Online Signature Verification: a Robust Approach for Persian Signatures

Mohammad Esmaeel Yahyatabar
Department of Electrical and Computer Engineering, Babol University of Technology, Babol, Iran
me.yahyatabar@stu.nit.ac.ir
Yasser Baleghi*
Department of Electrical and Computer Engineering, Babol University of Technology, Babol, Iran
y.baleghi@nit.ac.ir
Mohammad Reza Karami
Department of Electrical and Computer Engineering, Babol University of Technology, Babol, Iran
mkarami@nit.ac.ir

Received: 22/Jun/2014 Revised: 24/Dec/2014 Accepted: 17/Jan/2015

Abstract

In this paper, the specific trait of Persian signatures is applied to signature verification. Efficient features, which can discriminate among Persian signatures, are investigated in this approach. Persian signatures, in comparison with other languages signatures, have more curvature and end in a specific style. An experiment has been designed to determine the function indicating the most robust features of Persian signatures. To improve the performance of verification, a combination of shape based and dynamic extracted features is applied to Persian signature verification. To classify these signatures, Support Vector Machine (SVM) is applied. The proposed method is examined on two common Persian datasets, the new proposed Persian dataset in this paper (Noshirvani Dynamic Signature Dataset) and an international dataset (SVC2004). For three Persian datasets EER value are equal to 3, 3.93, 4.79, while for SVC2004 the EER value is 4.43.

These experiments led to identification of new features combinations that are more robust. The results show the overperformance of these features among all of the previous works on the Persian signature databases; however, it does not reach the best reported results in an international database. This can be deduced that language specific approaches may show better results.

Keywords: Online Signature Verification; Support Vector Machine; Robust Feature Extraction; Online Signature Dataset.

1. Introduction

Nowadays, biometric methods are more considered for identification. These methods strongly depend on inherent characters of people, thus they are highly reliable. A specific signature such as other biometric features is exactly associated to a specific person and this unique feature is applied to identification and verification. IBG1[1] reported that the signature modality is the second behavioral trait in commercial importance just after voice biometrics. Applications of online signature verification in legal (document authentication), medical (record protection), and banking sectors (cheque and credit card processing) are so common and increasing [2].

Signature verification consists of two types including static and dynamic verification. Shapes of signatures are available in static signature verification, i.e. recognition has to be done on a two-dimensional shape and the final decision is based on signature appearance. However, dynamic features are considered as well as appearance features in dynamic verification. In this method, Pressure sensitive tablet records the 2D coordination, pressure, Azimuth and Altitude of signatures in specific intervals of time.

Generally, feature extraction in signature verification is categorized into parametric and functional types [3]. In functional type, time sequences describing local properties of the signature are used for recognition[2],whereas other features such as global and shape-based features are included in parametric category.

1.1 Parametric Features

Many researchers have worked with different methods on parametric features. Pippin[4] used global features such as average pressure, average velocity, pen tip and number of curves. Curves in signatures are extracted and compared with Dynamic Time Warping (DTW) to find similarity between reference signature and input signature. With specifying a threshold value, similarity is analyzed and decision-making is done. Dehghani [5] presents a two-phases method for Persian signatures. The first phase consists of feature extraction based on fractal vector of signature pressure. Then the Adaptive Network based Fuzzy Inference System (ANFIS) is applied to classify Persian signatures, which passed first phase. Alizadeh [6] extracts 62 parametric features such as total signing duration, signature height, maximum of x and y and associated time. Extracted features were compared with two threshold values in two-stages in classification.

[1] International Biometric Group

* Corresponding Author

1.2 Functional Features

As mentioned before, some functions can be achieved from the tablet data package. Nanni [7] implements discrete wavelet transform on functions extracted from signature and decreases its dimensions using Discrete Cosine Transform (DCT). Gained vectors are considered as classifier inputs. One-dimensional signals of X(t) and Y(t) have been processed in parallel in [8]. The signature has been pre-processed by Mellin transform being scale invariant. Feature vectors have been extracted using Mel Frequency Cepstral Coefficient (MFCC) and feature dimensions have been decreased using Principal Component Analysis (PCA). Finally, decision is made using linear classifier. Muhammad Khan [9]proves that parts including middle velocity can be suitable criteria for proper verification. Middle velocity is categorized into low middle velocity and high middle velocity. Classifier inputs are considered using middle velocity features.

Marianela et al. [10,11] studied the discriminative power of combinations of most commonly time functions related to signing process. A consistency factor is defined to quantify the discriminative power of these different feature combinations. They presented experimental results that show there is a good correlation between the consistency factor and the verification errors, suggesting that consistency values could be used to select the optimal feature combination.

In [12], a new partitioning method is proposed for online signature verification. The partitions represent areas of high and low speed of signature and high and low pen's pressure. The method is performed on SVC2004 and BioSecure databases.

The One-Class Support Vector Machine (OC-SVM) based on independent parameters is used in [13]. This method is proposed for the situation when the forgery signatures are lack as counterexamples. In order to reduce the misclassification, a modification of decision function used in the OC-SVM is suggested.

In [14] a minimum distance alignment between the two signatures is made using dynamic time warping technique that provides a segment to segment correspondence. Fuzzy modelling of the extracted features is carried out in the next step. The experiments are carried out on SVC2004 and SUSIG databases.

The rest of the paper is organized as follows. Datasets are briefly described in the next section. In section 3, proposed method is presented. The results are described in section 4. Section 5 discusses about the proposed method and the conclusion and future works are given in the last section.

2. Datasets

To analyze the proposed method for Persian signatures, two datasets are used from [5] and [15] and a new own dataset named Noshirvani Dynamic Signature Dataset (NDSD) is produced and will be published. The method that is used for gathering data is the same as what is applied in SVC2004 [16] international dataset.

Datasets from [5] are the first used datasets in which signatures are generated from 40 people as well as 10 signatures per person are involved. The second dataset[15], the next used dataset in this paper, involves dynamic data that is generated from 50 signatures. Each person has registered 25 signatures and there are 15 forgery signatures per sample.

