International Journal of Computers, Communications & Control
Vol. I (2006), No. 1, pp. 17-24

Novel Features for Off-line Signature Verification

Banshider Majhi, Y Santhosh Reddy, D Prasanna Babu

Abstract: In this paper a novel feature extraction scheme has been suggested for offline signature verification. The proposed method used geometric center for feature extraction. Euclidean distance model was used for classification. This classifier is well suitable for features extracted and fast in computation. Method proposed in this paper leads to better results than existing offline signature verification methods. Threshold selection is based on statistical parameters like average and standard deviation (σ).
Keywords: Feature Extraction, Geometric Center, Euclidean Distance Model, Standard Deviation and Off-line Signature Verification.

1 Introduction

Signature verification is an important research area in the field of person authentication. We can generally distinguish between two different categories of verification systems: online, for which the signature signal is captured during the writing process, thus making the dynamic information available, and offline for which the signature is captured once the writing processing is over and, thus, only a static image is available[8]. 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. The variation among signatures of same person is called Inrea Personal Variation. The vatiation between originals and forgeries is called Inter Personal Variation[7].
In this paper we concentrated on Offline Verification System. Upto now many signature verification methods proposed based on different strategies but no verification system classified near forgeries which were classified by this method. And the main advantage of this algorithm is efficiency and computational complexity. For general purpose applications like smart cards we want quick and efficient verification system[2]. This method is based on the Geometric Center and signature strokes distribution. Section 2 discusses the feature extraction from signature. This is a recursive method which applying on signature recursively. A lot of work has been done in the field of automatic off-line signature verification. While a large portion of work is focused on random forgery detection, more efforts are still needed to address the problem of skilled forgery detection[6]. Our method will be the first verification system which seperates some skilled forgeries from originals.
This paper organized in the following sections: Section 1.1 provides the different types of forgeries. Section 2 introduces new feature extraction method. Section 3 discusses classification based on Euclidean distance model. Section 4 discussed about threshold selection. Section 5 shows training, testing and results and Section 6 gives conclusion and furthure working directions.

1.1 Types of forgeries

There are three different types of forgeries to take into account. The first, known as *random forgery* which writtn by the person who don't know the shape of original signature. The second, called *simple forgery*, is represented by a signature sample which written by the person who know the shape of original signature without much practice. The last type is *skilled forgery*, represented by a suitable imitation of the genuine signature model[3]. Each type of forgery requires different types of verification approach[4]. Hybrid systems have also been developed[9] Fig. 1 shows the different types of forgeries and how much they are varies from original signature[5].

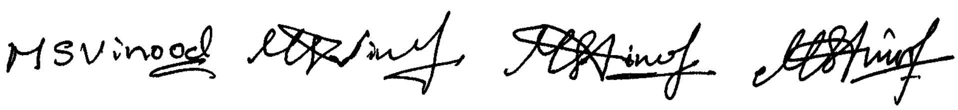

Figure 1: (a) Random Forgery (b) Simple Forgery (c) Skilled Forgery (d)Original Signature

By using this method we can easily eliminate random and simple forgeries. Some of the skilled forgeries also eliminated.

Copyright © 2006 by CCC Publications

Figure 2: (a) Before adjustment of signature (b) After adjustment of signature

2 Feature Extraction

The geometric features proposed by this paper are based on two sets of points in two-dimentional plane. Each set having six feature points which represent the stroke distribution of signature pixels in image. These twelve feature points are calculated by Geometric Center[1]. Vertical Splitting and Horizontal Splitting are two main steps to retrieve these feature points. Vertical Splitting is discussed in Section 2.2 and Horizontal Splitting is discussed in Section 2.3.
Before finding feature points we have to do some adjustments to the signature image. That is moving signature strokes to the center of the image which discussed in Section 2.1.

2.1 Moving signature to the center of image

In this step signatures are moving to the center of image. Because of this we can reduce intra-personal variations. Here first we have to find out the geometric center of the image and move the signature pixels such that the geometric center should reside at center of image. Fig. 2 shows the signature images before moving and after moving.

2.2 Feature points based on vertical splitting

Six feature points are retrieving based on vertical splitting. Here feature points are nothing but geometric centers. The procedure for finding feature points by vertical splitting is mentioned in Algorithm.

