Journal of Computer Science 4 (2): 111-116, 2008
ISSN 1549-3636

Improved Offline Signature Verification Scheme Using Feature Point Extraction Method

[1]Debasish Jena, [2]Banshidhar Majhi and [2]Sanjay Kumar Jena
[1]Centre for IT Education, Bhubaneswar, Orissa, India
[2]National Institute of Technology Rourkela, Orissa, India

Abstract: Signature verification is a technology that can solve security problems in our networked society. A new improved offline signature verification scheme is proposed which is based on selecting 60 feature points from the geometric centre of the signature and compares them with the already trained feature points. The classification of the feature points utilizes statistical parameters like mean and variance. The suggested scheme discriminates between two types of originals and forged signatures. The method takes care of skill, simple and random forgeries. The objective is to reduce the two vital parameters-False Acceptance Rate (FAR) and False Rejection Rate (FRR) which are normally used in any signature verification scheme. Comparative analysis has been made with standard existing schemes.

Key words: Offline signature, geometric centre, feature point, forgeries, euclidean distance model, FAR (False Acceptance Rate), FRR (False Rejection Rate).

INTRODUCTION

Signature verification is an important research area in the field of personal authentication. The recognition of human handwriting is important concerning about the improvement of the interface between human-beings and computers[1,8]. If the computer is intelligent enough to understand human handwriting it will provide a more attractive and economic man-computer interface. In this area signature is a special case that provides secure means for authentication, attestation authorization in many high security environment. The objective of the signature verification system is to discriminate between two classes: the original and the forgery, which are related to intra and interpersonal variability[1]. The variation among signatures of same person is called Intra Personal Variation. The variation between originals and forgeries is called Inter Personal Variation.

Signature verification is so different with the character recognition, because signature is often unreadable and it seems it is just an image with some particular curves that represent the writing style of the person. Signature is just a special case of handwriting and often is just a symbol. So it is wisdom and necessary to just deal with a signature as a complete image with special distribution of pixels and representing a particular writing style and not as a collection of letters and words[7].

A signature verification system and the techniques using to solve this problem can be divided into two classes: online and off-line[9]. In an online system, a signature data can be obtained from an electronic tablet and in this case, dynamic information about writing activity such as speed of writing, pressure applied, numbers of strokes are available[4,6]. In off-line systems, signatures written on study as has been done traditionally are converted to electronic form with the help of a camera or a scanner and obviously, the dynamic information is not available. In general, the dynamic information represents the main writing style of a person. Since the volume of information available is less, the signature verification using off-line techniques is relatively more difficult[2,3].

Our work is concerned with the techniques of off-line signature verification. The static information derived in an off-line signature verification system may be global, structural, geometric or statistical. We concern with offline signature verification which is based on geometric centre and is useful in separating skilled forgeries from the originals. The algorithms used have given improved results as compared to the previously proposed algorithms based on the geometric centre.

Types of Forgeries: There are three different types of forgeries to take into account. First one is random forgery which is written by the person who doesn't

Corresponding Author: Debasish Jena, Centre for IT Education, Bhubaneswar, Orissa, India

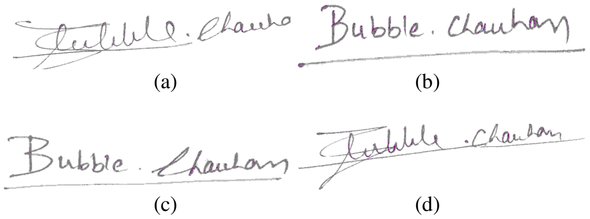

Fig. 1: (a): Original signature, (b): Random forgery, (c): Simple forgery and (d): skilled forgery

know the shape of original signature. The second, called simple forgery, which is represented by a signature sample, written by the person who knows the shape of original signature without much practice. The last type is skilled forgery, represented by a suitable imitation of the genuine signature model[4]. Each type of forgery requires different types of verification approach[5]. Hybrid systems have also been developed[10]. Fig. 1 shows the different types of forgeries and how much they vary from original signature[1].

Feature Extraction: The geometric features are based on two sets of points in 2-dimensional plane[7]. The vertical splitting of the image results thirty feature points (v1, v2, v3,…,v30) and the horizontal splitting results 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 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 centre. With reference to this point we extracted 60 feature points: 30 vertical and 30 horizontal feature points of each signature image.

