signature recognition, IPAN99, optimum seeking methods

Rafał DOROZ*, Maciej J. BARGIELSKI*

ON SOME OPTIMALIZATION OF SIGNATURE RECOGNITION

Signature recognition is one of the important problems nowadays. In paper we present known method of pattern (curves) recognition, i.e. algorithm IPAN99 and researches over its optimization; there are many control parameters which influence on recognition results. We present some quasi-optimal set of control parameter. Our next aim is to automatically find proper parameters. Thus some optimum seeking method for unimodale and multimodale function is proposed.

1. INTRODUCTION

Biometrics signature is one of the longest-known security techniques. The methods of automatic signature identification can be found in many areas of life. Signature has for many years adopted a form of determining the credibility (such as during operations related to running a bank account). Signature recognition on the basis of photos do not secure from signature forgery which is quite easy to carry. This causes that in important operations such as authorization a tablet is often used. This device during signing is capable to measure many parameters, such as a pressure pen on tablet surface, position of pen, the angles at which the person holds a pen. It is also able to determine the dynamics of signature. This complex analysis of signatures may be more precise then analysis the signatures submitted on paper and thus forgery is less probable.

There are a number of ways to specify the similarities [8]. Their action are based on the calculation of distance such as Euclidean, [4] or Mahalanobis, [4,10,12]. They may also use different models of neural networks or Hidden Markov Models (HMM). There are also methods based on an analysis of the signatures object characteristics (usually edges or angles points) and reducing the problem to match these points. The maximum size of the signature attributes vector representing the points obtained from a tablet can reach several thousand numbers. This is a lot even for practical reasons (such as storage and transmission of data, biometric verification carried out by mobile devices or smart cards, etc.). Therefore reduction of the vector size is done. Reduction of data can be achieved by identifying the various important points on contour or some attribute functions of a signature. Usually those are the points representing the curves or corners of signature contour. The main problem is to identify such points, which can keep the shape of the object.

The best-known algorithms seeking the points of highest curvature are: Rosenfeld and Johnston 1973 [RJ73], Rosenfeld and Weszka 1975 [RW75], Freeman and Davis 1977 [FD77], Beus and Tiu 1987 [BT87], IPAN99.

In this work an algorithm IPAN99 which has confirmed its high efficiency is used [2,3]. Unfortunately, his disadvantage is the need for proper selection of parameters. The paper proposes a method of testing the similarity of signatures based on detecting and analyzing the signatures of specific points and the choice of parameters IPAN99 algorithm to detect these points.

2. THE IPAN99 ALGORITHM

The algorithm IPAN99 defines the corner as a point of a given curve, if we are able to inscribe in that curve a triangle with a given size and angle of the vertical gap. Curve is represented as a sequence of points P_i.

In the first stage, the algorithm reviews the string of points and selects the candidates for corners. At every point P algorithm is trying to inscribe into the curve a few triangles of different arms lengths (P^-, P, P^+) constructed using a few simple rules:

$$\begin{aligned} d_{min}^2 &\le \left|p - p^+\right|^2 \le d_{max}^2 \\ d_{min}^2 &\le \left|p - p^-\right|^2 \le d_{max}^2 \\ \alpha &\le \alpha_{max}, \end{aligned} \tag{1}$$

where:

$\left|p - p^+\right| = |a| = a$ the distance between points P and P+,

$\left|p - p^-\right| = |b| = b$ the distance between points P and P^-,

$\alpha \in [-\pi, \pi]$ the triangle angle of vertical crack defined as follows: $\alpha = \arccos\frac{a^2 + b^2 - c^2}{2ab}$

* University of Silesia. Institute of Informatics, 41-200 Sosnowiec, Będzińska 39, Poland

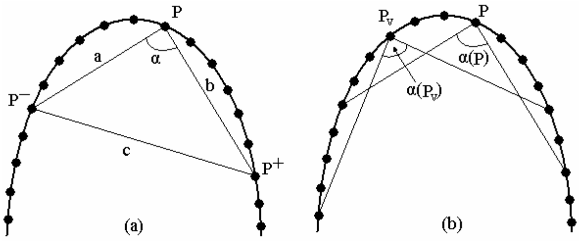

Fig. 1. Detection of points with the highest curvature by means of the IPAN99

Triangle, which meets the conditions (1) is called an acceptable one. The searching for an acceptable triangle is stared from the point P on the outside that is from the shortest length of the triangle sides and stops, if any part of the conditions (1) is fulfilled (therefore the restricted number of neighbouring points is taken into consideration). Among all acceptable triangles, the one with the smallest angle of vertical crack is selected $\alpha(P)$.

Point P is attributed to "the sharpness" expressed as $\pi - |\alpha(P)|$. If no acceptable triangle can not be found, the point P is rejected and has not assigned any sharpness (in practice, sharpness of zero is assigned).

In the second stage, the redundant candidates are eliminated. The candidate point P is rejected, if it has a shaper neighbour that is P_v point, which is also a candidate, and which was assigned a greater strength of the corner:

$$\alpha(P) > \alpha(P_v)$$

P_v belongs to the neighbourhood of P point, if $|P - P_v|^2 \leq d^2_{min}$. Value $|P\text{-}P_v|^2$ which is the distance between two points of greatest curvature will be denoted as D_{IPAN}.