Algorithm

This is the procedure for generating feature points based on verical splitting.
Input: Static signature image after moving the signature to center of image
Output: $v_1, v_2, v_3, v_4, v_5, v_6$ (feature points)
(a)Split image with vertical line at the center of image then we will get left and right parts of image.
(b)Find geometric centers v_1 and v_2 for left and right parts correspondingly.
(c)Split left part horizontal line at v_1 and find out geometric centers v_3 and v_4 for top and bottom parts of left part currespondingly.
(d)Split right part horizontal line at v_2 and find out geometric centers v_5 and v_6 for top and bottom parts of left part currespondingly.

Fig. 3 shows the feature points retrieved from signature image and O is the center of image. These features we have to calculate for every signatrure image in both training and testing.

2.3 Feature points based on horizontal splitting

Six feature points are retrieving based on horizontal splitting. Here feature points are nothing but geometric centers. The procedure for finding feature points by horizontal splitting is mentioned in Algorithm.

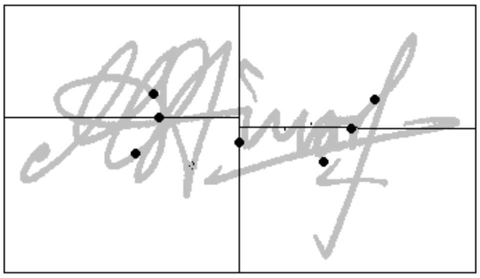

Figure 3: Feature points based on vertical splitting

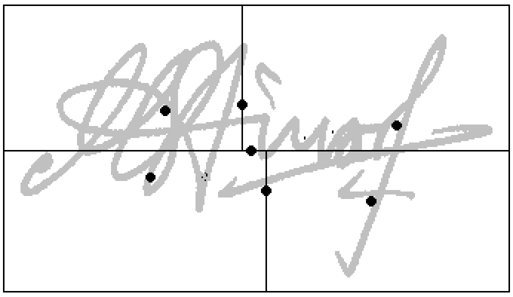

Figure 4: Feature points based on horizontal splitting

Algorithm

This is the procedure for generating feature points based on horizontal splitting.
Input: Static signature image after moving the signature to center of image
Output: $h_1, h_2, h_3, h_4, h_5, h_6$ (feature points)
(a)Split image with horizontal line at the center of image then we will get top and bottom parts of image.
(b)Find geometric centers h_1 and h_2 for top and bottom parts correspondingly.
(c)Split top part with vertical line at h_1 and find out geometric centers h_3 and h_4 for left and right parts of top part curresspondingly.
(d)Split bottom part with vertical line at h_2 and find out geometric centers v_5 and h_6 for left and right parts of left part curresspondingly.

Fog. 4 shows the feature points retrieved from signature image and O is the center of image. These features we have to calculate for every signatrure image in both training and testing. Now total twelve feature points ($v_1, ..., v_6$ and $h_1, ..., h_6$) are calculated by vertical and horizontal splittings. In Section 4 we will see how each feature point can classify.

3 Classification

In this paper features are based on geometric properties. So we used 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 eucliden distance model is descibed in Section 3.1. In threshold calculation these distances are useful.

3.1 Euclidean distance model

Let $A(a_1, a_2, ..., a_n)$ and $B(b_1, b_2, ..., b_n)$ are two vectors of size n. We can calculate $distance(d)$ by using equarion 1.

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

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

4 Threshold

Individual thresholds for vertical splitting and horizontal splitting. Here we proposed one method for threshold selection which used in Section 5.1. Fig. 5 shows the variations in single curresponding feature points of training signatures. Let n is the number of training signatures and $x_1, x_2, ..., x_n$ are curresponding single feature points of training signatures(taking one curresponding feature point from each signature). x_{median} is the median of n features from n signatures. Let $d_1, ..., d_n$ are distances defined here,

$$\begin{aligned} d_1 &= distance(x_{median}, x_1) \\ d_2 &= distance(x_{median}, x_2) \\ &\vdots \\ d_n &= distance(x_{median}, x_n) \end{aligned} \tag{2}$$

Two main parameters we used in threshold calculation are d_{avg} and σ. Equations 3 and 4 shows the calculation

