Matching between Important Points using Dynamic Time Warping for Online Signature Verification

Mitra Hamedanchian Mohammadi, Karim Faez

***Abstract*— Online signature verification is one of the biometric features which can be used as a common method for identity verification. According to the previous studies, calculation of similarities between the signatures with an extended regression approach, as compared with the Euclidean distance and Dynamic Time Warping (DTW), gives a better measure of the similarity of two signatures. For this purpose the time length of the signals corresponding to the two signatures should be the same.**

Using of all point matching strategy to unify the time length of the signals reduce the distinction level between genuine and forged signatures. Therefore, in this study, to maximize the distinction between the genuine and forged signatures, a method based on the correspondence between important points in the direction of warp for the time signal provided, extended regression is used to calculate the similarity index of the two signatures. This system was tested on a set of signatures of the First International Signature Verification Competition (SVC2004) database. With this method the verification error of 6.33% is obtained with professional forged signatures. But there was no failure for random forged signatures.

***Index Terms*— Extended regression, Important points, DTW, Online signature, Professional forged signature, Verification error.**

I. Introduction

In a society in which a lot of information are electronically stored and delivered, on-line and real time verification of identity is essential [1]. To confirm the identity of individuals, the physical characteristics such as fingerprints, face, palm, iris or behavioral characteristics such as speech and signature can be used. Although the signature of a person has changed over the time and the possibility of its forgery is easier than the forgery of fingerprints and iris of the eye, given that the signature in different countries and cultures is more usual and accepted as a sign of identity, and also can be processed with more speed. Thus, use of signature to verify and confirm the identity of individuals has been taken more into consideration [2].

Signature verification systems are divided into two groups: on-line and off-line signature verification. In offline signature verification systems, there is only information about the shape and spatial features of the signatures, whereas in on-line signature verification systems, in addition to the signature shape information, time information of the signature path is also used [3]. Given that it is much harder to forge the dynamic characteristics of a signature, as compared to its spatial characteristics; thus on-line signature verification systems carry relatively lower errors. On-line signature verification methods and procedures are divided into two groups of parametric and functional methods. The parametric methods use general characteristics and features for signature representation and the features of a reference signature are compared with the features of test signature and the final decision concerning the original or forged signatures is taken. However, in the functional method, a signature pattern is defined as functions of time and its features are compared locally. These comparisons can be performed point by point or segment by segment. Functional signature verification methods convey lower verification error and a great deal of studies have been conducted on these methods, some of which are reviewed in the following sentences.

In a study conducted by Ali-Zadeh et al [4], they developed a new method for signature verification using parametric features based on selection of an optimal threshold. After preprocessing, for each signature, 62 parametric features are derived from horizontal place, x(t), vertical place, y(t) and pen down and up signals which are obtained from a digitizer plane. After extracting the genuine signature characteristics from the training set, the mean and variance for each of the signers are calculated and then stored as a reference feature. The weighted Euclidean distance between each feature of a signer and the mean feature of the reference signature is compared with an appropriate threshold and then the feature will be verified or rejected. Finally, the number of verified features listed for an individual is compared with another threshold, which is different for each signature and has a proper value. At the final step, the signature will be verified or rejected.

Yoon et al [5] used geometric extremum points for

Manuscript received January 10, 2012.

M. Hamedanchian Mohammadi; Electrical, Computer Engineering and Information Technology Department, Islamic Azad University Qazvin Branch, Qazvin, Iran (Corresponding author; e-mail: Mitra.Hamedanchian@gmail.com.)

K. Faez; Electrical Engineering Department, Amirkabir University of Technology, Tehran, Iran (e-mail: Kfaez@aut.ac.ir).

segmentation of signatures and with changing the standard algorithm of Dynamic Programming (DP), obtained the matching pattern between segments and finally extracted several features for each segment. Then, by use of neural networks, he obtained the similarity between two signatures. In this system, 5 original signatures are used to generate a sample signature. This method was tested on a signature set consisted of 6790 original signatures from 271 persons and for the presence of random forged signatures. The EER was obtained as 94/1% for this system.

Nakanashi [6] used the functions of the signature path and angle of the pen moving on the surface to verify signatures and similarities between these functions in different bands of details features and approximation by adaptive signal processing and then a combination of them was used for signature verification. He tested this method on 200 genuine signatures and 200 forged signatures from 4 persons and reported the EER of 3.5% for this verification method.

