Advances in Computer Science and Information Technology (ACSIT)
p-ISSN: 2393-9907; e-ISSN: 2393-9915; Volume 3, Issue 4; April-June, 2016, pp. 339-344

http://www.krishisanskriti.org/Publication.html

Enhanced Offline Signature Verification Scheme Using Euclidean Distance for Feature Point Extraction

***Neelima Nausaliya*[1] *Ajay Khuteta*[2]**

[1,2]*Poornima College of Engineering, Jaipur, Rajasthan, 302020, India*
E-mail: [1]neelimanausaliya@gmail.com, [2]khutetaaajay@poornima.org

Abstract—*In this paper a novel scheme for offline signature verification has been proposed. The scheme is based on extracting multiple feature points from the geometric centre of the signature and comparing them with the existing trained feature points. The assortment of the feature points employs statistical parameters such as mean and variance. The suggested scheme distinguishes between two types of signatures i.e. original and forged. The method takes care of skill, simple and random forgeries. The objective of the work focuses on the reduction of the two crucial parameters False Acceptance Rate (FAR) and False Rejection Rate (FRR) usually used in any signature verification system. In the end comparative analysis has been done w.r.t. standard existing schemes.*

1. INTRODUCTION

Signature verification is considered as quite an important research area in the field of personal authentication. The recognition of an individual's handwriting is a significant concern which in turn enhances the interface between human-beings and computers [1, 8]. If the computer is intellectual enough to comprehend human handwriting, then it will deliver a more economic man-computer interface. In this extent signature is an exceptional case which provides secure means for authentication, authorization, attestation, in various high security environment. The objective of a signature verification system is to distinguish between two classes: the original and the forgery, which are related to inter and intrapersonal variability [1]. The disparity among signatures of same individual is called Intra Personal Variation whereas the variation between originals and forgeries is referred as Inter Personal Variation.

Signature verification is so diverse with the character acknowledgment, as it is often unreadable, which makes it just an image with some specific curves that signify the writing style of the individual. Signature is a distinct case of handwriting and often represented as a symbol. So, it is wise and necessary to assume a signature as a complete image with distinctive distribution of pixels and demonstrating a particular writing style and not just a collection of letters and words [7].

A signature verification system and the methods used to solve this problem can be bifurcated into two classes: online and off-line [9]. In the online system, a signature data can be acquired through an electronic tablet where the dynamic information on writing activity such as speed of writing, numbers of strokes, pressure applied, are available[4, 6]. In off-line systems, signatures inscribed on paper, done traditionally are transformed into electronic form with the assistance of a camera or a scanner. Here, the dynamic information is not available. In general, the dynamic information signifies the key writing style of a person. Since the volume of data available is less, the signature verification using off-line methods is relatively more difficult [2, 3].

Our work is comprehensive of the techniques for off- line signature verification. The static information resulting in an off-line signature verification scheme might be global, geometric, structural, or statistical. We implement offline signature verification which is focused on geometric centre and is beneficial in sorting out skilled forgeries from the originals. The algorithms implimented, have provided improved results in comparison with the previously suggested algorithms based on the geometric centre.

This paper is structured in the following sections. Section 1.1 delivers the different types of forgeries. Section 2, introduces a novel feature extraction technique. Section 3 debates over classification based on Euclidean distance model. Section 4 discusses over threshold selection. Section 5 depicts training, testing and then results. Section 6 delivers the concluding remarks.

1.1 Types of Forgeries

There are three diversified types of forgeries to take into consideration. First is called random forgery which is written by an individual who is unaware of the form of original signature. The second, known as simple forgery, is signified by a signature sample, is written by the individual who knows the shape of original signature without much practice. The third type is skilled forgery, designed by a proper imitation of the genuine signature model [4]. Every type of forgery demands several types of verification approaches [5]. Hybrid systems have been developed for the same [10]. Fig. 1 shows the various types of forgeries and how much they differ from original signature [1].

Fig. 1: (A) Original signature, (B) Random forgery,

(C) Simple forgery, (D) Skilled forgery

2. FEATURE EXTRACTION

Now, the geometric features rely on 2 sets of points in a 2-dimensional plane [7]. The vertical splitting of the image gives 30 feature points(v1,v2,v3,……..,v30) and the horizontal splitting has also thirty feature points(h1,h2,h3,…….,h30) .These feature points are achieved that are relative to a central geometric point of the image. Then the centered image is examined from left to right where the total number of black pixels is calculated. The same process is carried from top to bottom to calculate the black pixels. The image is divided into two halves w.r.t. the number of black pixels through two lines horizontally and vertically which intersect at a point described as a geometric centre. With reference to this point, we have extracted 60 feature points: 30 each for vertical and horizontal feature points of every signature image.