Specific features of Persian signatures are tested to identify and analyze in this study. This method has been applied for international dataset called SVC2004. The mentioned dataset includes 1600 signatures generated from 40 people in which 20 forgery and 20 genuine signatures are involved per person[16].

2.1 Data Acquisition (NDSD)

The signatures of NDSD database -newly introduced in this work- in are acquired by WACOM INTUOS4 digitizing tablet. The tablet sent a data package including pen tip coordination, pressure, azimuth and altitude angles (see figure 1). Data are sent every 10 milliseconds and signers sign in a plate with size of 129×96 mm. The pen senses 2048 levels of pressure. The interface software was programmed by visual basic software and the tablet was connected to computer with USB port.

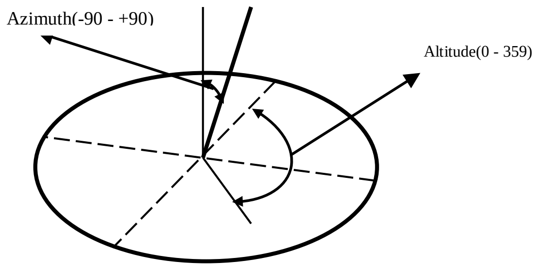

Fig. 1. Azimuth and Altitude angles.

NDSD dataset was prepared in Digital Signal Processing Laboratory of Babol Noshirvani University of technology. 55 students of computer and electrical engineering department participated in producing the dataset. Each person signed 65 signatures in two different times with more than 3 days interval. People signed in two situations of standing and sitting. The dataset users were in range of 18 to 40 years old. Seven signers of them were left handed and 23 signers were female.

Fifteen professional forgers forged all signatures. Two types of professional forging were performed. In the first type, the signers could see just the shape of the genuine signatures of people and had enough time to practice for forging. In this type of forging, forger person should guess the signature path and other dynamic features. In the second type, Forgers had dynamic information of signatures and tried to forge shapes and dynamic information of signatures. The signatures path and pen tip velocity were animated for forgers and the signatures intensities were proportional with the pen pressure on the tablet. In both types, forgers tried at least 15 times to forge signatures before recording the dataset signatures. Ultimately, forty forgery signatures were recorded from four forgers assigned for each genuine signature. In figure 2 some genuine and forgery signatures from NDSD dataset are illustrated.

NDSD dataset are illustrated.

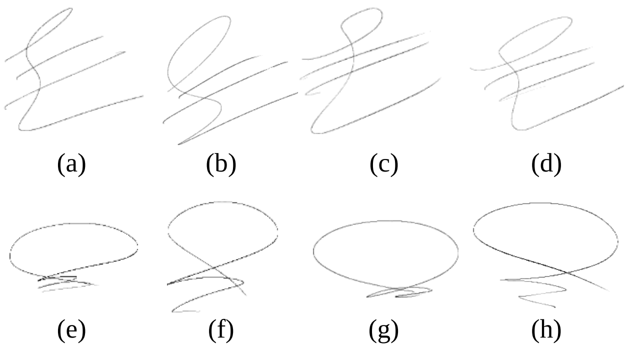

Fig. 2. figures 2.(a,b) , 2.(e,f) are genuine signature and 2.(c,d) , 2.(g,h) are their forgery signatures respectively.

3. Proposed Method

The proposed biometric verification system in this paper can be sketched as in figure 3. Signatures acquisition from the set, extracting features and making decision are three main stages of the method.

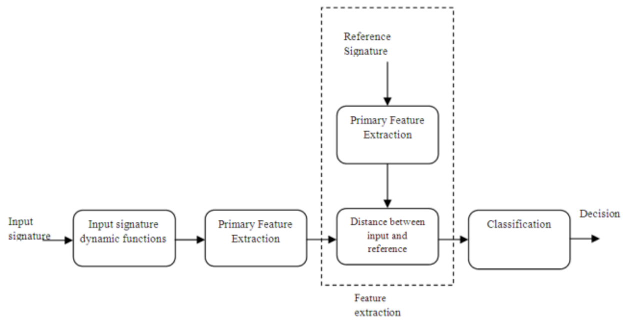

Fig. 3. Block Diagram of the proposed method.

In the proposed method, at first, specified features of Persian signatures are considered. The stability experiment of dynamic features has been carried out on Persian available signatures. Robust and reliable features of all Persian signatures are recognized. Besides, the results obtained from the experiment as well as two other features are the base of primary feature extraction. SVM classifier inputs are the distance between the input signature primary features and the reference signature features. The classification will be improved due to assigning small numbers to all genuine signatures and large numbers to forgeries through this process.

In the following sections, the proposed algorithm is described in detail:

3.1 Persian Signatures

Persian signatures are significantly different from other language signatures. In other languages, the shape of signatures is close to the names whereas Persian signatures are made by some lines, curves, and signs and almost different from the people's names. Some of the most important features in signatures of some languages such as Persian language can be pointed as following features:

- Using more curves in signatures
- More discrete lines than other languages
- Distribution on length and width (against some languages that are on a straight line)

Figure 3 (a-d) illustrates some Persian signatures from NDSD dataset. Some international signatures from SVC2004 are shown in figure 4 (e-h). The different characteristic mentioned above can be seen in the figure.

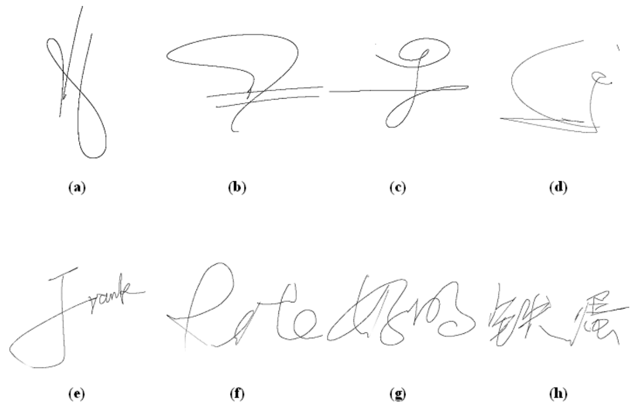

Fig. 4. Signatures 4.a to 4.d are Persian signatures from NDSD Persian dataset and in 4.e to 4.h includes some samples of SVC2004 international dataset.