Lee [7] developed an extended regression analysis to calculate the similarity between two signatures. Considering the fact that time length of signals related to two signatures are different, he used the matching of all points for unifying the time length of signals for the two signature that reduces the distinction between the genuine and forged signatures.

In this paper, to increase the distinction between genuine and forged the signature, the important points approach was used to unifying the time length of signals and the extended regression was used to calculate the similarity between signatures.

This paper consists of 8 sections of which, in section 2 the applied preprocessing procedures are explained. In section 3, feature extraction, and in section 4, calculation of the similarity between two signatures is discussed. Sections 5 and 6 are devoted to the training of the verification algorithm and decision, respectively. In section 7, the verification system is evaluated and finally section 8, is devoted to conclusions.

II. PREPROCESSING

In preprocessing stage, the size of signatures is normalized and the rotation angle is removed.

A. Normalization of signature size

In the signature verification system, if the digitalizing pages, which users use them, have different sizes, then the sizes of signatures must be normalized, since, a person changes the size of his/her signature proportionally to the available space for them. Difference in the size of signature causes different problems for comparison of the signatures. To resolve these problems, the X and Y signals become normalized using the (1).

$$x(n) = \frac{x^*(n) - m(x)}{x^*_{\max} - x^*_{\min}} \qquad (1)$$

$$y(n) = \frac{y^*(n) - m(y)}{y^*_{\max} - y^*_{\min}}$$

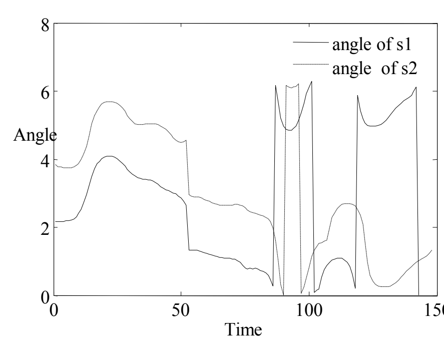
In (1), $x^*(n)$، $y^*(n)$ are the coordinates of the pen at the time n and $x(n)$، $y(n)$ denote the normalized values. $x^*_{\max}$ ، $x^*_{\min}$ and $m(x)$, denotes the maximum, minimum and average of signal X, respectively. And $y^*_{\max}$ ، $y^*_{\min}$, and $m(y)$, shows the maximum, minimum, and average of the signal Y, respectively.

B. Removing the rotation angle using DTW

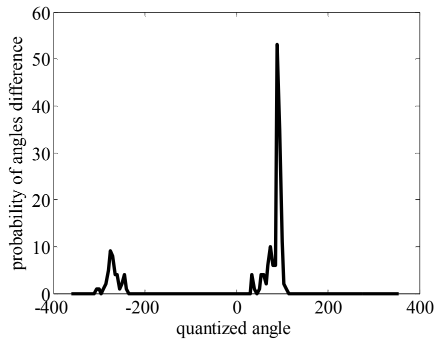
In this paper, prior to the feature extraction and calculation of the similarity between two signatures, the rotation angle between them was identified and then removed. This is done to reduce the variation within the class.

Transforming the signature into polar coordinates, signals of r and θ describe a signature. If we rotate a signature in polar coordinates by angle α, then the function of r would remain unchanged and only function of θ will shift by α or $2\pi - \alpha$.

We use this property of polar coordinates to eliminate the rotation angle. Due to a lack of time one to one correspondence between the two signatures, their point by point comparison is not possible. Therefore, firstly, we find correspondence between “r functions” of two signatures with the use of DTW algorithm, Then we use the obtained correspondence path to unify the θ signals time length. How to find match and unifying the signals length of time in section (4) will be discussed in details. Following the unifying the time length of the signals θ for two signatures, their difference can be used to determine the rotation angle. To do this, we found the difference frequency of signals θ of the two signatures in the intervals of 5º. The angle by which frequency diagram is maximum, defines the rotation angle. The reason for this choice of 5º intervals is that the maximum error of the determined rotation angles is 2.5º and signature verification system is not sensitive to this error rate. In the Fig. 1. (a) and Fig. 1.(b), signals θ of two signatures and their corresponding difference frequency are shown. These signatures are rotated 90º relative to each other.