Processing of the Signature: The geometric features proposed here are based on two sets of points in two-dimensional plane. Each set having thirty feature points which represent the stroke distribution of signature pixels in image. These sixty feature points are calculated by Geometric Center. Vertical Splitting and Horizontal Splitting are two main steps to retrieve these feature points. Before finding feature points we have to do some adjustments to the signature image[1]. The processing of the signature is discussed below.

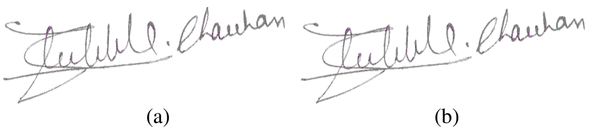

Fig. 2: Captured signature before adjustment (a) and after adjustment (b)

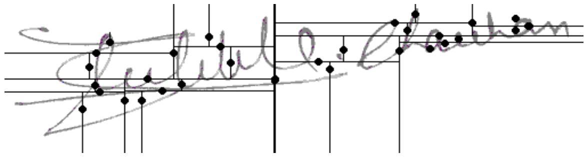

Fig. 3: Vertical splitting of the signature image

Moving signature into the centre of image: The signature is moved to the centre by taking the signature image into a fixed calculated frame and the unnecessary white spaces are removed without affecting the signature image such that the image is in the middle of the frame. For this first we divide the whole frame of the signature into 10*10 square row-wise and column-wise and find the 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. Thus squares of unnecessary white spaces are removed and then the image is restored in the fixed frame as shown in Fig. 2.

Thirty feature points are obtained based on vertical splitting w.r.t. the central feature point. The procedure for finding vertical feature points are given below.

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

Output: Vertical feature points: v1, v2, v3, v4,…v29, v30.

The steps are:

- Split the image with a vertical line passing through the geometric centre (v0) which divides the image into two halves: Left part and Right part.
- Find geometric centers v1 and v2 for left and right parts correspondingly.
- Split 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.
- We again split each part of the image through their geometric centers to obtain feature points v7, v8, v9,….v13, v14.
- Then we split each of the parts once again to obtain all the thirty vertical feature points (as shown in Fig. 3).

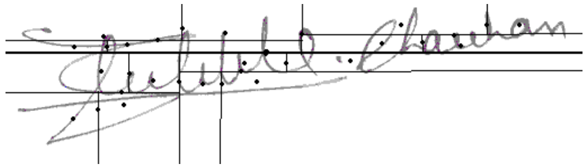

Fig. 4: Horizontal splitting of the signature image

Feature points based on Horizontal Splitting: Thirty feature points are obtained based on horizontal splitting w.r.t. the central feature point. The procedure for finding horizontal feature points are given below.

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

Output: Horizontal feature points: h1, h2, h3, h4,….h29, h30.
The steps are:

- Split the image with a horizontal line passing through the geometric centre(h0) which divides the image into two halves: Top part and Bottom part.
- Find geometric centers h1 and h2 for top and bottom parts correspondingly.
- 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.
- We again split each part of the image through their geometric centers to obtain feature points h7,h8,h9,…….h13,h14.
- Then we split each of the parts once again to obtain all the thirty vertical feature points (as shown in Fig. 4).

Classification: In this study 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 (a1,a2…..an) and B (b1,b2….bn) are two vectors of size n. We can calculate distance (d) by using Eq. 1:

$$\text{Dis tan ce}(d) = \sqrt{\sum_{t=1}^{n}(a_t - b_t)^2} \qquad (1)$$

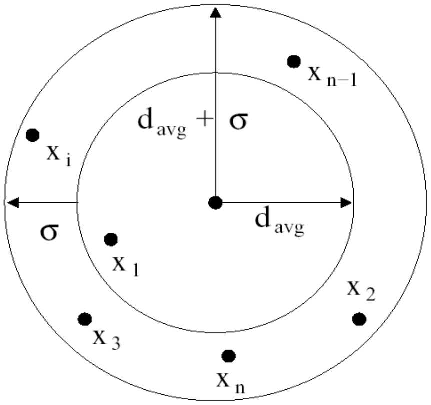

Fig. 5: d_{avg} (average distance) and s (standard deviation) derivation from distances

In our application, vectors are feature points on plane. So d is the simple distance between two points.