2.1 Processing of the Signature

We propose the geometric features that are based on two sets of points in a 2-dimensional plane. Each set has thirty feature points that represent the stroke distribution of signature pixels in the image. These sixty feature points are further calculated by Geometric Center.

Horizontal and vertical splitting are two major measures needed to retrieve these feature points. Before finding feature points, we have to make some alterations to the signature image [1]. The processing of the signature is given below.

2.2 Shifting signature to the centre of image

The signature is shifted to the centre by taking the image into an adjusted calculated frame and the redundant white spaces are removed without affecting the signature image in a way that the image comes in the middle of the frame. At first we split the whole frame of the signature into 10*10 square row and column-wise to find the variance (signature has to be binary consisting only black and white pixels). Whenever a square block has a zero variance, we eliminate that square, else it's restore. Thus squares of redundant white spaces are eliminated and then the image is reestablished in the fixed frame as depicted in Fig. 2.

2.3 Feature points through Vertical Splitting

Thirty feature points are collected from vertical splitting

the central feature point. This procedure for determining vertical feature points is given below: **Algorithm:**

Input: Static signature image after relocating it to the centre of the fixed sized frame.

Output: Vertical feature points: v1, v2, v3, v4, …, v29,

v30.

The steps are:

1) Divide the image with a vertical line passing through the geometric centre (v0) that divides the image into two halves: Left and Right part.
2) Find geometric centers v1 and v2 for left and right parts correspondingly.
3) Split the determined parts with horizontal lines through v1 and v2 to get four parts: Top-left, Top-right, Bottom-left and Bottom-right parts from which we attain v3, v4 and v5, v6.
4) Now, we split each part of the image through their geometric centers to find feature points v7, v8, v9, …, v13, v14.
5) Then we divide each of the parts once again to acquire all the 30 vertical feature points (as shown in Fig. 3).

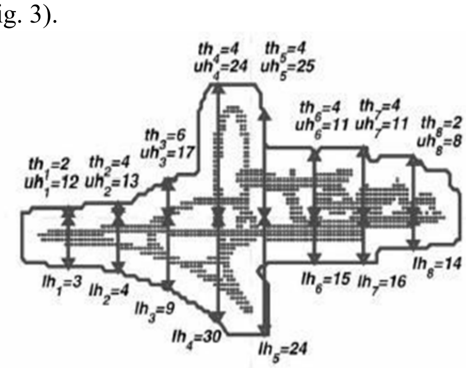

Fig. 3: Vertical splitting of the signature image

2.4 Feature points based on Horizontal Splitting Thirty feature points are collected after horizontal splitting w.r.t. the central feature point. The procedure for obtaining horizontal feature points is given below: **Algorithm:**

Input: Static signature image after relocating it to the centre of the frame of fixed size.

Output: Horizontal feature points: h1, h2, h3, h4,…, h29,

h30.

The steps are:

1) Break the image with a horizontal line passing through the geometric centre(h0) that divides the image into two halves: Top and Bottom part.
2) Find geometric centers h1 and h2 for top and bottom parts respectively.
3) Divide the top and bottom part with vertical lines as h1 and h2 so as to split 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) Now we divide each part of the image through their geometric centers to obtain feature points h7, h8, h9, ..., h13, h14.
5) Then, we divide these parts to obtain all the thirty vertical feature points (as shown in Fig. 3).

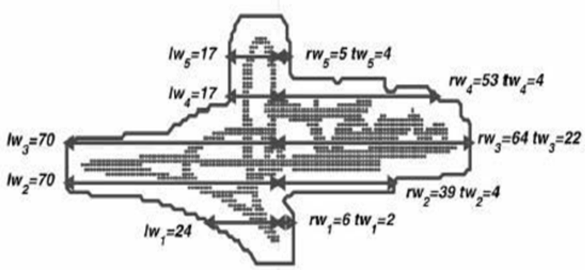

Fig. 4: Horizontal splitting of the signature image

3. CLASSIFICATION

In this paper, characteristics are based on the geometric properties. Hence, we use Euclidean distance model for the distribution. This is the distance between the pair of vectors of size 'n'. Vectors are only the feature points that were extracted with the size 2. Below is the procedure to calculate the distance Euclidean distance model. These distances are useful in threshold calculation.

3.1 Euclidean distance model

Let *A*(*a*1,*a* 2…..*an*) and *B*(*b*1,*b*2….*bn*) be the two vectors of size *n*. We can determine *distance*(*d*) by using equation 1.

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

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

(1) In our application, vectors are feature points on the plane. So *d* is the distance between these two points.

4. THRESHOLD

We have estimated individual thresholds for vertical and horizontal splitting. Here, we have suggested one method for threshold selection. Fig. 5 exhibits the variations in the identical feature points of training signatures. Let *n* be the number of training signatures and *x1, x2, …., xn* can be the corresponding single feature points of our training signatures. *xmedian* is the median of *n* features from *n* signatures.