(a)



(b)

Fig. 1. (a) Signals θ of two signatures with 90° rotation with respect to each other, (b) Frequency of signal differences for two signals θ

To remove the rotation, we rotate the signature S2 by $-\alpha$ using (2).

$$\begin{bmatrix} x(n) \\ y(n) \end{bmatrix} = \begin{bmatrix} \cos(-\alpha) & -\sin(-\alpha) \\ \sin(-\alpha) & \cos(-\alpha) \end{bmatrix} \begin{bmatrix} x^*(n) \\ y^*(n) \end{bmatrix} \tag{2}$$

III. Feature Extraction

In this paper, the signals of x, y, v_x and v_y can be used as functional features. The digital signals of x and y are registered directly with the digitalizing page and used after the preprocessing stage as the functional features. The functions of v_x and v_y can be determined using (3).

$$V_x(n) = \sum_{\tau=1}^{\tau=3} \frac{x(n+\tau) - x(n-\tau)}{\tau} \tag{3}$$

$$V_y(n) = \sum_{\tau=1}^{\tau=3} \frac{y(n+\tau) - y(n-\tau)}{\tau}$$

IV. Calculation of Similarity Between Two Signatures

Used to calculate the similarity between the signatures we used an extended regression. Extended regression directly determines the similarity levels of two multi-dimensional sequences, but it can only calculate the similarities between the signals with the same time length. If the signals have different lengths of time, their time lengths should be the unified to the corresponding points of them can be matched on each other [7]. DTW algorithm is used to find the corresponding points of two signals [8].

A. Finding corresponding points of two signals with DTW

Suppose we have two signals, of which the length of signal A equals n and the length of signal B equals m. To find corresponding points of the two signals by use of an algorithm based on Dynamic Time Warping (DTW) method; first, a $n \times m$ matrix is formed which its (i, j) element is determined by (4):

$$d(i,j) = \sqrt{\sum_{k=1}^{n} (A_{k,i} - B_{k,j})^2} \tag{4}$$

To find the corresponding points of two signals, we find a path on which the sum of the elements of matrix d from the (1,1) element to (m, n) element is minimum. Such a path with the above condition is called warping path. Warping path W is an integrated set of elements of the matrix d and represents the details of the match between the signals A and B. The K-th element of the W path is presented by (5).

$$\mathrm{w_k} = (\mathrm{i(k), j(k)}) \quad \max(m,n) < k < m+n-1 \tag{5}$$

Optimal path W, is the path that minimizes (6).

$$\mathrm{DTW}(\mathrm{X_1, X_2}) = \min \sqrt{\sum_{t=1}^{K} \mathrm{d(w_t)}} \tag{6}$$

From the algorithm DP, the (7) is used to find the corresponding points of two signals.

$$\gamma(\mathrm{i,j}) = \mathrm{d}(\mathrm{X_{1,i}, X_{2,j}}) + \min \begin{cases} \gamma(\mathrm{i}-1, \mathrm{j}-1) \\ \gamma(\mathrm{i}-1, \mathrm{j}) \\ \gamma(\mathrm{i}, \mathrm{j}-1) \end{cases} \tag{7}$$

Where, $\gamma(n,m)$ shows the total distance between the two signals [8]. Using DTW, the matching path between the two signals is obtained and by use of the matching path, the time length of the signals will be the same. Fig. 2 shows the matching path of the signals A and B.

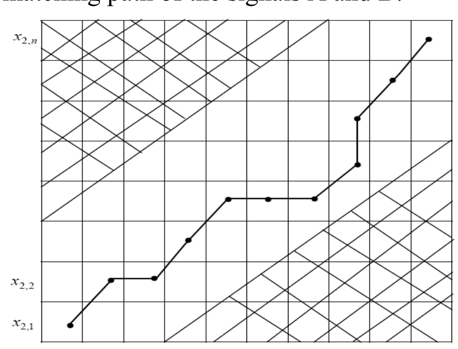


Fig. 2. Matching path of the signals A and B

B. The proposed method for unifying the time length of signals

Lee for unifying the time length of signals proposed the following procedure:

If $x_{1,i}$ in signal A matches k > 1 points of the signal B, then it will repeat k-1 times and if $x_{2,i}$ matches k points of the signal A, then it will be extended by the same method [7].

With this algorithm, two signals with the same time length are obtained which their corresponding points are matched to each others. However, this algorithm reduces the distinction between the genuine and forged signatures. In this paper, to keep the distinction between original and forged signatures, a method is developed based on important points matching for unifying the time length of the signals.