Figure 4 illustrates more curves in Persian signatures that are mentioned before. While signing, the velocity of pen has special state and usually the velocities in these curves are more than the others. In addition, it is deemed that a Persian signer is moving his/her wrist and fingers more than other languages. This additional motion may lead to more discrepancy in various iterations. However, considering the smooth motion of hand, dynamic features of signatures are persistent enough in the specific zone of signature (i.e., curves). The experiments of this study indicate that all signatures with the motion on vertical and horizontal directions have close dynamic features in specific curves. This issue triggers the authors to do more experiments regarding dynamic features of signatures.

The general idea of this work is to find velocity, acceleration and pressure functions of signatures, segmenting the functions to different ranges and finally a comparison between segmented functions of corresponding curves of input signatures and the reference signature.

However, some questions should be answered. What ranges of these functions can be selected for this experiment? How to find the reference signature? Scale and recording angle variant are another hazard in this work. These issues are to be discussed in the rest of the paper

3.2 Dynamic Features Stability Experiment

The introduced device is used to record the dynamic signatures outputs including the coordination, pressure, Azimuth and altitude in the specified interval. In this work, coordination and pressure of points are used.

Pressure: two resolution levels of 1024 or/and 2048 for pressure of points are directly available.

Velocity: considering the constant interval of time for signature record, as in equations (1) it is possible to calculate their velocity by calculating the difference between length and width.

$$\left.\begin{aligned} V_x(i) &= x(i) - x(i-1) \\ V_y(i) &= y(i) - y(i-1) \end{aligned}\right\} \quad V(i) = \sqrt{V_x^2(i) + V_y^2(i)} \quad \begin{aligned} x(0) &= x(1) \\ y(0) &= y(1) \end{aligned} \tag{1}$$

In the above equation, *x(i)* and *y(i)* are defined as *i*th samples coordination. V_x and V_y indicate the velocities in the direction of *x* and *y* respectively. *v(i)* represents the velocity in points *i*. The measure of *v* has the same size of *x* and *y*. All values of vector *v* are positive.

Acceleration: as velocity, acceleration is calculated by the difference between velocities. Equations (2) and (3) are the associated formulas.

$$a(i) = V(i) - V(i-1) \tag{2}$$

$$\mathrm{V}(0) = \mathrm{V}(1) \tag{3}$$

a(i) is the acceleration of *i*th point and V(i) is the velocity of *i*th point.

3.2.1 Length Equalization

One of the problems ahead is that the signatures recorded by a person do not have the same length even in small sequential times. In other words, many factors such as standing or sitting of a person may affect on signing. It is necessary for a signature verification system to consider these factors. These are typical in realistic scenarios. Therefore, the best way to consider these factors is equalization to a reference.

Because of equal time interval of samples in signing and velocity vector independence from pen tip direction, the length can be calculated by sum of all points' velocities. Using equations (4) to (7) is the way to reach signature length.

$$\begin{aligned} X(i) &= V(i).t + X(i-1) \\ X(i+1) &= V(i+1)t + X(i) \end{aligned} \tag{4}$$

$$\begin{aligned} &\vdots \\ X(N) &= V(N)t + X(N-1) \end{aligned} \tag{5}$$

$$\Rightarrow \mathrm{X(N)} = \sum_{\mathrm{i}=1}^{\mathrm{N}} \mathrm{V(i)t} \tag{6}$$

$$t = cte \Rightarrow X(N) = \sum_{i=1}^{N} V(i) \tag{7}$$

In equations (4) to (7), X(i) indicates length of signature from initiation to the point i. V(i) is the velocity of the point and *t* is defined as time. Considering the relative velocity and position, the constant value *t* is removed from the equations. All signatures recorded by each person follow this procedure and curve lengths of all signatures are obtained. The objective is to equalize the lengths. So the average length is considered as a reference for the person's signatures. This average value is assumed for all signatures. New x and y are computed using following equations.

$$\mathrm{r} = \frac{\mathrm{L}}{\ell} = \frac{\sqrt{\mathrm{X}^2 + \mathrm{Y}^2}}{\mathrm{x}^2 + \mathrm{y}^2} \tag{8}$$

Substituting X=r.x, Y=r.y, equation (8) will change to equation (9):

$$r = \frac{\sqrt{r^2.x^2 + r^2.y^2}}{\sqrt{x^2 + y^2}} \tag{9}$$

r is defined as reference length in proportion to current signature length. *L*, *X* and *Y* are also defined as curve length, reference length and reference width respectively that *r*, *x*, *y* are transformed to them. So if *r* is multiplied by *x* and *y* functions, all signatures of a person will have same length. *r* is different for each user signature.

3.2.2 Rotation Normalization

The angle of recording signature is another issue. Identical signing angle is essential for correct verification. All signatures are matched with binary image that the signature pixels are depicted with white colour. First, a signature is considered as a reference randomly and angles of all signatures are equalized to the angle of this reference signature. All signatures are rotated from – 90 to +90. In each step, cross correlation of rotated signature and reference is calculated and analyzed. Angle with the maximum correlation amount is considered to rotate the signature. This rotation is done around the centroid of signature. Figure 5 shows rotated signatures of a sample signature for 7 different angles. In figure 5 the intensity of signatures are proportional with their cross correlation with the reference signature.

