handwritten signatures,
dynamic signature features,
linear regression method,
dynamic time warping

Rafał DOROZ*, Monika MITAS**

NEW METHODS TO DETERMINE SIMILARITY OF SIGNATURES BASED ON LOCAL EXTREMES

Authentication based on handwritten signature is one of the most accepted authentication systems based on biometry. In this paper a method for the automatic verification of on-line handwritten signatures using three similarity measures is described. The proposed approach, is based on extreme values and dynamic features of the signature. In investigations proposed coefficients together with the factor R^2 were connected and new signature recognition quality has been achieved.

1. INTRODUCTION

Biometric technologies attempt to make use of people's biological characteristics that identify and match these, mainly for security. We can distinguish several types of biometric identification types e.g.: face - the analysis of facial characteristics; fingerprint - the analysis of an individual's fingerprints; hand geometry - analysis of the shape of the hand and the length of the fingers; signature - the analysis of the way a person signs his name; vein - the analysis of pattern of veins in the back of the hand and the wrist; voice - the analysis of the tone, cadence and frequency of a person's voice[4,5].

All biometric techniques operating principle are very similar. In the first stage, device captures a sample of the biometrics, then features are extracted and converted into a mathematical form [8]. In the process of identification, a sample is compared with stored templates. There is no perfect biometric technique. Some biometric techniques can be more or less suitable for particular applications.

The efficiency of biometric technologies can also be improved by combining few techniques. In our research new methods based on that assumption have been proposed. Each of these methods analyzes dynamic features of the signature.

The signature data analyzed in our research were produced by means of the SigLite LCD 4x3 tablet, where the static and dynamic signature features can be captured. It allows to measure some features, such as: signature time intervals, pen-pressure and pen location on the tablet surface. The sizes of the active area, where the signature can be putted are restricted to: 112 x 64mm. The speed of the data conversion is: 377 points per second, and device resolution is: 410 ppi. The signatures signed by 30 different people were the set of the input data for the research. Each of them made 5 signatures every several days. Finally, the signature database contains 150 signatures.

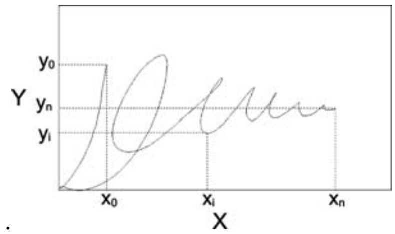

Fig.1 Coordinates (x_i, y_i) of the analysed signature.

2. SIGNATURE SIMILARITY DETERMINATION USING LINEAR REGRESSION METHOD

One of the methods, which is most often used for determination of the functional dependency between two data strings is the method of the smallest squares [6]. This method enables form mathematical formulas describing the relation between the studied elements.

In the described method as the measure for the quality of point adjustments of different strings, the coefficient R^2 was adopted, where the variables X and Y can be correlated: [6]

* University of Silesia. Institute of Informatics, 41-200 Sosnowiec, Będzińska 39, Poland
** Silesian University of Technology, The Faculty of Mathematics and Physics, Institute of Mathematics, 44-100 Gliwice, Kaszubska 23, Poland

$$R^2 = \frac{\left[\sum_{i=1}^{n}\left(x_i - \bar{x}\right)\left(y_i - \bar{y}\right)\right]^2}{\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2 \sum_{i=1}^{n}\left(y_i - \bar{y}\right)^2}; \tag{1}$$

where values of R^2 have the following interpretation:

1. Reflexivity, i.e. $R^2(X,X) = 1$,
2. Symmetry, i.e. $R^2(X,Y) = R^2(Y,X)$,
3. $R^2 \in [0,1]$ where R^2=1 means the full linear conformity (high similarity), and R^2=0 means the lack of linear conformity (the lack of similarity).

3. SIMILARITY MEASURE R^2

One of the coefficients, used to compare two signatures, is R^2. To widen the possibilities and to improve the working of the program, two new signature similarity measures were elaborated. In the next part of the paper, they will be called EVP and MEV. Using EVP and MEV it is possible to calculate the similarity between two signatures basing on the marker's movement in the vertical and horizontal direction and the pressure of the marker.