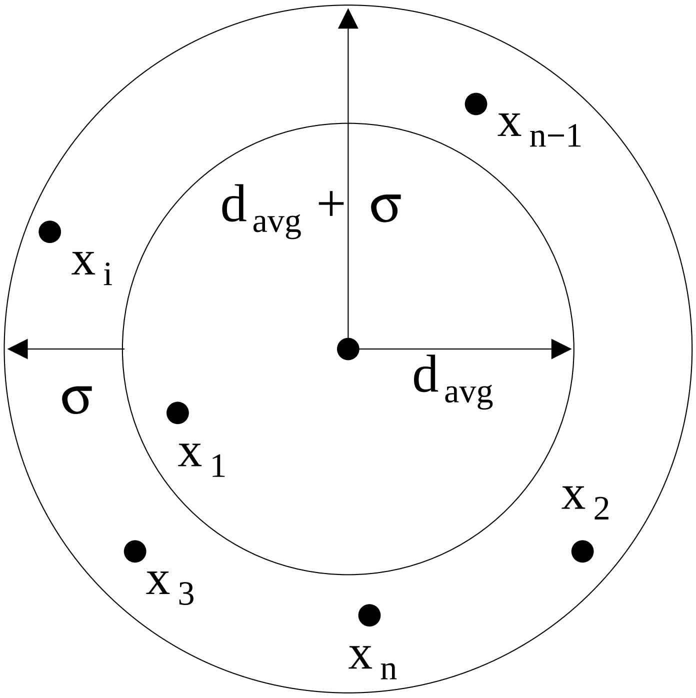

Figure 5: d_{avg} (average distance) and σ (standard deviation) derivation from distances

of these two parameters.

$$d_{avg} = average(d_1, d_2, ..., d_n) \tag{3}$$

$$\sigma = SD(d_1, d_2, ..., d_n) \tag{4}$$

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

$$threshold(t) = \sqrt{\sum_{i=1}^{6} (d_{avg,i} + \sigma_i)^2} \tag{5}$$

5 Experiments & Results

For experiment we took 30 original signatures from each person and selected 9 for training. These original signatures are taken in different days. Forgeries taken by three persons and 10 from each. Total 21 originals and 30 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.

5.1 Training

Let n signatures are taking for training from each person. There are 12 feature points from each original signature, 6 are taken by vertical splitting (Section2.2) and 6 are taken by horizontal splitting (Section2.3). Individual thresholds and patterns are calculating for vertical splitting and horizontal splitting. Pattern points based on vertical splitting are shown below.

$$\begin{aligned}
v_{pattern,1} &= median(v_{1,1}, v_{2,1}, \ldots, v_{n,1}) \\
v_{pattern,2} &= median(v_{1,2}, v_{2,2}, \ldots, v_{n,2}) \\
v_{pattern,3} &= median(v_{1,3}, v_{2,3}, \ldots, v_{n,3}) \\
v_{pattern,4} &= median(v_{1,4}, v_{2,4}, \ldots, v_{n,4}) \\
v_{pattern,5} &= median(v_{1,5}, v_{2,5}, \ldots, v_{n,5}) \\
v_{pattern,6} &= median(v_{1,6}, v_{2,6}, \ldots, v_{n,6})
\end{aligned} \tag{6}$$

Where $v_{i,1}, v_{i,2}, \ldots, v_{i,6}$ are vertical splitting features of i^{th} training signature sample. Threshold based on vertical splitting is shown below.

$$v_{threshold} = \sqrt{\sum_{i=1}^{6} (vd_{avg,i} + \sigma_{v,i})^2} \tag{7}$$

In equation 9 $vd_{avg,i}$ is same as average distance and $\sigma_{v,i}$ is same as standard deviation shown in Section 4. Pattern points based on horizontal splitting are shown below.

$$\begin{aligned}
h_{pattern,1} &= median(h_{1,1}, h_{2,1}, \ldots, h_{n,1}) \\
h_{pattern,2} &= median(h_{1,2}, h_{2,2}, \ldots, h_{n,2}) \\
h_{pattern,3} &= median(h_{1,3}, h_{2,3}, \ldots, h_{n,3}) \\
h_{pattern,4} &= median(h_{1,4}, h_{2,4}, \ldots, h_{n,4}) \\
h_{pattern,5} &= median(h_{1,5}, h_{2,5}, \ldots, h_{n,5}) \\
h_{pattern,6} &= median(h_{1,6}, h_{2,6}, \ldots, h_{n,6})
\end{aligned} \tag{8}$$

Where $h_{i,1}, h_{i,2}, \ldots, h_{i,6}$ are horizontal splitting features of i^{th} training signature sample. Threshold based on horizontal splitting is shown below.