Threshold: We have calculated individual thresholds for vertical splitting and horizontal splitting. Here we proposed one method for threshold selection. Fig. 5 shows the variations in single corresponding feature points of training signatures. Let n is the number of training signatures and x1, x2, ….,xn are corresponding single feature points of training signatures (taking one corresponding feature point from each signature). x_{median} is the median of n features from n signatures.

Let d1,d2,….,dn are distances defined here:

$$\begin{aligned} d1 &= \text{distance}(x_{median}, x1) \\ d2 &= \text{distance}(x_{median}, x2) \\ &.... \\ dn &= \text{distance}(x_{median}, xn) \end{aligned} \qquad (2)$$

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

$$d_{avg} = \text{average}(d1, d2, ..., dn) \qquad (3)$$

$$\sigma = \text{SD}\ (d1, d2, ..., dn) \qquad (4)$$

Like this total thirty different feature points are there for both vertical and horizontal splitting based on average distance (d_{avg}) and standard deviation (s). Eq. 5 shows the main formula for threshold.

$$\text{Threshold}(t) = \sqrt{\sum_{t=1}^{t^{30}}(d_{avg,t} - \sigma_t)^2} \qquad (5)$$

RESULTS AND DISCUSSION

For experiment we took 24 original signatures from each person and selected 8 for training. These original signatures are taken in different days. Forgeries taken by three persons and 8 from each. Total 16 originals and 24 forgeries for each person signature are going to be tested. There are two thresholds (one based on vertical splitting and another based on horizontal splitting) for each person signature.

Training: Let n signatures are taking for training from each person. There are 60 feature points from each original signature, 30 are taken by vertical splitting (Section 2.3) and 30 are taken by horizontal splitting (Section 2.4). Individual thresholds and patterns are calculated for vertical splitting and horizontal splitting. Pattern points based on vertical

splitting are shown below:

$$\begin{aligned}
&\text{Vpattern},1 = \text{median}(v1,1,v2,1,...,vn,1)\\
&\text{Vpattern},2 = \text{median}(v1,2,v2,2,...,vn,2)\\
&\text{Vpattern},3 = \text{median}(v1,3,v2,3,...,vn,3)\\
&0\ \text{Vpattern},4 = \text{median}(v1,4,v2,4,...,vn,4)\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&\text{Vpattern},29 = \text{median}(v1,29,v2,29,...,vn,29)\\
&\text{Vpattern},30 = \text{median}(v1,30,v2,30,...,vn,30)
\end{aligned}\tag{6}$$

Where vi,1,vi,2,……,vi,30 are vertical splitting features of ith training signature sample. Threshold based on vertical splitting is shown below:

Now we have to calculate the Vd which is the distance of the first feature point off all the training signatures from the geometric centers.

$\text{Vd}_n = \text{Distance}(x_{\text{median}}, x_{n)}$

So,

$\text{Vd}1,1 = \text{Distance}(\text{Vpattern},1\ ,\ V_{1,1})$

$\text{Vd}1,2 = \text{Distance}(\text{Vpattern},1\ ,\ V_{1,2})$

$\text{Vd}1,3 = \text{Distance}(\text{Vpattern},1\ ,\ V_{1,3})$

…………………………………..

…………………………………..

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

Therefore,

$$\text{davg}1 = \text{Average}(\text{Vd}1,1,\ \text{Vd}1,2,\ \text{Vd}1,3,\ \ldots\ldots\text{Vd}1,n)$$

where, davg1 is the average of Vd's for the first feature point of n signatures.

Similarly we calculate davg2, davg3,…….,davg29, davg30 for the 2nd, 3rd, ………,29th and 30th feature points.

Now we know that, Variance,

$$\sigma = \text{SD}\ (d1,d2,\ldots\ldots,dn)$$

Therefore,

$\sigma1 = \text{SD}\ (\text{Vd}1,1,\ \text{Vd}1,2,\ldots\ldots,\text{Vd}1,n)$

$\sigma2 = \text{SD}\ (\text{Vd}2,1,\text{Vd}2,2\ldots\ldots,\text{Vd}2,n)$

…………………………………….

…………………………………….

…………………………………….