This is done in three steps; in the first step, from the first signature, zero-crossing points of the velocity signals X and Y are derived as important points of the signature. In the second step, on the matching path, the corresponding points with the important points in the second signal are determined. Finally, in the third step, by use of the important points matching, the time length of the signals will be unified.

First step: extraction of important points

To find the important points in the velocity signals of the first signature, their zero-crossing points will be defined according to the following (8). In other words, the zero-crossing points of the velocity signal along the direction of X and Y are considered as the important points.

$$V_x(n)\times V_x(n-1)<0 \quad OR \quad V_y(n)\times V_y(n-1)<0 \tag{8}$$

The n-th important point from the signal A is presented by the element $e(1,n)$.

Second step: finding the optimal matching for the corresponding points with important points in the next signatures

Considering the fact that we are going to use the matching of important points to unify the time length of the two signals, a wrong matching can impose critical problems for the comparison of the two signals. Therefore, finding the optimal matching between important points in the first signature with their corresponding points in the second signature. To find the matching of important points, we adapt the following procedure:

With the algorithm DTW, we find the matching path of the signals A and B and represent them with X1 and X2, respectively. The important points found in the signal A are determined on the matching path of the X1, and then their corresponding points on the path X2 in the signal B are found. It is possible to map several points from the signal B on the one important point of the signal A and in this condition, the mean of the mapped points on that important point should be consider as its corresponding point. It is shown in the following example.

For further explanation, the performance of this method is presented with a numerical example. The occurrence time of the important points of the signal A is presented in Table I and Table II shows the corresponding point of an important point on a section of the matching path obtained by DTW method.

TABLE I
Occurrence time of the important points in signal A

Important Points (A)	10	17	24	33	41	45

TABLE II
Representation of the corresponding point in the signal B1 on a section of the matching of the time axis of signals A and B

A1	7	8	9	10	**10**	10	11	12	13
B1	7	8	9	10	**11**	12	13	14	15

Third Step: Unifying the time length of signals

For the unifying the time length of two signals, we place the important points, and the points corresponding with important points inside the array C with $2\times N$ dimension in which N is equal to the number of important points. The pair of $C(1,n)$ and $C(2,n)$ show the n-th matching. In addition to the matching of the important points with the corresponding points, we match the first and last points of two signals. Suppose the time length of the signal A is equal to p and for the signal B is q. To match the two signals the time axis of the signal A is kept constant and just change the time axis of the signal B according to (9).

$$\text{If } 0<n<C(2,1) \tag{9}$$

$$n' = 1+\left[\frac{C(1,1)}{C(2,1)}\times(n-1)\right]$$

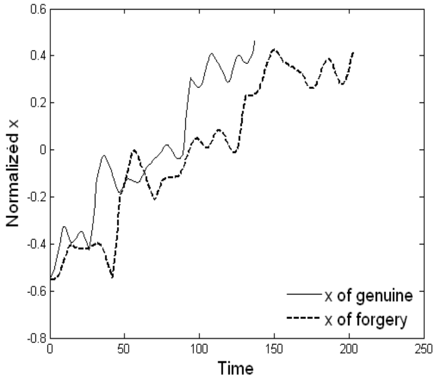
$$\text{If } C(2,m)\le n<C(2,m+1)$$

$$n' = C(1,m)+\left[\frac{C(1,m+1)-C(1,m)}{C(2,m+1)-C(2,m)}\times\left(n-C(1,m)\right)\right]$$

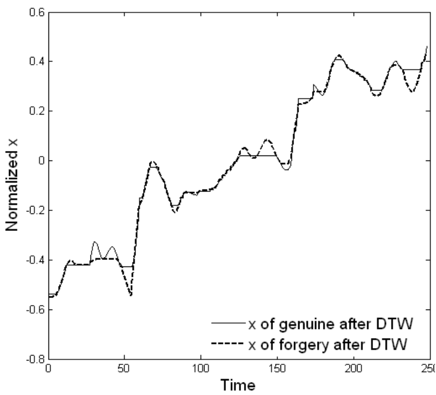
$$\text{And if } C(2,N)\le n\le q$$

$$n' = C(1,N)+\left[\frac{p-C(1,N)}{q-C(2,N)}\times\left(n-C(2,N)\right)\right]$$