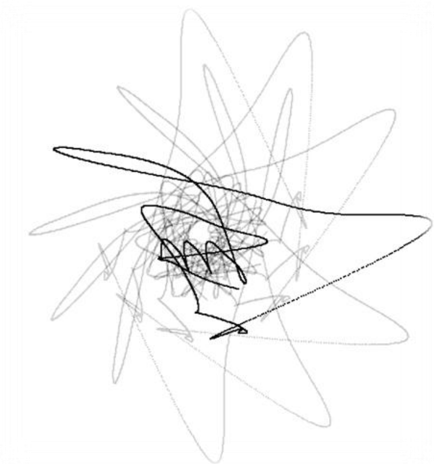

Fig. 5. Rotating signatures around their centroid. The intensity of signatures are proportional to their cross correlation with the reference signature.

3.2.3 Functional Segmentation

Velocity, pressure and acceleration histogram are closely similar to normal distribution. Histogram of sample signature acceleration is shown in figure 6.

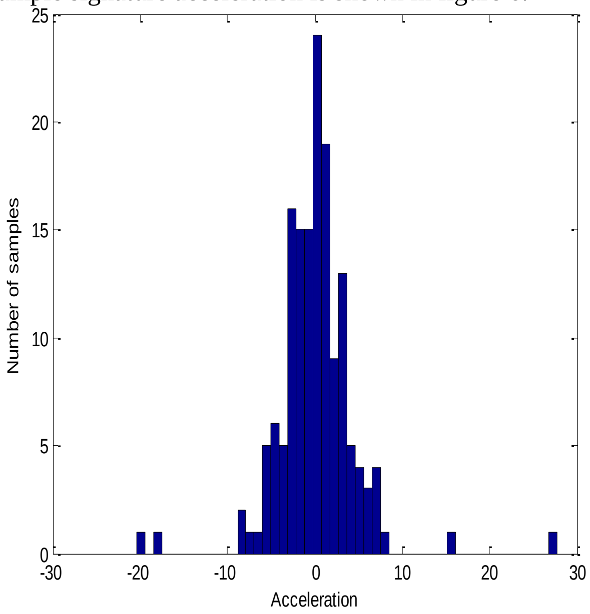

Fig. 6. acceleration histogram of a signature

Therefore, mean and standard deviation of each dynamic parameter are calculated from formulas (10) and (11):

$$m_f^i = \frac{1}{N}\sum_{j=1}^{N} f(j) \tag{10}$$

$$\sigma_f^i = \left(\frac{1}{N}\sum_{j=1}^{N}(m_f^i(j) - f(j))\right)^{\frac{1}{2}} \tag{11}$$

As a result, all segmentation zones of functions are calculated by two above-mentioned parameters. Boundary values, i.e. $m^i - \sigma^i, m^i$ and $m^i + \sigma^i$ divide a signature into four zones. This way is done for all three functions, i.e., velocity, pressure, and acceleration. To improve the visualization of these segments, four colours are assigned to each segmented zone. Table 1 shows the colours associated areas.

Table 1. Colours used in signature segmentation and associated areas.

Segmented colour	Function area
Red	$f > m^i - \sigma^i$
Green	$m^i - \sigma^i < f < m^i$
Blue	$m^i < f < m^i + \sigma^i$
white	$f > m^i + \sigma^i$

Parameter f is the considered function. For instance, if the velocity is the intended function, all zones of the signature that their velocity is less than $m^i - \sigma^i$ are shown in red, values between $m^i - \sigma^i$ and m^i are shown by green, values between m^i and $m^i + \sigma^i$ are in blue and values more than $m^i + \sigma^i$ are shown by white. Background pixels are black. Figure 7 shows two genuine signatures coloured by this way.

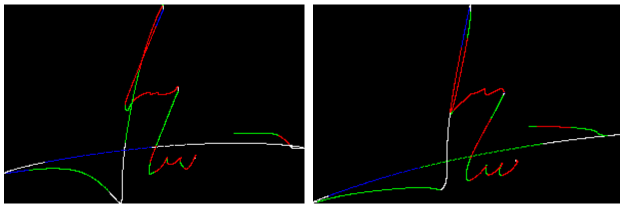

Fig. 7. Signature segmentation with respect to its acceleration

Now this data includes signatures with identical curve length, identical record angle, and four coloured types of zones. The conformity between forgery signature and genuine signature is analyzed in the following.

3.2.4 Conformity

All signatures recorded by a person are conformed. More conformity of the zones with same colour leads to more stability of the selected boundary for the specific feature.

Due to better observation and regardless of trivial changes of lines, signatures are thickened by morphological dilation. Images of the signatures are separately conformed according to the colours. The quantities of genuine signatures lines pixels that conformed on other signatures lines are counted for each colour and indicated by $C_{g,color}^{i,function}$ .The area of all conformed genuine signatures are calculated in pixel and defined as $A_{g,color}^{i,function}$ (i indicates the signers number, *function* is function names like pressure, velocity and acceleration and g shows the parameters are calculating for genuine signatures). It is obvious that more $R_{g,color}^{i,function} = \frac{C_{g,color}^{i,function}}{A_{g,color}^{i,function}}$ leads to more stability of range and type of dynamic function for genuine signature. For more confidence, it is done on the forgery signatures existing in the dataset. The values of $C_{f,color}^{i,function}$ and $A_{f,color}^{i,function}$ are calculated. As expected, the value of $R_{f,color}^{i,function} = \frac{C_{f,color}^{i,function}}{A_{f,color}^{i,function}}$ is not large in the latter set.

$S_{color}^{i,function} = \frac{R_{g,color}^{i,function}}{R_{f,color}^{i,function}}$ expresses the conformity of considered feature and separation between forgery signature and genuine signature for one person, e.g., when twenty third person of recorded signatures considered, the value of $S_{green}^{23,p}$ indicates the proportion of genuine signatures and separation between genuine signatures and forgery signatures via the pressure between $[m_p^{23} - \sigma_p^{23}, \sigma_p^{23}]$ of twenty third person.