4. SIMILARITY MEASURE EVP (EXTREME VALUES PROPORTIONS)

The measure EVP bases on extreme values. If in two signatures, that are being compared, the number of extreme values is different, then it is equalized. From the more numerous set of extreme values we take only a subset, which contains the most distinctive extreme values [3]. While elaborating the measure, we first of all were guided by the assumption, that if two signatures come from the same person, the proportions between the following minima and maxima do not change. Let k be the number of extreme values, $a_1, a_2, ..., a_k$ the values of the consecutive extreme values in the first signature, and $b_1, b_2, ..., b_k$ the analogous values in the second signature. Let also $r_1, r_2, ..., r_{k-1}$ be the distances between the abscissas of the consecutive extreme values in the first signature and $s_1, s_2, ..., s_{k-1}$ the analogous values in the second signature.

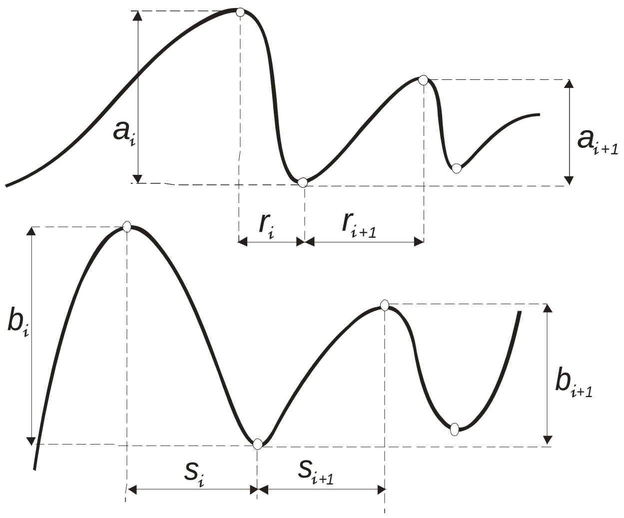

Fig.2 . The values and distances between the abscissas of the consecutive extreme values

Let's now define:

$$\forall i \in \{1,2,...,k-1\} \;\; A_i = \frac{a_i}{a_{i+1}}; \tag{2}$$

$$\forall i \in \{1,2,...,k-1\} \;\; B_i = \frac{b_i}{b_{i+1}}; \tag{3}$$

$$\forall i \in \{1,2,...,k-2\} \;\; R_i = \frac{r_i}{r_{i+1}}; \tag{4}$$

$$\forall i \in \{1,2,...,k-2\} \;\; S_i = \frac{s_i}{s_{i+1}}. \tag{5}$$

then:

$$EVP = \prod_{i=1}^{k-1} \min\left\{\frac{A_i}{B_i}, \frac{B_i}{A_i}\right\} \cdot \prod_{i=1}^{k-2} \min\left\{\frac{R_i}{S_i}, \frac{S_i}{R_i}\right\}. \tag{6}$$

5. SIMILARITY MEASURE MEV (MEAN AND EXTREME VALUES)

Measure MEV, similarly as measure EVP, bases on the local extreme values. Furthermore, now it is necessary to calculate the average value of every course, which will be denoted as m and global extreme values max_1, max_2, min_1, min_2 [10]. Let the range of the graph be denoted as R and defined as $R=max\{max_1, max_2\}-min\{min_1, min_2\}$. Similarly as by EVP, it is required to equalize the number of extreme values, but this time it is essential to get sure, that in every course the first extreme value is a minimum. Let k be the final number of extreme values, $x_{11}, x_{12},...,x_{1k}$ the values of ordinates of the consecutive extreme values in the first signature and $x_{21}, x_{22},...,x_{2k}$ the analogous values in the second signature.

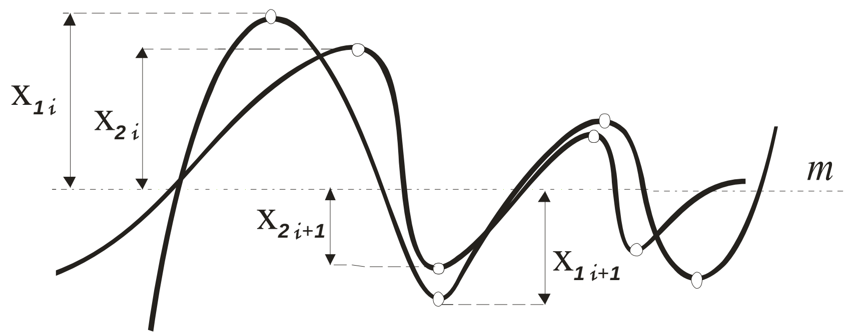

Fig.3 The values of ordinates of the consecutive extreme values

Let's now define:

$$\forall i \in \{1,2,...,k\} \;\; X_i = \frac{|x_{1i} - x_{2i}|}{m_i}; \tag{7}$$

where:

$$m_i = \begin{cases} x_{2i} - m & if \quad x_{2i} \ge x_{1i} \ge m; \\ x_{1i} - m & if \quad x_{1i} \ge x_{2i} \ge m; \\ m - x_{2i} & if \quad x_{2i} \le x_{1i} \le m; \\ m - x_{1i} & if \quad x_{1i} \le x_{2i} \le m; \\ R & if \quad x_{2i} \ge m \ge x_{1i}; \\ R & if \quad x_{1i} \ge m \ge x_{2i}. \end{cases} \tag{8}$$

Finally:

$$MEV = \prod_{i=1}^{k} X_i \tag{9}$$

is the newly defined similarity measure.

6. DYNAMIC TIME WARPING (DTW)

Some inconvenience of the above described methods is the possibility of calculating similarity for only two sequences with the same number of n points. The biometric data used in the signature analysis have different lengths. Dynamic time warping (DTW) [4,9] is an elastic matching technique, that allows nonlinear alignment of sequences and can overcome mentioned limitation.

To align two sequences $X = (x_1, x_2, ..., x_n)$ $Y = (y_1, y_2, ..., y_m)$ using DTW, we construct a matrix $n \times m$ whose $d(i, j)$ element is the Euclidean distance between x_i and y_j (Fig. 2):

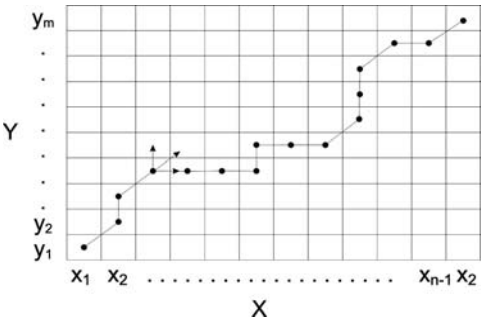

Fig.2 The matrix adjustment plan of an exemplar path

$$D(i,j) = \min \begin{cases} D(i-1, j-1) + d(x_i, y_i) \\ D(i-1, j) + d(x_i, y_i) \\ D(i, j-1) + d(x_i, y_i) \end{cases} ; \tag{10}$$

where $d(x_i, y_i) = \sqrt{(x_i - y\ )^2}$ is the cost of the series matching.

To find the best match between these two sequences we can find a path through the grid which minimizes the total distance between them [4]. Figure 2 shows the optimal paths obtained by comparing two signatures.

The DTW is a simple technique well known in the research community. For this reason the details of this method were omitted.

7. LINEAR AND QUADRATIC FISHER'S CLASSIFICATION FUNCTION

Discrimination variables are calculated using PCA. Let k be the number of classes, p the number of dimension, μ_i the average in the i-th class, Σ_i the covariance matrix in the i-th class and q_i the probability of the occurrence of an element from the i-th class. N_i elements are drawn from each class [7]. Then, the quadratic Fisher's classification function is given by:

$$u_i(x) = -\frac{1}{2} \cdot \Delta_i^2(x) - \frac{1}{2} \cdot \ln|\Sigma_i| + \ln q_i ; \tag{11}$$

where:

$$\Delta_i^2(x) = (x - \mu_i)^T \cdot \Sigma_i^{-1} \cdot (x - \mu_i) . \tag{12}$$

If $\forall i\ q_i = const$, we can omit ln q_i.

The algorithm starts by taking the examined element x_0. Then for every previously worked out u_i, $u_i(x_0)$ is calculated. The element is assigned to this class, for which the value of $u_i(x_0)$ is the highest.

If $\forall i\ \Sigma_i = const$, it is allowed to use the linear Fisher's classification function, which is given by:

$$e_i(x) = -\frac{1}{2} \cdot d_i^2(x) + \ln q_i ; \tag{13}$$

where:

$$d_i^2(x) = (x - \mu_i)^T \cdot S^{-1} \cdot (x - \mu_i) ; \tag{14}$$

S-joined covariance matrix.

Moreover, the following assumptions have to be fulfilled:

a) $N_i \geq p + 1$,

b) $N - k - p - 1 > 0$,

where $N = \Sigma N_i$.