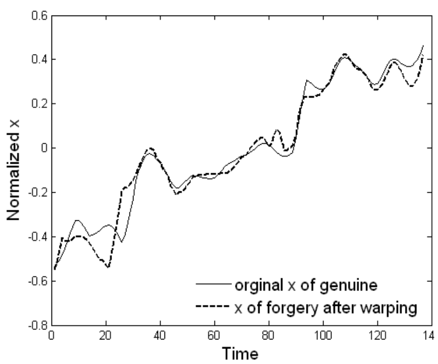
In the Fig. 3. (a), signals X of one genuine and one forged signature and in Fig. 3. (b) and Fig. 3. (c), the results of the unifying of the time length of the signals with DTW method [7] and the proposed method are shown.



(a)



(b)



(c)

Fig. 3. (a) Signals x of the genuine and forged signatures, (b) Unifying the time length with DTW method, (c) Unifying the time length with the proposed method

Fig. 3. (b) and Fig. 3. (c) show that the proposed method, in comparison with the method described by reference [7], increases the distinction level of the genuine and forged signatures.

C. *Similarity calculation*

Following the unifying procedure of the time length of the signals, corresponding to the two signatures, we place the values of the x، y، v_x and v_y of one signature in the matrix g and the corresponding signals of the another signature in the matrix F and then we normalized these matrices by use of (10).

$$G_{ji} = \frac{g_{ji}}{\sqrt{\sum_{i=1}^{N}(g_{ji})^2}} \qquad F_{ji} = \frac{f_{ji}}{\sqrt{\sum_{i=1}^{N}(f_{ji})^2}} \tag{10}$$

The matrices of G and F have the dimension of $4 \times n$. With the (11), the similarities between the two signatures are determined.

$$similarity = \frac{[\sum_{j=1}^{M}(\sum_{i=1}^{N}(G_{ji}-\overline{G_j})(F_{ji}-\overline{F_j}))]^2}{\sum_{j=1}^{M}\sum_{i=1}^{N}(G_{ji}-\overline{G_j})^2\sum_{j=1}^{M}\sum_{i=1}^{n}(F_{ji}-\overline{F_j})^2} \tag{11}$$

In (11), M=4, and $\overline{G_j}$ and $\overline{F_j}$ represent the mean of the J-th dimension for the sequences of G and F. N denotes the time length of the recorded signature in the matrix G.

V. Signature Verification System Training

In some methods of signature verification in the stage of training of the system, a model for signature is derived [5], but in this paper, the purpose of the training is determination of the decision boundaries. To do this procedure, 5 genuine signatures of a person are used and the similarities between these signatures are calculated two by two and the average of ten obtained similarities is used for determination of the decision boundaries.

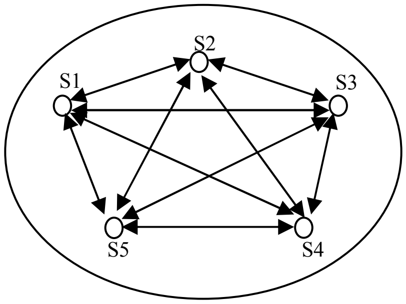
Decision boundary related to the signatures of the i-th person is determined by the (12).

$$T_i = \alpha \times ms\_i \qquad ms\_i = \frac{\sum similarity}{10} \tag{12}$$

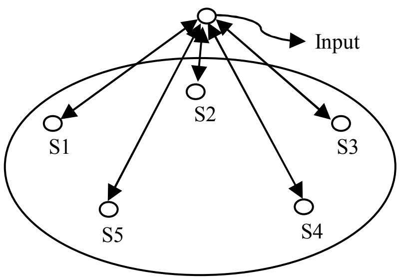
In (12), T_i is a decision boundary for the signatures of an i-th person and α is an experimental coefficient which is determined from the error graphs, based on the required security level.

VI. Decision Making

In the verification stage, similarity of the input with each of 5 training-step signatures is calculated and the mean value of the 3 first higher similarities, are considered as the similarity index of the input signature with the training stage signatures. This index is determined experimentally. Similarities of an input signature with training samples assigned to the i-th person are presented with score i. Procedure of signature comparisons in the step of training and decision is presented in Fig. 4. (a) and Fig. 4. (b), respectively.