These values are computed for all people and finally in equation (12):

$$S_{color}^{function} = \frac{1}{I}\sum_{i=1}^{I} S_{color}^{i,function} \tag{12}$$

Where I is total number of signers and $S_{color}^{function}$ expresses final parameter that is defined as measure of suitableness for the selected feature and this value is obtained from average of $S_{color}^{i,function}$ for all signers of the dataset.

Table 2 shows the results of experiment.

Table 2. Results of "Dynamic features stability experiment"

	Persian Signatures Datasets			International Dataset
Datasets Parameters	Dataset 1 (NDSD)	Dataset 2 (Dehghani) [5]	Dataset 3 (Zoghi) [15]	Dataset 4 (SVC2004) [16]
S_{red}^{V}	1.7248	0.3699	0.9651	1.1480
S_{green}^{V}	1.9499	1.3940	1.9425	1.2528
S_{blue}^{V}	1.9599	1.5659	1.9941	1.2245
S_{white}^{V}	0.0291	1.3048	1.9542	0.0225
S_{red}^{a}	1.8576	1.3977	1.8470	1.1194
S_{green}^{a}	1.4356	0.0232	0.4384	1.1344
S_{blue}^{a}	0.4320	0.0253	0.4559	1.1375
S_{white}^{a}	1.7898	1.3665	1.9820	0.0321
S_{red}^{p}	1.3813	0.0382	0.0134	1.2626
S_{green}^{p}	1.9458	1.4281	1.8343	1.1862
S_{blue}^{p}	1.7097	1.4230	1.8439	1.2027
S_{white}^{p}	0.2330	1.0725	1.8233	0.1233

3.3 Feature Extraction

Considered features for a sample generally include the difference between one or several parameters based on a single template signature in this paper. In fact, during the process described as follows, a signature is specified as a reference signature. The closer signature to the reference signature results in more possibility to be the genuine signature.

Pressure, velocity, acceleration and angular velocity are analyzed in the following. So, extracted features are considered as functional features. Features are classified into three categories that are described in the following.

3.3.1 Critical Samples

As mentioned before, Persian signatures have distinctive features in specific areas, e.g. these signatures often have more curves compared to other signatures. Velocity, acceleration and pressure have specific state in these areas. Therefore, it is possible to compare the curves of the signature associated to specific area of triple functions (i.e. velocity, acceleration, and pressure) with the curves extracted from reference signature.

As for this experimental result, each of triple functions is analyzed in four zones. The best range of the best function was explored. Based on this observation, pen tip velocity in range of m and $m+\sigma$ is the best criterion that is more stable in genuine Persian signatures.

Figure 8 illustrates the velocity diagram of signature. Suppose that the goal is separation of the samples that their velocity is more than average velocity and less than its standard deviation. After calculation of velocity and sample separation from signature, the comparison is applied among samples associated to the considered curves.

In spite of the experiment, the samples were used instead of associated signature curves. Because considering samples, cover another point that is important in signature verification. The point is total samples of signature that indicates total time of recording a signature.

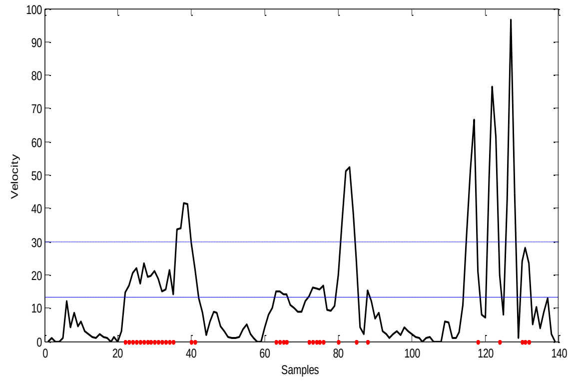

Fig. 8. The black curve is velocity curve of a sample and blue dashed lines are for illustrating boundary points.

3.3.2 Maximum Velocity Area

As noted before (section 3.1), the maximum value of velocity and its position in signature could be essential and play an important role in Persian signatures. Therefore, by windowing the velocity functions and shifting them in time axis, a number is assigned to each window.

The window with maximum number is recorded for the considered signature. The difference between these numbers and recorded number of reference signature specify the next feature. "Maximum velocity area" feature formulas are represented in equation (13) to (17).

$$rect(a,b)=\begin{cases}1 & a-\frac{b}{2}<n<a+\frac{b}{2}\\ 0 & o.w\end{cases} \tag{13}$$

$$T_{i,j}^{k}=rect(j\times(N-C),N)\times f_i^k \tag{14}$$

$$S_{i,j}^{k}=\frac{1}{N}\sum_{j=1}^{N}T_{i,j}^{k} \tag{15}$$

$$S_i^k=\arg\max_{j=1:N}(S_{i,j}^k) \tag{16}$$

$$feature2_i^k=distance(S_i^k,S_{2\_Template}^k) \tag{17}$$

As noted in equation (13), the function *rect* is a rectangular function centralized in *a* and including *b* samples. In equations (14) to (17), *j* indicates window number. *N* and *C* indicate length and the overlapping respectively. $T_{i,j}^{k}$ represents separated window from *i*th signature related to *k*th person. $S_{i,j}^{k}$, S_i^k, 2-*Template* and *feature2* indicate average value of samples in window,

window number with the maximum $S^k_{i,j}$, the second extracted feature for reference signature and finally the second extracted feature respectively.

3.3.3 Relative Angular Velocity

A forger might do a dynamic forge if he/she knows dynamic features or can do shape based forgery. However, forging dynamic features and signature shape simultaneously is too hard even if the forger has all signature information and signature shape. In fact, relative angular velocity is changing signature line for two sequential samples.

