*

A support vector machine (SVM) is a tool used for classification and regression prediction and is based on machine learning theory in order to maximize predictive accuracy. The main aim of SVM is to draw a decision plane among a set of objects having different class memberships and classify them. There are two broad categories of classifiers one is linear and another is non-linear.SVM falls into the category of linear classifier. In case the data set is non-linear, SVM uses one of the four kernel functions to map the data such that they are linearly separable. A SVM generally aims at producing a large margin hyper plane, i.e. the perpendicular distance between the nearest point from the hyper plane and the hyper plane must be maximum. However in the real life scenario there exists over lapping data set and hence the SVM relies on loss functions. These loss functions ignore the errors that are present within certain range of the true value. Hard margin, L1 soft margin, L2 soft margin are the widely used epsilon intensive loss functions [8].

In other words, SVMs measure the complexity of hypotheses according to the margin, which separates the data. Thus, even with many features present, we can apply SVMs if input data is separable with a wide margin using functions from the hypothesis space [9].

III. CONCLUSION

From the above study it is clear that different methods are used for the signature verification. 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. In order to achieve more accuracy & optimize run time result can be achieved through Support Vector Machine classifier.

REFERENCES

[1] Zareen, F.J., and Jabin, S., “A Comparative Study of the Recent Trends in Biometric Signature Verification”, 2013IEEE.

[2] Haque, M.A. and Ali,T., “Improved Offline Signature Verification Method Using Parallel Block Analysis", 2012 International Conference on Recent Advances in Computing and Software Systems.

[3] Khalifa O., Alam M. K., Abdalla A. H. , An Evaluation on Offline Signature Verification using Artificial Neural Network Approach. 2013 International Conference On Computing, Electrical And Electronic Engineering (Icceee).

[4] Neerja Arora, Anil Kumar, Charu Jain, GMM For Offline Signature Forgery Detection, 2014 5th International Conference- Confluence The Next Generation Information Technology Summit (Confluence).

[5] Vu Nguyen, Yumiko Kawazoey, Tetsushi Wakabayashiy, and et. Al, Performance Analysis of the Gradient Feature and the Modified Direction Feature for Off-line Signature Verification, 2010 12th International Conference on Frontiers in Handwriting Recognition.

[6] Vaibhav Shah, Umang Sanghavi, Udit Shah Dwarkadas, Off-line Signature Verification Using Curve Fitting Algorithm with Neural Networks.

[7] M. Manoj Kumar, N. B. Puhan, Offline Signature Verification using the Trace Transform, 2014 IEEE International Advance Computing Conference (IACC).

[8] Kruthi.C, Deepika.C.Shet, Offline Signature Verification Using Support Vector Machine, 2014 Fifth International Conference on Signals and Image Processing.

[9] Emre Özgündüz,Tülin Şentürk and M. Elif Karslıgil, Off-Line Signature Verification And Recognition By Support Vector Machine.