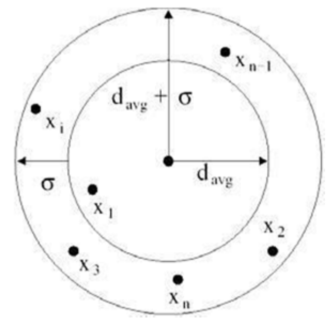

Fig. 5: Derivation of s (standard deviation) and davg (average distance) from distances

Let *d*1,*d*2,….,*d*n are distances given here,

$d1 = distance(x_{median};x1)$ $d2 = distance(x_{median};x2)$

…...

$dn = distance(x_{median};xn$ (2)

Two main factors we used in threshold calculation are d_{avg} and s. Equations 3 and 4 illustrates the calculation of these two parameters.

$d_{avg} = average(d1,d2,….,dn)$ (3)

$\sigma = SD(d1,d2,……,dn)$ (4)

There are total thirty different feature points forwhere davg1 is the average of Vd's for the first feature both vertical and horizontal splitting based on averagepoint of n signatures. Similarly we calculate davg2, distance (d_{avg}) and standard deviation (*s*). Equation 5 davg3, …, davg29, davg30 for the 2^{nd}, 3^{rd}, …, 29^{th}, and

shows the main formula for threshold. 30^{th} feature points.

Now we know that, Variance,

σ= *SD* (*d*1, *d*2,……, *dn*)

5. EXPERIMENTS AND RESULTS

(5)

Therefore,

$\sigma 1 = SD\ (Vd1,1; Vd1,2, \ldots, Vd1,n)$

$\sigma 2\ SD = (Vd2,1; Vd2,2, \ldots, Vd2,n)$

For research, we took 20 original signatures from……

every person and selected 10 for training. These original signatures were taken in different days. 4 forgeries taken by three persons each. Total 20 originals and 12

forgeries for each person signature are going to be tested.

$\sigma 30\ SD = (Vd30,1; Vd30,2, \ldots\ldots, Vd30,n)$

Hence,

There are two thresholds (one based on vertical splitting(7)

and another based on horizontal splitting) for each personNow we apply the same process to calculate the hthreshold signature.implimenting the horizontal feature points.

5.1. Training

$hpattern,1 = median(h1,1; h2,1; \ldots\ldots; hn,1)$

$hpattern,2 = median(h1,2; h2,2; \ldots\ldots; hn,2)$

Let n signatures be taken for training from each person.$hpattern,3 = median(h1,3; h2,3; \ldots\ldots; hn,3)$

There are 60 feature points from every original signature,$hpattern,4 = median(h1,4; h2,4; \ldots\ldots; hn,4)$ (8)

30 are taken by vertical splitting (Section 2.3) and 30 by horizontal splitting (Section 2.4). Individual patterns and thresholds will be calculated for vertical and horizontal splitting. Pattern points are based on vertical splitting.

$hpattern,29 = median(h1,29; h2,29; \ldots\ldots; hn,29)$

$hpattern,30 = median(h1,30; h2,30; \ldots\ldots; hn,30)$

$Vpattern,1 = median(v1,1; v2,1; \ldots\ldots; vn,1)$ $Vpattern,2 = median(v1,2; v2,2; \ldots\ldots; vn,2)$ $Vpattern,3 = median(v1,3; v2,3; \ldots\ldots; vn,3)$

$Vpattern,4 = median(v1,4; v2,4; \ldots\ldots; vn,4)$ (6)

……

$Vpattern,29 = median(v1,29; v2,29; \ldots\ldots; vn,29)$ $Vpattern,30 = median(v1,30; v2,30; \ldots\ldots; vn,30)$

where $vi,1; vi,2; \ldots\ldots; vi,30$ are taken as vertical splitting features of i^{th} training signature sample. Threshold as per vertical splitting is given below. Now, we will calculate the Vd, the distance of the first feature point of every training signature from the geometric centers.

$Vd_n = Distance\ (x_{median}, x_n)$

So,

$Vd1,1 = Distance(Vpattern,1\ ,\ V_{1,1})$ $Vd1,2 = Distance(Vpattern,1\ ,\ V_{1,2})$ $Vd1,3 = Distance(Vpattern,1\ ,\ V_{1,3})$

……

$Vd1,n = Distance(Vpattern,1\ ,\ V_{1,n})$

Therefore,

$davg1 = Average\ (Vd1,1; Vd1,2; Vd1,3; \ldots\ldots Vd1,n)$

Where $hi,1; hi,2; \ldots\ldots; hi,30$ are known as horizontal splitting features of ith training signature sample. Threshold based on horizontal splitting is shown below:

$$h_{threshold} = \sqrt{\sum_{i=1}^{30} (hd_{avg,i} + \sigma_{h,i})^2}$$

We will collect pattern points and thresholds of both vertical and horizontal. These values are useful in testing.