This feature is calculated with the formulas that are given in equations (18) to (21).

$$\omega' = \frac{\Delta\theta}{\Delta t} \xrightarrow{\Delta t=cte} \omega = \Delta\theta = \frac{\Delta y}{\Delta x} \tag{18}$$

$$\omega^k_{i,n} = \frac{y^k_{i,n} - y^k_{i,n-1}}{x^k_{i,n} - x^k_{i,n-1}} \tag{19}$$

$$\omega^k_{\omega\_Template,n} = \frac{y^k_{3\_Template,n} - y^k_{3\_Template,n-1}}{x^k_{3\_Template,n} - x^k_{3\_Template,n-1}} \tag{20}$$

$$feature3^k_i = distance(\omega^k_i - \omega^k_{3\_Template}) \tag{21}$$

In equations(18) to (21), ω' is angular velocity, x and y are samples coordination, $\omega^k_{i,n}$ is relative angular velocity of *i*th signature of signer number *k* in point number *n*. In addition, distance function is distance between signatures. *3-template* and *feature3* are reference signature of third category and third category feature.

Figure 9 illustrates extracted features of Zoghi Dataset's signatures respectively by blue dots and red triangles.

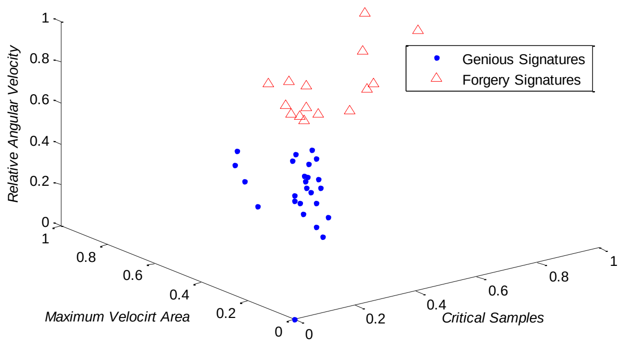

Fig. 9. Spatial illustration of features extracted from signatures. Blue dots are spatial representation of genuine signatures and red triangles illustrate the forgery signatures.

Two points are important in the proposed feature extraction.

a. The first two categories (critical samples and maximum velocity area) are applied to express dynamic features of Persian signatures. Relative angular velocity is complementary for two previous categories. These three categories show both signature features (dynamic features and shape-based) and type (Persian signature) behavior.
b. Calculation of difference between input signature samples and reference signature sample is required for all three categories feature extraction. Therefore, a similarity criterion is used. Dynamic Time Warping and Euclidean distance were options for this work. Because of less time consuming and good result, Euclidean distance is selected. Therefore, distance in all parts of this paper is Euclidean distance.

3.4 Reference Selection

3.4.1 Reference Signature Selection

A signature is selected as a reference in extraction of each feature and this signature is considered the best signature sample according to that feature.

To select reference signature, several genuine signatures are randomly selected as template signatures. Then the distances between each two template signatures are computed using the specified feature. The signature that sum of its distance to other signatures is less than other signatures is selected as specific reference for the feature [9].

3.4.2 Reference Vector Selection

As mentioned before, signatures of one person may extremely vary in different tries. It is reasonable to expect that a signature with great difference may be recorded as a genuine signature. Since selecting template signatures is done randomly, these unreliable signatures have chance to be selected as reference signature. Unreliable Reference Signatures (URS) cause poor verification. In order to eliminate URS another method is represented. In reference vector selection, all of the template signatures participate in producing a reference signature. In this method after selecting templates signatures, the difference between each signatures pair is computed. Sum of difference between a signature and other templates is calculated. Uniformly, a number as a total difference is assigned to each template signature. The smaller number shows more similarity to other templates and should have more effect on reference selection. Differences between the maximum values are assigned to templates and each signature achieves impact weights of templates. Weighted average with these weights indicates to the reference vector for each feature.

3.5 Classifier

3.5.1 Support Vector Machine

Support vector machine [1] is a new tool to pattern recognition. Basically, SVM uses a hyper plane to separate two classes[17]. Support vector machine leads to decrease structural risk. Whereas artificial neural network decreases

[1] SVM

experimental risk [18]. This point causes to increase generalization and better training with few train samples.

As shown in figure 8, a classifier with linear separation is needed. Support vector machine separates data into genuine and forgery data.

3.5.2 Decision Making

Support vector machine is applied as a classifier indecision-making phase. Seven genuine samples and seven forgery samples regarding Zoghi [15] dataset, NDSD and SVC and six genuine samples and six forgery samples regarding Dehghani's [5] dataset are used to train classifier. Genuine data used in training stage are the features extracted from template signatures as noted (in reference signature selection section). As a result eight samples (4 genuine and 4 forgery signatures) from dataset1, 26 samples (18 genuine and 8 forgery signatures) from dataset2, 91 samples (58 genuine and 33 forgery signatures) from NDSD and 26 signatures (13 genuine and 13 forgeries) from SVC are used to test the proposed algorithm.

4. Results

Three criteria are used to verify the performance of signature verification algorithms, i.e. False Acceptance Rate (FAR), that is the error rate that the classifier incorrectly claims acceptable and False Rejection Rate (FRR) that is rate of incorrect rejections. Equal Error Rate (EER) is defined as error percentage when FRR and FAR are equal. EER is considered as the main criteria to study the performance of the algorithms. FAR and FRR change by the variation in classifier threshold.

A hyper plane is defined to separate two class patterns. In SVM, to achieve verification error rates as algorithms measure, a threshold value is applied to classification results before final decision making (sign function). Actually, the hyper plane bias is used as the threshold to change error rates (FAR and FRR). The Euclidian distance between patterns and separating hyper plane are the class membership degrees.

Two types of threshold (General and Individual) are applied to data to achieve error rates. A discussion on both types of thresholds and their associated verification results will come in the next section.

As stated before in section "Dynamic features stability experiment" and "Feature Extraction", three categories of features have been extracted from each signature sample. Each feature expresses three major characteristic of Persian signatures. One of triple functions and range of the function amount used in "critical samples" is selected in this experiment. Section "maximum velocity area" extracts special behavior of curves in Persian signatures and third category is shape-based feature.

Table 3 represents the result of verification.

