

Computer Science And Engineering

University of Dhaka



Thesis report

Topic: Offline Signature Verification Scheme

Presented by:

Md. Mizanur Rahman (60)

Md. Faruk Hossain (84)

Md. Mahmudul Hasan (28)

Supervisor

Dr. Amin Ahsan Ali

Assistant Professor , Department of Computer Science & Engineering
University of Dhaka

A thesis paper submitted for the fulfillment of the partial requirement of Course CSE-3205(Math for Computer Science) for the Degree of Honours in Computer Science and Engineering ,University of Dhaka.

Abstract:

This report will present an investigation into one of the common used biometrics, personal signatures. From the perspective of system identification, the signature dynamics will be studied, and one novel verification algorithms will be discussed in details. We verified our algorithms using both public dataset and private dataset. The result will be shown and compared. Furthermore, one unfinished algorithms will also be proposed.

Index:

1.Introduction-----	3
2.Signature verification concept-----	4
3.Types of forgery-----	4
4.Methodology-----	5
4.1.Verification-----	6
4.2.Asymmetric binary dissimilarity-----	7
4.3.Signature verification-----	8
4.4.Result-----	9
5.Real data experiment-----	12
6.Further Investigation-----	13
6.1.Resizing alignment-----	13
6.2.Angular alignment-----	13
7.Conclusion-----	14
8.References-----	14

Introduction:

Handwritten signatures are widely accepted as a means of personal authentication and verification. So legality most documents like bank cheques, visa application and academic certificates, attendance register monitoring need to have authorized offline handwritten signatures. Today's society where forgery is rampant, there is the need for an automatic Handwritten signature verification (HSV) system to complement visual verification. Biometrics is the technological means that enables the identification or true verification of an individual from its physical or behavioral characteristics depending on their nature. It is classified into two categories namely behavioral and physiological. Where physiological biometrics measure some physical features of the subject like fingerprints, iris, hand and finger geometry which are stable over time. With the use of edge direction histogram derived from the edge map of the picture, only a small number of most possible intra prediction modes are chosen. Therefore the fast mode decision algorithm helps to speed up intra coding significantly. Usually, two acquisition modes are used for capturing the signature, which are off-line mode and on-line mode, respectively. The offline mode allows generating a handwriting static image from a scanning document and used for analysis. In contrast, the on-line mode allows generating from pen tablets or digitizers and analysis is based on dynamic information such as force, speed and rushing HSV systems are suited for forgery detection as they are cheap and non-intrusive and provide a direct link between the writer's identity and the transaction. The objective of signature verification systems is to differentiate between original and forged signature, which is related to intrapersonal and interpersonal variability. Intra-personal variation is variation among the signatures of the same person and inter-personal is the variation between the originals and the forgeries. There will always be slight variations in a human's handwritten signature, the consistency generated by natural motion and practice over time generates a recognizable pattern that makes the handwritten signature suitable for biometric identification.

Signature verification concept:

A signature is any written specimen in a person's own handwriting meant to be used for identification. A signature verification (SV) system authenticates the identity of any person, based on an analysis of his/her Signature through a set of processes which differentiates a genuine signature from a forgery signature. The precision of signature verification systems can be expressed by two types of error: the percentage of genuine signatures rejected as forgery which is called False Rejection Rate (FRR).dealing with any signature verification system, we take FRR as its performance estimate parameters.

Types of Forgery:

A signature forgery means an attempt to copy someone else's signature and use them against him to steal his identity there can be basically three types of forgeries. Both offline and online systems are used to detect various types of forgeries.

Classification of forgeries

Signature forgeries are classified as follows :

- 1) Random/simple or zero effort:** The forger doesn't have the shape of the writer signature but comes up with a draw of his own. He may derive this from the writer's name. This forgery accounts for majority of forgery cases though it's easy to detect with naked eyes.
- 2) Simple /casual forgery:** The forger knows the writer's signature shape and tries to imitate it without much practice.
- 3) Skilled forgeries:** This is where the forger has unrestricted access to genuine signature model and comes up with a forged sample. The skilled forgery category has been classified further into amateur and professional forgery. A professional forgery is done by a person with professional expertise in handwriting analysis and is able to come up with high quality forgery. The amateur forgeries are again categorized in the context of online verification into home improved and over the shoulder forgeries. In home improved the forger has a paper copy of the signature and has ample time to practice at home. The

reproduction is based on static features of the image. And over the shoulder forgeries are produced when immediately the forger has witnessed the writer make a genuine signature; the forger in this case has dynamic properties of signature and spatial image.

Methodology:

1. Data Acquisition For offline signature verification system, images of the signatures are scanned using a digital scanner. Scanned images are stored digitally for offline processing. There will be two images of signature will be stored as original signatures to compare with testing signature.

2. Preprocessing The purpose of pre-processing phase is to make signatures standard and ready for feature extraction. The pre-processing stage primarily involves some of the following steps:

- 1) **Noise reduction:** A noise filter is a normalization that applied to remove the noise caused during scanning and improves the quality of document.
- 2) **Resizing:** The image is cropped. Only 300 by 300 images will be allowed to be compared here. It may happen same person a sign in different point on the image which does not match the stored signature position like figures below:

(Note: Starting position means the lowest row and column index of image holds color points of signature)



Figure: Original Signature
(The starting position of the signature is (8,96) in original signature.)



Figure: Testing Signature

(The starting position of the signature is (56,88) in testing signature.)

It was needed to make the starting position of the testing signature is (8,96) before compare with original signature without changing if shape.