$\sigma30 = \text{SD}\ (\text{Vd}30,1,\ \text{Vd}30,2,\ldots\ldots,\text{Vd}30,n)$

Hence,

$$\text{vthreshold} = \sqrt{\sum_{t=1}^{30}(\text{vd}_{\text{avg},t} - \sigma_{v,t})^2}\tag{7}$$

Now we apply the same procedure to calculate the $\text{h}_{\text{threshold}}$ using the horizontal feature points.

$$\begin{aligned}
&\text{hpattern},1 = \text{median}(h1,1,h2,1,...,hn,1)\\
&\text{hpattern},2 = \text{median}(h1,2,h2,2,...,hn,2)\\
&\text{hpattern},3 = \text{median}(h1,3,h2,3,...,hn,3)\\
&\text{hpattern},4 = \text{median}(h1,4,h2,4,...,hn,4)\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&\text{hpattern},29 = \text{median}(h1,29,h2,29,...,hn,29)\\
&\text{hpattern},30 = \text{median}(h1,30,h2,30,...,hn,30)
\end{aligned}\tag{8}$$

where, hi,1,hi,2,……,hi,30 are horizontal splitting features of ith training signature sample. Threshold based on horizontal splitting is shown below:

$$\text{hthreshold} = \sqrt{\sum_{t=1}^{30}(\text{hd}_{\text{avg},t} - \sigma_{h,t})^2}\tag{9}$$

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

Testing: When new signature comes for testing we have to calculate features of vertical splitting and horizontal splitting. Feature points based vertical splitting are vnew,1, vnew,2, vnew,3, vnew,4,… vnew,29, vnew,30. Distances between new signature features and pattern feature points based on vertical splitting are shown below:

$$
\begin{aligned}
&\text{vdnew,1} = \text{distance(vpattern,1,vnew,1)}\\
&\text{vdnew,2} = \text{distance(vpattern,2,vnew,2)}\\
&\text{vdnew,3} = \text{distance(vpattern,3,vnew,3)}\\
&\text{vdnew,4} = \text{distance(vpattern,4,vnew,4)}\\
&\dots\dots\dots\dots\dots\dots\dots\dots\\
&\dots\dots\dots\dots\dots\dots\dots\dots\\
&\dots\dots\dots\dots\dots\dots\dots\dots\\
&\text{vdnew,29} = \text{distance(vpattern,29,vnew,29)}\\
&\text{vdnew,30} = \text{distance(vpattern,30,vnew,30)}
\end{aligned}
\qquad (10)
$$

For classification of new signature we have to calculate vdistance and compare this with vthreshold. If vdistance is less than or equal to vthreshold then new signature is acceptable by vertical splitting:

$$\text{vdis}\tan\text{ce} = \sqrt{\sum_{t=1}^{30} \text{vd}_{\text{new},t}^{2}} \qquad (11)$$

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:

$$
\begin{aligned}
&\text{hdnew,1} = \text{distance(hpattern,1,hnew,1)}\\
&\text{hdnew,2} = \text{distance(hpattern,2,hnew,2)}\\
&\text{hdnew,3} = \text{distance(hpattern,3,hnew,3)}\\
&\text{hdnew,4} = \text{distance(hpattern,4,hnew,4)}\\
&\dots\dots\dots\dots\dots\dots\dots\dots\\
&\dots\dots\dots\dots\dots\dots\dots\dots\\
&\dots\dots\dots\dots\dots\dots\dots\dots\\
&\text{hdnew,29} = \text{distance(hpattern,29,hnew,29)}\\
&\text{hdnew,30} = \text{distance(hpattern,30,hnew,30)}
\end{aligned}
\qquad (12)
$$

For classification of new signature we have to calculate h_{distance} and compare this with $h_{\text{threshold}}$. If h_{distance} is less than or equal to $h_{\text{threshold}}$ then new signature is acceptable by horizontal splitting:

$$\text{hdis}\tan\text{ce} = \sqrt{\sum_{t=1}^{30} \text{hd}_{\text{new},t}^{2}} \qquad (13)$$

New signature features have to satisfy both vertical splitting and horizontal splitting thresholds.