Table 3. Result of the proposed algorithm on three Persian dataset and international SVC2004 dataset and comparison with best previous result

	Datasets	Best previous EER (%)	Proposed algorithm EER (%)
Persian Datasets	First Dataset (Dehghani)	3.95[5]	3.12
	Second Dataset (Zoghi)	11.2[15]	3.98
	NDSD	-	4.26
International Dataset	SVC2004	3[8]	4.58

5. Discussion

In previous section, the results of signature verification with seven genuine and seven forgery signatures (except for first dataset with six genuine and six forgery signatures) are illustrated. This method was implemented with different numbers of training samples on Persian datasets. Results (Table 4) show that the algorithm performance is relatively acceptable for few training data.

The top of this paragraph illustrates a sub-subheading.

Table 4. Results (EER) of verification with different number of training samples. Half of training numbers on each row are from genuine signatures and other signatures are from forgery group.

		Dataset1 (Dehghani)	Dataset2 (Zoghi)	NDSD
Number of training sample	4	8.26	7.66	10.83
	8	6.41	5.11	7.14
	12	3.12	4.51	5.13
	14	3.04	3.98	4.26
	20	-	2.32	4.01
	40	-	-	3.17

Of course, the verification by few training data is unreliable. Nevertheless, it can show the reliability of features and capability of leading classifier to discriminate the classes. Large variety of people signatures and especially Persian signatures may cause lower rate of verification with few number of training samples. As noted in "selecting reference" section, if the template signatures are selected from limited number of signatures, the probability of selecting bad references will increase. Hence the algorithm process should run repeatedly for achieving correct performance rate.

As illustrated in Table 2, the selected zone and function $(m^i < V < m^i + \sigma^i)$ of all three Persian signatures datasets is equal, while it is different from the international one. Since all three categories are based on Persian signatures datasets, the presented method is not significant for international SVC2004 dataset. This method was applied on SVC2004 dataset and as expected, the result of verification was not better than previous works. In Table 2 it can be seen that the S^V_{blue} is the parameter selected for Persian dataset and S^V_{green} is for international one. Since the signature samples are not enough, this criterion is not reliable. However both two features (selected features from "critical samples" category) with two other feature categories are applied to SVC2004 dataset. The results (see Table 5) show that SVC2004 selected parameter with two other Persian-based features leads to better results than the other one.

The results of $S_{color}^{function}$ in section 3.2.4 show that $m^i < V < m^i + \sigma^i$ is the best option for extracting critical samples of Persian signatures. This result was tested on three available Persian datasets and triple function and different zones selected for extracting critical samples feature in the test.

Table 5. Verification results of SVC2004 with two different parameters (Persian and SVC2004's based) of first "critical samples" category

Parameter	EER (%)
S_{blue}^{V} (Persian based)	4.58
S_{red}^{P} (For SVC)	4.39

Figure 10 illustrates the verification results for different options of critical samples feature and two other constant features mentioned in feature extraction section.

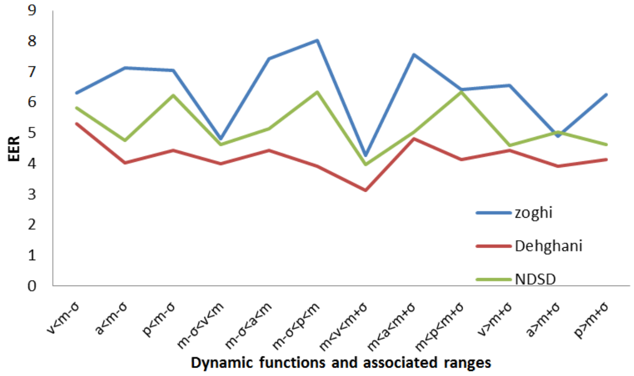

Fig 10. Verification result with different critical sample feature

As mentioned in previous section the threshold can be chosen for all writers or set individually one for each signer. A common threshold is used for the entire enrolment data from all the signers. This threshold is applied to a set containing all data.

To adopt the verification process to the single signers' properties, a signer dependant threshold should be applied. In Table 6the results of the two threshold types are listed.

Table 6. Comparison between applying general and individual thresholds

	Datasets	General Threshold EER (%)	Individual Threshold EER (%)
Persian Datasets	First Dataset (Dehghani)	3.78	3.12
	Second Dataset (Zoghi)	4.52	3.98
	NDSD	4.64	4.26
International Dataset	SVC2004	5.55	4.58

In "Dynamic features stability experiment" critical curves have been used and because of dependency to time factor, critical samples are used instead. However, both types of critical samples and critical curves were used as features and their results are illustrated in Table 7.

Table 7. Comparison between critical curves and critical samples as third used feature

Datasets	Critical Curves Feature EER (%)	Critical Samples Feature EER (%)
First Dataset (Dehghani)	6.02	3.12
Second Dataset (Zoghi)	6.25	3.98
NDSD	6.53	4.26

Another key feature in an identification system is its usability. Actually real forgery signatures are not available and an identification system must be independent of forgery signatures. The purpose of the system is to use genuine signature for verification. The proposed algorithm is based on simulating forgery signatures by random patterns. As illustrated in figure 9 forgery signatures are at large distance from genuine with high scatters. The random points with normal distribution and equal mean and standard deviation are used as representative of forgery signatures. Table 8 shows mean and standard deviation of three independent features explained in section 3.3 of forgery signatures in three Persian datasets.

Table 8. Normal distribution parameters of forgery signatures in three Persian datasets

Datasets	Forgery Signatures Features Mean			Forgery Signatures Features Standard Deviation		
	1	2	3	1	2	3
First Dataset (Dehghani)	0.20	0.33	0.62	0.24	0.28	0.22
Second Dataset (Zoghi)	0.45	0.31	0.45	0.27	0.24	0.20
NDSD	0.41	0.53	0.72	0.27	0.30	0.21
Average	0.36	0.39	0.6	0.26	0.28	0.21

Their average values are used to produce random patterns independently. In the last row of Table 8 final normal distribution parameters for Persian signatures are shown. Verification results of the proposed algorithm with only three genuine signatures are illustrated in Table 9.