(a)



(b)

Fig. 4. (a) Training of the verification System,
(b) Decision making

To confirm or reject input signature which is claimed to be belonged to the i-th person, if the condition of score–i > T_i is fulfilled, then the input signature will be verified, otherwise, it will be rejected.

VII. EXPERIMENTAL RESULTS

Evaluation of signature verification system needs a signature dataset. We used the signatures of the First International Signature Verification Competition (SVC2004) dataset which is consisted of 1600 signatures from 40 persons. For each person, 20 genuine signatures and 20 forged signatures are collected. This set was available to the public in the internet website of the competition [9]. Five genuine signatures are employed for the training of the verification system, and 10 genuine signatures and 20 forged signatures (according to the competition conditions) are used to evaluate the signature verification system.

For evaluation of the signature verification system, two types of error are defined: First type of error is represented with FAR. It shows the verification rate of forged signature and the second type of error is represented by FRR and it shows the rejection rate of the genuine signatures. These two types of errors are proportional to each other inversely; i.e. if the decision boundary changes somehow to FAR decreases, then the value of FRR will increase. Therefore, the charts of the FAR and FRR are drawn versus the variation of the decision boundary and the intersectional point of two graphs, which is called EER, is used for the evaluation of the signature verification system. Fig. 5 shows false acceptance rate and false rejection rate with different values of α for the proposed system, when applying our method on SVC2004 dataset.

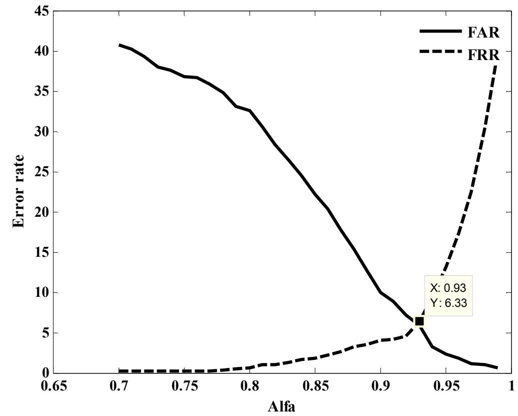


Fig. 5. False acceptance rate (solid line) and False rejection rate (dotted line) for proposed system.

Based on the required security level, α can be changed. For example, when the rate of payment is getting more in the electronic payments, verification of forgery signature makes a lot of problems for bank. Therefore, FAR is more important than FRR. We recommend that for these applications, α is set larger than 0.93 that leads to smaller value of FAR in the charge of increasing the value of FRR. For example, $\alpha = 095$ results in FAR=2% and FRR=13.09%.

A. Comparison with some other methods

The major difference between the reference [7] and the proposed algorithm is the unification of the time length of the signals. These two signature verification systems were tested on the signature set of the SVC2004 dataset. The result of this experiment and some other related work that reported their results on SVC2004 are presented in Table III

TABLE III
Error rates for comparison with some other methods

Signature verification system	EER(%)
The proposed algorithm	6.33
Reference [7]	14.21
Best SVC2004 [9]	5.50
Reference [10]	7
Reference [11]	10.63

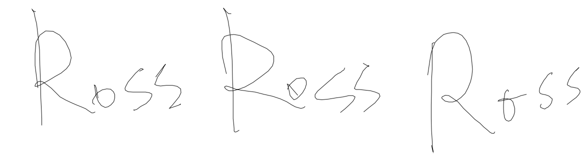
The results show that the proposed method for the unifying of the time length of the signals reduces the error rate by 55%, as compared to the method described by the reference [7] and it has a superior performance compared to other methods. Also the proposed signature verification system for skilled forgery signatures is ranked 2nd in comparison with the other teams participating in the first international signature verification competition. Furthermore, this method for the verification of the random forgery signatures yields no verification error and it will be ranked first.

B. Discussion

For signature verification systems, some of the signatures are verified with error. We found some reasons reported in the following.

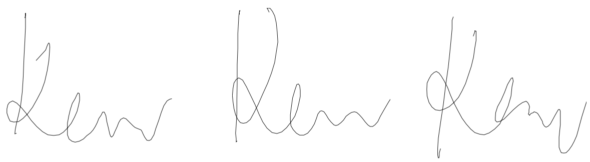
It is easy to forge some signatures that are the name of signers or signatures that have simple shapes as shown in Fig.6 and Fig. 7, respectively.

Some complex signatures have high intra-class changes (Fig. 8). Therefore some genuine signatures are rejected, hence the FRR value of the verification system are increased.