5.2. Testing

When new signature is up for testing we calculate features of vertical and horizontal splitting. Feature points based vertical splitting are $vnew;1$, $vnew;2$, $vnew;3$, $vnew;4, \ldots\ldots\ldots$ $vnew;29$, $vnew;30$. Distances between the new signature features and pattern feature points are related to vertical splitting are given below:

$vdnew;1 = distance(vpattern;1; vnew;1)$
$vdnew;2 = distance(vpattern;2; vnew;2)$
$vdnew;3 = distance(vpattern;3; vnew;3)$

$vdnew;4 = distance(vpattern;4; vnew;4)$ (10)

……

$vdnew;29 = distance(vpattern;29; vnew;29)$
$vdnew;30 = distance(vpattern;30; vnew;30)$

To arrange new signature validation we calculate $vdistance$ and compare this with $vthreshold$. If $vdistance$ is equal to or less than $vthreshold$, then new signature is suitable as per vertical splitting.

$$v_{distance} = \sqrt{\sum_{i=1}^{30} vd_{new,i}^2} \qquad (11)$$

Feature points based vertical splitting are $hnew;1$, $hnew;2$, $hnew;3$, $hnew;4$,…… $hnew;29$; $hnew;30$. Distances between new signature features and pattern feature points based on horizontal splitting are shown below:

$hdnew;1 = distance(hpattern;1;hnew;1)$
$hdnew;2 = distance(hpattern;2;hnew;2)$
$hdnew;3 = distance(hpattern;3;hnew;3)$
$hdnew;4 = distance(hpattern;4;hnew;4)$ (12)
……
$hdnew;29 = distance(hpattern;29;hnew;29)$ $hdnew;30 = distance(hpattern;30;hnew;30)$

For classification of new signature we have to calculate

$h_{distance}$ and compare this with $h_{threshold}$. If $h_{distance}$ is less than or equal to $h_{threshold}$ then new signature is acceptable by horizontal splitting.

$$h_{distance} = \sqrt{\sum_{i=1}^{30} hd_{new,i}^2} \quad (13)$$

New signature features have to fulfill the vertical splitting and horizontal splitting thresholds criteria.

5.3 Results

False Acceptance Rate (FAR) and *False Rejection Rate (FRR)* are the two factors used for measuring performance of any signature verification scheme. FAR is calculated by equation 14 and FRR is measured by equation 15.

$$FAR = \frac{number\ of\ forgeries\ accepted}{number\ of\ forgeries\ tested} \times 100 \quad (14)$$

$$FRR = \frac{number\ of\ originals\ rejected}{number\ of\ originals\ tested} \times 100 \quad (15)$$

Table1: Comparative analysis of FAR

Forgery	**Scheme**	**Scheme with 12 feature point**	**Scheme with 60 feature point**
RANDOM	5.61	2.08	0.43
SIMPLE	16.39	9.75	0.98
SKILL	19.3	16.36	2.08

Table2: Comparative analysis of FRR

FALSE REJECTION RATE(FRR)	
Existing Scheme	19.1
Existing Scheme(12 feature point)	11.58
Proposed Scheme(60 feature point)	4.83

6. CONCLUSION

The Algorithm which requires 60 feature points is more proficient and gives more precise results than the existing Techniques and sustains against the skilled forgeries. The results are taken as the FAR that is quite less when compared to the FARs of the already existing techniques. We also compared our proposed algorithms with different techniques based on feature extraction (12 feature points) and procedures demanding Polar and Cartesian coordinates. However, as our algorithm requires 60 feature points for calculating threshold, a small variation of a signature structure gives a large difference in the values of threshold distance from the calculated geometric center. Consequently, in our algorithm the FRR value is enhanced. So, it becomes important for a user to put his signature with utmost care in order to avoid a large variation of his signature w.r.t. his training signatures. Failing which there are chances of rejection of an original signature. Additionally, as we have extracted 30 feature points by each vertical and horizontal splitting, for calculating the threshold value, the time complexity is more than the time complexity of the existing technique that uses just 12 feature points.

REFERENCES

[1] 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.

[2] 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.

[3] J Edson, R. Justino, F. Bortolozzi and R. Sabourin, "A comparison of SVM and HMM classifiers in the off-linesignature verification", *Pattern Recognition Letters* 26, 1377-1385, 2005.

[4] 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.

[5] 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.

[6] 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.

[7] 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.

[8] 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.
[9] 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.
[10] 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.