- 3) **Binarization:** it is the process of transformation from color to grayscale and then converts to binary image means a 2D array with only 0 or 1.
- 4) **Clutter Removal:** Any unconnected black dots are removed before processing and this is done by masking.

Verification:

A binary variable has only two states: 0 or 1, where 0 means that the variable is absent, and 1 means that it is present. Treating binary variables as if they are interval-scaled can lead to misleading clustering results. Therefore, methods specific to binary data are necessary for computing dissimilarities. One approach involves computing a dissimilarity matrix from the given binary data. If all binary variables are thought of as having the same weight, we have the 2-by-2 contingency table.

Contingency Table

		Original Signature		
		1	0	sum
Testing signature	1	q	r	$q + r$
	0	s	t	$s + t$
	sum	$q + s$	$r + t$	p

Where:

q , is the number of variables that equal 1 for both signature.

r , is the number of variables that equal 1 for testing signature but that are 0 for testing signature.

s , is the number of variables that equal 0 for testing signature but that are 1 for testing signature.

t , is the number of variables that equal 0 for both signature.

Now, dissimilarity between original and testing signature can be measured in two ways :

1. Symmetric Binary Dissimilarity:

$$d(i, j) = \frac{r + s}{q + r + s + t}$$

2. Asymmetric Binary Dissimilarity:

$$d(i, j) = \frac{r + s}{q + r + s}$$

Asymmetric binary dissimilarity:

A binary variable is asymmetric if the outcomes of the states are not equally important. Given two asymmetric binary variables, the agreement of two types of Data in Cluster Analysis 1s (a positive match) is then considered more significant than that of two 0s (a negative match). In this case if at a certain position of original and testing pixel array is 0 it will not be counted as matching.

For example: if a white image having no signature is being scanned and compared with original signature using **Symmetric Binary Dissimilarity** equation the dissimilarity % will be very low because the variable t will be very high and q will be 0. It may happen that white image will be accepted.

So, **Asymmetric binary dissimilarity** equation should be used to compute error percentage.

$$d(i, j) = \frac{r + s}{q + r + s}$$

Signature Verification:

There are two original signatures of a person stored in a system. When the testing signature appears, compare the testing signature with both the original signatures to find Asymmetric binary dissimilarity. Average the dissimilarity values. If the error rate is low, accept it; otherwise, reject it.



Fig: Two original signatures

Result:

Experiment on Same person:



Fig: Testing Signature (Same person).

Result(same person) :

Error percentage of signatures of same person

16.666666666666664

Accepted

Experiment on different person:



Fig:Testing Signature (forgeries).



Result(forgeries) :

Error percentage of signatures of forgeries

Average dissimilarity : 94.67656795204438

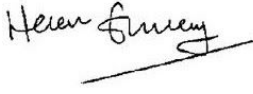
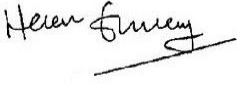
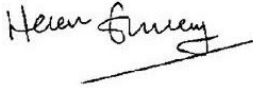

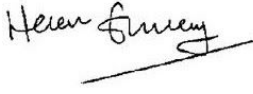
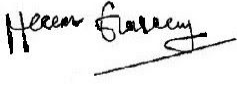
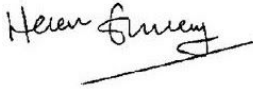
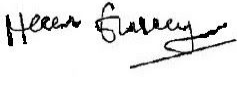
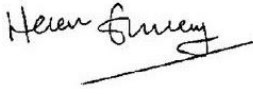
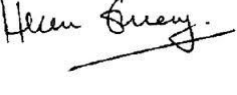
Not Accepted

Test Case 1:

Real Signature	Testing Signature	Error %	Accept/Reject
		16.6667	Accepted

		19.8255	Accepted
		22.1849	Not Accepted
		37.2080	Not Accepted
		94.67657	Not Accepted

Test Case 2:

No	Real Signature	Testing Signature	Error %	Accept/Reject
1			12.0127	Accepted
2			16.7193	Accepted
3			19.7135	Accepted
4			30.896	Not Accepted
5			76.5315	Not Accepted

Real data experiment:

To complete this report, we finally got some real world data by ourselves to test our algorithm. The data was collected manually, so we were only able to get a small set of signatures. We totally collected 2 genuine signatures from two users and check it 5 times with different persons for each signature. Sample plots are below :



Here according to the implemented software that we have made the first two signatures are genuine where is the last signature is fraud as the error percentage is too high.

Further Investigation:

According to the requirement of the signature verification software there are some features which need to be implemented. Two of them are:

Resizing alignment:

Resizing the digital signature is a prime requirement which ensures the accuracy of the signature verification. This can be attained through proper alignment of the signature with zoom in or zoom out feature. This features needs further investigation.

Angular alignment:

There is a probability that the signature is not angularly aligned with the test signatures. Therefore this rotational features can increase the accuracy of the verification.

Conclusion:

This paper presents a brief work on off-line signature recognition & verification. Two approaches are discussed here where one of them is implemented through Java programming . As we could observe that lots of work has already been done in the field of signature verification; there are still many challenges in this research area. The non-repetitive personality of variation of the signatures, because of age, sickness, geographic location and some extent the emotional state of the person, accentuates the

problem. Another problem associated with this category is, for security reasons, it is not very easy to make a signature dataset of real documents such as banking documents, and academic certificates are available for signature verification community. Publicly available signature datasets of real documents would make it possible for researchers to achieve a better performance in this field.

References:

<http://ijarcet.org/wp-content/uploads/IJARCET-VOL-2-ISSUE-8-2497-2503.pdf>
<https://web.stanford.edu/~boyd/vmls/vmls.pdf>