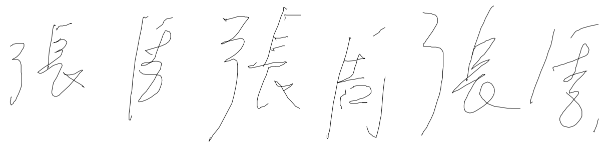
(a) (b) (c)

Fig. 6. signature that is the person's name
(a) Training signature,
(b) genuine signature was rejected,
(c) forged signature was verified



(a) (b) (c)

Fig. 7. a sample of the signatures that are simple
(a) Training signature,
(b) genuine signature was rejected,
(c) forged signature was verified



(a) (b) (c)

Fig. 8. The complex signature with the high intra-class changes
(a) Training signature,
(b) genuine signature was rejected,
(c) forged signature was verified

We recommend using two signatures for verification of these types of signatures. For simple signatures, both of them must be verified where for complex one, verification of one signature is enough.

VIII. CONCLUSIONS

Distinction increase between genuine and forged signatures, improves the efficacy of the signature verification system. In this paper, by use of important points matching for unification of time length of signals and the calculation of the similarity level by the application of an extended regression, the distinction level between the genuine and forged signatures is increased significantly. By use of this developed method, the EER percentage for the signatures of the SVC2004 and professional forged signatures, were obtained 6.33%, while the value of the EER for the all points matching method for unification of the time length of the signals and also calculation of the similarity, was reported 14.21%. The proposed signature verification system for skilled forgery signatures is ranked 2nd in comparison with the other teams participating in the first international signature verification competition. Furthermore, this method for the verification of the random forgery signatures yields no verification error and it will be ranked first.

REFERENCES

[1] G.Gupta, R.Joyce, "Using position extrema points to capture shape in on-line handwritten signature verification" Pattern Recognition, vol.40, pp. 2811 – 2817, 2007

[2] A.Kholmatov, B.Yanikoglu, "Identity authentication using improved online Signature verification method" Pattern Recognition, vol.26, pp.2400–2408, 2005

[3] M. Talal Ibrahim, M. Aurangzeb Khan, K. Saleem Alimgeer, M. Khalid khan, I.A.Taj, L.Guan "Velocity and Pressure-based partitions of horizontal and veritical trajectories on-line signature verification" Pattern Recognition, vol.43, pp.2817–2832, 2010

[4] A. Alizadeh, T.Alizadeh, Z.Daei, "Optimal Threshold Selection for Online Verification of Signature" IMECS international Conference of Artificial Intelligence and Applications, Hong-Kong, vol.1, pp.98-102, March 2010

[5] J.Lee, H.Yoon, J.Sob, B.Chon, Y.K.Chong, "Using geometric extrema for segment-to-segment characteristics comparison in online signature verification" Pattern Recognition, vol.37, pp.93–103, 2004

[6] I.Nakanishi, H.Sakamoto, Y.Itoh, Y.Fukui, "Multi-matcher online signature verification system in DWT domain" Proc. Int. Conf. on Acoustics, Speech, and Signal Processing, vol.2, pp. 965–968, March 2005

[7] H.Lei, S.Palla, V.Govindaraju, "ER2: An intuitive similarity measure for on-line signature verification" IWFHR '04: Proceedings of 9th International Workshop on Frontiers in Handwriting Recognition (IWFHR'04), IEEE Computer Society, pp.191-195, Tokyo, October 2004

[8] M. Zoghi, V. Abolghasemi, "Persian Signature Verification Using Improved Dynamic Time Warping-based Segmentation and Multivariate Autoregressive Modeling" Proc. IEEE, 15th workshop of Statistical Signal Processing (SSP'09), pp.329-332, 2009.

[9] The First International Signature Verification Competition, 2004, (SVC2004), available in http://www.cs.ust.hk/SVC2004.

[10] M. Adamski, Kh. Saeed, "Online signature classification and its verification system", Proc. IEEE, 7-th Computer Information Systems and Industrial Management Applications (CISIM'08), pp.189-194, June 2008.

[11] A. Flores-Mendez, M. Bernal-Urbina, "Dynamic Signature Verification through the Longest Common Subsequence Problem and Genetic Algorithms", Proc. IEEE, Evolutionary Computation Conferences (CEC), Barcelona, pp. 1-6, July 2010.