$$h_{threshold} = \sqrt{\sum_{i=1}^{6} (hd_{avg,i} + \sigma_{h,i})^2} \tag{9}$$

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

5.2 Testing

When new signature comes for testing we have to calculate features of vertical splitting and horizontal splitting. Feature points based vertical splitting are $v_{new,1}$, $v_{new,2}$, $v_{new,3}$, $v_{new,4}$, $v_{new,5}$, $v_{new,6}$. Distances between new signature features and pattern feature points based on vertical splitting are shown below.

$$\begin{aligned} vd_{new,1} &= distance(v_{pattern,1}, v_{new,1}) \\ vd_{new,2} &= distance(v_{pattern,2}, v_{new,2}) \\ vd_{new,3} &= distance(v_{pattern,3}, v_{new,3}) \\ vd_{new,4} &= distance(v_{pattern,4}, v_{new,4}) \\ vd_{new,5} &= distance(v_{pattern,5}, v_{new,5}) \\ vd_{new,6} &= distance(v_{pattern,6}, v_{new,6}) \end{aligned} \tag{10}$$

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

$$v_{distance} = \sqrt{\sum_{i=1}^{6} vd_{new,i}^2} \tag{11}$$

Feature points based vertical splitting are $h_{new,1}$, $h_{new,2}$, $h_{new,3}$, $h_{new,4}$, $h_{new,5}$, $h_{new,6}$. Distances between new signature features and pattern feature points based on vertical splitting are shown below.

$$\begin{aligned} hd_{new,1} &= distance(h_{pattern,1}, h_{new,1}) \\ hd_{new,2} &= distance(h_{pattern,2}, h_{new,2}) \\ hd_{new,3} &= distance(h_{pattern,3}, h_{new,3}) \\ hd_{new,4} &= distance(h_{pattern,4}, h_{new,4}) \\ hd_{new,5} &= distance(h_{pattern,5}, h_{new,5}) \\ hd_{new,6} &= distance(h_{pattern,6}, h_{new,6}) \end{aligned} \tag{12}$$

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.

$$h_{distance} = \sqrt{\sum_{i=1}^{6} hd_{new,i}^2} \tag{13}$$

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

5.3 Results

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 equation 14 and FRR is calculated by equation 15.

$$FAR = \frac{number\ of\ forgeries\ accepted}{number\ of\ forgeries\ tested} \times 100 \tag{14}$$

$$FRR = \frac{number\ of\ originals\ rejected}{number\ of\ originals\ tested} \times 100 \tag{15}$$

Table 1 shows the *False Acceptance Rate* of our method for different types of forgeries. Table 2 shows the *False Rejection Rate* for original sigature.

Table 1: False Acceptance Rate (FAR)

Forgery Type	FAR(%)
Random Forgeries	2.08
Simple Forgeries	9.75
Skilled Forgeries	16.36

In general there are different thresholds for different types of forgery detections. But here threshold is same for random, simple and skilled forgeries. Because this method is mainly eliminating random and simple forgeries.

Table 2: False Rejection Rate (FRR)

Signature	FRR(%)
Original Signatures	14.58

6 Conclusion

This method performs much better than any other off-line signature verification methods. Future direction in this is classifying the skilled forgeries correctly. For this we have to approach novel classification method.

Figure 6: Feature points based on vertical splitting of depth 2

Figure 7: Feature points based on horizontal splitting of depth 2

For better classification we can again split the sub-parts of Fig.3 using vertical splitting and Fig.4 using horizontal splitting. Then instead of six featire points we can get 24 feature points for each vertical and horizontal splittings. Fig.6 shows the vertical splitting of depth 2. Fig.7 shows the horizontal splitting of depth 2.

References

[1] 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.

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

[3] 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.

[4] 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.

[5] 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.

[6] 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", *Patter Recognition* 36, 2003, pp. 91–101.

[7] Migual A. Ferrer, Jesus B. Alonso and Carlos M. Travieso, "Off-line Geometric Parameters for Automatic Signature Verification Using Fixed-Point Arithmetic", *IEEE Tran. on Pattern Analysis and Machine Intelligence*, vol.27, no.6, June 2005.

[8] 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.

[9] 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.

Banshider Majhi, Y Santhosh Reddy, D Prasanna Babu
Department of CSEA
NIT Rourkela
India 769008
E-mail: ysantosh@rediffmail.com