False Acceptance Rate (FAR) and False Rejection Rate (FRR) are the two parameters using for measuring performance of any signature verification method. FAR is calculated by Eq. 14 (Table 1) and FRR is calculated by Eq. 15 (Table 2):

$$\text{FAR} = \frac{\text{Number of forgeries accepted}}{\text{Number of forgeries tested}} \times 100 \qquad (14)$$

$$\text{FAR} = \frac{\text{Number of originals rejected}}{\text{Number of originals tested}} \times 100 \qquad (15)$$

Table 1: Comparative analysis of FAR

Forgery type	Existing scheme	Existing scheme with 12 feature point	Proposed 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

Table 2: Comparative analysis of FRR

False rejection rate (FRR)	
Existing scheme	19.10
Existing scheme (12 feature point)	14.58
Proposed scheme (60 feature point)	20.83

CONCLUSION

The Algorithm which is based on the 60 feature points is more efficient and gives more accurate results than the existing Techniques and survives against the skilled forgeries. The algorithm results the FAR which is very much less as compared to the FARs of the previously existing techniques. We compared our algorithms with other techniques based on feature extraction(12 feature points) and techniques based on Polar and Cartesian coordinates. But as our algorithm takes 60 feature points for threshold calculations, a small variation of a signature results in a large change in the values of threshold distance from the geometric center. Therefore in our algorithm the FRR value is increased. So it is important for a user to sign his signature with utmost care so that there is not a large variation of his signature to his training signatures. Otherwise there is a probability of rejection of an original signature. Moreover, since we have extracted 30 feature points by vertical splitting and 30 by horizontal splitting for the calculation of the threshold value, the time complexity is higher than the time complexity of the existing technique which uses 12 feature points for threshold calculations.

REFERENCES

1. Banshidhar Majhi, Y. Santhosh Reddy and D. Prasanna Babu, 2006. Novel features for off-line signature verification. Int. J. Comput. Commun Control, 1: 17-24. http://www.journal.univagora.ro/download/pdf/20.pdf.

2. Brault, J.J. and R. Plamondon, 1993. Segmanting handwritten signatures at their perceptually important points. IEEE Trans. Pattern Anal. Mach. Intel., 15: 953-957. http://doi.ieeecomputersociety.org/10.1109/34.232079.
3. Edson, J., R. Justino, F. Bortolozzi and R. Sabourin, 2005. "A comparison of SVM and HMM classifiers in the off-line signature verification. Pattern Recognit. Lett., 26: 1377-1385. http://dx.doi.org/10.1016/j.patrec.2004.11.015.
4. Edson, J., R. Justino, F. Bortolozzi and R. Sabourin, 2001. An off-line signature verification using HMM for random, simple and skilled forgeries. In: 6th International Conference on Document Analysis and Recognition, September, pp: 1031-1034. http://doi.ieeecomputersociety.org/10.1109/ICDAR.2001.95394
5. Edson, J., R. Justino, F. Bortolozzi and R. Sabourin, 2002. The interpersonal and intrapersonal variability influences on off-line signature verification using HMM. Proceeding XV Brazilian Symposium Computer Graph and Image Processing, pp: 197-202. http://ieeexplore.ieee.org/iel5/8352/26315/01167143.pdf?arnumber=1167143.
6. Edson, J., R. Justino, A. El Yacoubi, F. Bortolozzi and R. Sabourin, 2000. An off-line Signature Verification System Using HMM and Graphometric features. DAS, pp: 211-222. http://www.livia.etsmtl.ca/publications/2000/JustinoDAS.pdf.
7. Fang, B., C.H. Leung, Y.Y. Tang, K.W. Tse, P.C.K. Kwok and Y.K. Wong, 2003. Off-line signature verification by the tracking of feature and stroke positions. Patter Recognit., 36: 91-101. http://dx.doi.org/10.1016/S0031-3203(02)00061-4.
8. Migual, A.F., B.A. Jesus and M.T. Carlos, 2005. Off-line geometric parameters for automatic signatureverification using fixed-point arithmetic. IEEE Tran. Pattern. Anal. Mach. Intel., 27: pp. 993-97http://dx.doi.org/10.1109/TPAMI.2005.125
9. Plamondon, R. and S.N. Srihari, 2000. Online and offline handwriting recognition: A comprehensive survey. IEEE Tran. Pattern Anal. Mach. Intel., 22: 63-84. http://doi.ieeecomputersociety.org/10.1109/34.824821.
10. Zimmer, A. and L.L. Ling, 2003. A hybrid on/off line handwritten signature verification system. In: 7th International Conference on Document Analysis and Recognition, 3-6 August 2003, Edinburgh, Scotland, UK. IEEE Computer Society 2003pp:424-428. http://doi.ieeecomputersociety.org/10.1109/ICDAR.2003.1227702.