Table 9. Verification results with three genuine signature and producing random features instead of forgeries

Datasets	EER (%)
First Dataset (Dehghani)	16.67
Second Dataset (Zoghi)	12.14
NDSD	5.17

Due to independence of signers and forgers, and low computational complexity, the method can be practical for real world applications.

6. Conclusion and Future Works

An online signature verification based on special properties of Persian signatures is presented. Usually Persian signers move their wrist and fingers more than signers of other languages do and these motions cause variation in dynamic features. An experiment has been designed to explore robust features and the best one has been selected. Two dynamic features and relative angular velocity are extracted from signatures and the distance from reference signatures are used as the input to classifier. A linear SVM is used to classify signatures. The results of verification illustrated that acceptable EER was achieved.

Signature is the behavioral biometric that changes in different tries. In the proposed algorithm, distance between reference and input signature is the verification basis. Selecting bad reference leads to bad verification. In the algorithm, the probability of selecting improper signature as reference is not zero. It is expected that an intelligent method that specifies best representation of genuine signature cause less EER. Also using only two or three genuine signatures that make verification system more practical will be possible if the mentioned method works.

Acknowledgments
The author would like to thank the anonymous reviewers and editors for the useful and constructive comments which have improved the article.

References

[1] International Biometric Group(IBG). Available: http://www.biometricgroup.com
[2] A. K. Jain, P. Flynn, and A. A. Ross, Handbook of Biometrics. USA: Springer, 2008.
[3] K. R. Radhika, G. N. Sekhar, and M. K. Venkatesha, "Pattern recognition techniques in on-line hand written signature verification - A survey," in Multimedia Computing and Systems, 2009. ICMCS '09. International Conference on, 2009, pp. 216-221.
[4] C. E. Pippin, "Dynamic Signature Verification using Local and Global Features," 2004 July 2004.
[5] E. Dehghani and M. E. Moghaddam, "On-line signature verification using ANFIS," in Image and Signal Processing and Analysis, 2009. ISPA 2009. Proceedings of 6th International Symposium on, 2009, pp. 546-549.
[6] T. A. A. Alizadeh, Z. Daei, "Optimal Threshold Selection for Online Verification of Signature," presented at the IMECS2010, Hong Kong, 2010.
[7] L. Nanni, E. Maiorana, A. Lumini, and P. Campisi, "Combining local, regional and global matchers for a template protected on-line signature verification system," Expert Systems with Applications, vol. 37, pp. 3676-3684, 2010.
[8] A. Fallah, M. Jamaati, and A. Soleamani, "A new online signature verification system based on combining Mellin transform, MFCC and neural network," Digital Signal Processing, vol. 21, pp. 404-416, 2011.
[9] M. A. U. Khan, M. K. K. Niazi, and M. A. Khan, "Velocity-Image Model for Online Signature Verification," Image Processing, IEEE Transactions on, vol. 15, pp. 3540-3549, 2006.
[10] M. Parodi and J. C. Gómez, "Legendre polynomials based feature extraction for online signature verification. Consistency analysis of feature combinations," Pattern Recognition, vol. 47, pp. 128-140, 2014.
[11] M. Parodi, Go, x, J. C. mez, M. Liwicki, and L. Alewijnse, "Orthogonal function representation for online signature verification: which features should be looked at?," Biometrics, IET, vol. 2, pp. 137-150, 2013.
[12] K. Cpałka, M. Zalasiński, and L. Rutkowski, "New method for the on-line signature verification based on horizontal partitioning," Pattern Recognition, vol. 47, pp. 2652-2661, 2014.
[13] Y. Guerbai, Y. Chibani, and B. Hadjadji, "The effective use of the one-class SVM classifier for handwritten signature verification based on writer-independent parameters," Pattern Recognition, vol. 48, pp. 103-113, 1// 2015.
[14] A. Q. Ansari, M. Hanmandlu, J. Kour, and A. K. Singh, "Online signature verification using segment-level fuzzy modelling," Biometrics, IET, vol. 3, pp. 113-127, 2014.
[15] M. Zoghi and V. Abolghasemi, "Persian signature verification using improved Dynamic Time Warping-based segmentation and Multivariate Autoregressive modeling," in Statistical Signal Processing, 2009. SSP '09. IEEE/SP 15th Workshop on, 2009, pp. 329-332.
[16] D.-Y. Yeung, H. Chang, Y. Xiong, S. George, R. Kashi, T. Matsumoto, et al., "SVC2004: First International Signature Verication Competition," ed. http://www.cse.ust.hk/svc2004/results.html, 2004.
[17] C. Gruber, T. Gruber, S. Krinninger, and B. Sick, "Online Signature Verification With Support Vector Machines Based on LCSS Kernel Functions," Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on, vol. 40, pp. 1088-1100, 2010.
[18] K.-j. Kim, "Financial time series forecasting using support vector machines," Neurocomputing, vol. 55, pp. 307-319, 2003.

Mohammad Esmaeel Yahyatabar is a Ph.D at Iran University of Science & Technolgy, Tehran, Iran. He received M.Sc degree in Electronic Engineering from Babol University of Technology, Babol, Iran. His research interests are: signature verification and non-destructive testing.

Yasser Baleghi is an Assistant Professor of Electronic Engineering at Babol University of Technology. He holds a Ph.D degree in Electronic Engineering from Iran University of Science & Technology. His research interests are evolvable and adaptive hardware, image processing and fault tolerant system design.

Mohammad Reza Karami received the B.Sc degree in Electronic Engineering in 1992, M.Sc. degree in Signal Processing in 1994, and the Ph.D degree in Biomedical Engineering from I.N.P.L d'Nancy of France. He is now the Associate at the Department of Electrical & Computer Engineering, Babol University of Technology. Since 1998 his research is in signal and speech processing. He has published 80 articles in journals and conferences.