8. RESULTS

The aim of this paper was to study new similarity measures which take advantage of dynamic features. The results of experiments consisted in the investigation of on-line similarity measure using proposed coefficients were compared with results got applying by the R^2 factor.

Experiments, in which the proposed coefficients and the factor R^2 were connected, were conducted additionally.

The experiments were performed on a database consisting of 360 genuine signatures and 240 skilled forgery signatures from 40 individuals. Table 1 shows the EER percentage values obtained for the experiments.

Table 1. EER's values of particular coefficients.

	EVP	***MEV***	***R^2***	***EVP+R^2***	***MEV+R^2***	***EVP+MEV+R^2***
X	11.4	12,1	9,5	10,7	11,1	7,4
Y	7,5	8,0	6,3	7,6	6,6	5,6
P	12,5	9,5	7,3	10,8	7,2	6,1
V_X	8,8	11,9	4,2	5,1	7,2	4,9
V_Y	9,9	12,8	3,3	5,0	6,4	2,8
V_P	21,8	16,2	2,7	6,1	3,6	3,8

EVP method achieves an equal error rate (EER) of 7,5% when only *Y* signature coordinate was considered. From Table 1, it is noted that under the same test conditions, the connected methods EVP, MEV and R^2 shows improvement about 1-2%. When the dynamic features (V_X, V_Y, V_P) are used, the lowest EER is 8,8% for EVP and 11,9% for MEV and 2,7% for R^2. In signature verification system, which uses the EVP, MEV and R^2, the EER is significantly lower.

We have also noticed that the results for method EVP are generally slightly better than those of MEV. This seems to imply that additional dynamic information including local extreme values is very useful and can lead to satisfying results.

9. CONCLUSIONS

In proposed methods of signature analysis, collection of signatures with the pattern where compared. We have studied the effects on the performance (EER) combining different methods. The work focused on studying helpful verification methods. The EVP and MEV methods provide bigger errors than R^2. The investigations showed that the best results had been got through the connection of the new coefficients with the R^2 factor. Obtained results look very promising. In the near future we plan comparisons of our method with other measures and a new multivariate algorithm has been introduced.

BIBLIOGRAPHY

[1] HUANG B.Q., KECHADI M-T., A Fast Feature Selection Model for Online Handwriting Symbol Recognition, Proceedings of the 5th International Conference on Machine Learning and Applications (ICMLA'06), 2006.

[2] ASIF MASOOD1 SHAIQ A., Role of Corner Detection in Capturing Shape Outlines, 7th International Conference on Computer Vision, Pattern Recognition and Image Processing.

[3] FENG H., CHOONG WAH CH.., Online signature verification using a new extreme points warping technique, Pattern Recognition Letters archive Volume 24 , Issue 16 (December 2003), pp 2943-2951, 2003 .

[4] DOROZ R., PORWIK P., PARA T., WRÓBEL K., Dynamic signature recognition based on velocity changes of some features., International Journal of Biometrics, Vol. 1, No. 1, ISSN 1755-8301, pp. 47-62.

[5] PORWIK P., The Compact Three Stages Method of the Signature Recognition., Proceeding of 6th International Conference on Computer Information Systems and Industrial Management Applications (CISIM'07), Ełk. 2007, pp. 282-287.

[6] HANSHENG L., SRINIVAS P., VENU G., ER^2: An Intuitive Similarity Measure for On-Line Signature Verification, Proceedings of the Ninth International Workshop on Frontiers in Handwriting Recognition, pp. 191–195, 2004.

[7] BILLINGS S., LEE S., Nonlinear Fisher discriminant analysis using a minimum squared error cost function and the orthogonal least squares algorithm, 2002 Elsevier Science, pp. 263-270.

[8] AL-SHOSHAN, A.I, Handwritten Signature Verification Using Image Invariants and Dynamic Features, Computer Graphics, Imaging and Visualisation, International Conference on Publication, 2006, pp. 173-176.

[9] PIYUSH SHANKERA A., RAJAGOPALAN A.N., Off-line signature verification using DTW, Image Processing and Computer Vision Lab, Department of Electrical Engineering, Indian Institute of Technology Madras, India 2006

[10] GUPTAA G.K., JOYCEB R.C., Using position extreme points to capture shape in on-line handwritten signature verification, Faculty of Information Technology, 2006.