Algorithm parameters:

d_{min} – the parameter restricting the length of sides from the “bottom”. Small values effect that algorithm reacts to small corners.

d_{max} – the parameter restricting the length of sides from the “top”. It is necessary to avoid false acute angles created by distant points of the curve

α_{max} – boundary angle specifying the minimal acuteness, which has to have a point in order to classify it as the candidate for the corner.

3. SELECTION OF PARAMETERS

Each signature was examined as a series of points representing the values of x, y signature coordinates. Then, based on an algorithm IPAN99 characteristic points were selected. IPAN99 algorithm allows easy changing the parameters affecting the sensitivity of the algorithm that is the number of detected characteristic points. Below are examples of the influence of these parameters. Each of the charts shows the x coordinates of two signatures together with the characteristic points founded by algorithm. Their number strongly depends on parameters of the algorithm IPAN99. Definitely the greater amount of points shows chart 2a. One can see that the points were located in specific places which are relevant and intuitive to describe function. This graph is also characterized by high density of points lying close to each other. For comparison in chart 2b number of points lying close to each other is by far less, while preserving the discovery of key points. This is the optimal situation; therefore it can be assumed that the parameters established for this measurement have been chosen very well. The worst situation is graphed by 2c. In this case there are a small number of specific points but unfortunately they have been detected at the point which does little to reflect the shape of the investigated signature. In contrast to previous charts restoration of the functions on the basis of founded characteristic points is practically impossible.

a)
b)
c)

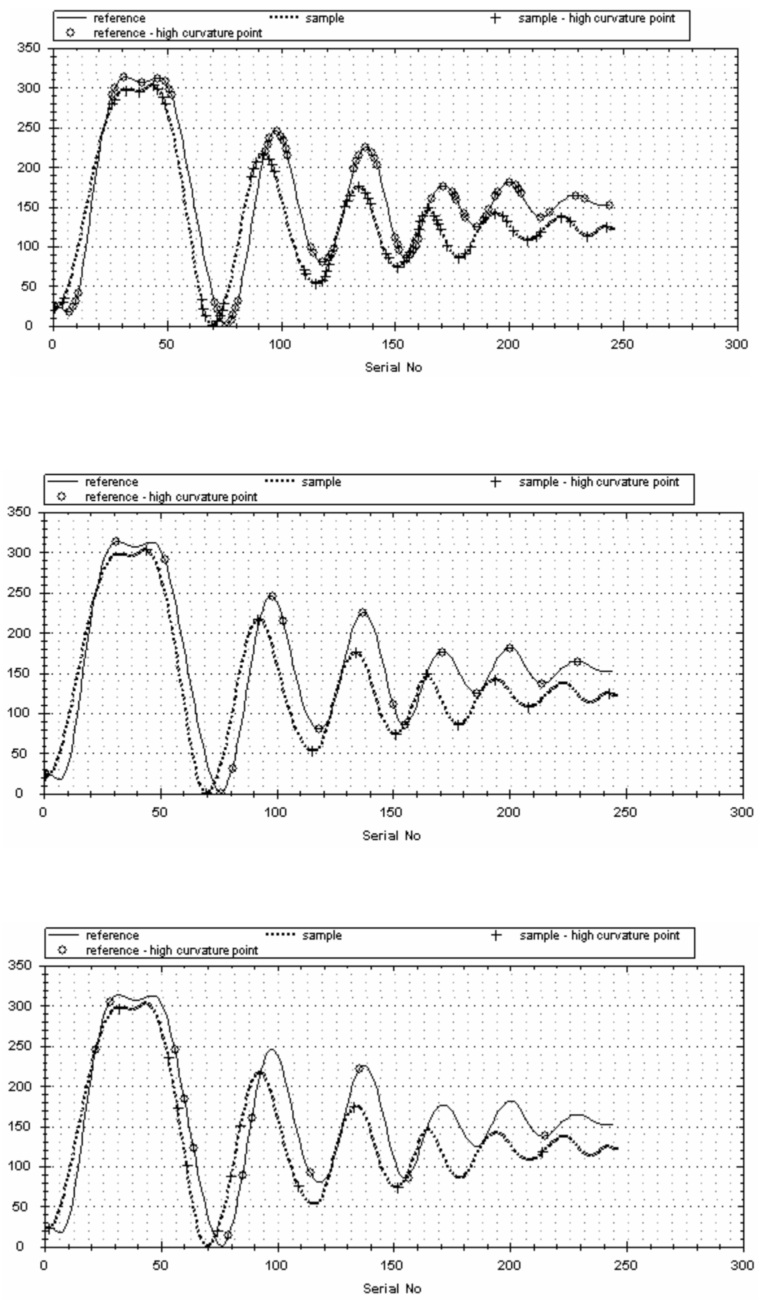

Fig. 2. Charts of x coordinate of two signatures indicating characteristic points.
(a) the parameters of algorithm IPAN99: $d_{min} = 15$, $d_{max} = 60$, $\alpha_{max} = 50$, $D_{IPAN} = 50$;
(b) algorithm parameters IPAN99: $d_{min} = 15$, $d_{max} = 50$, $\alpha_{max} = 80$, $D_{IPAN} = 65$;
(c) algorithm parameters IPAN99: $d_{min} = 15$, $d_{max} = 60$, $\alpha_{max} = 90$, $D_{IPAN} = 65$;

On the basis of characteristic points the similarity of signatures were designated. As a measure of similarity the two-dimensional linear correlation of strings *X, Y* was chosen, known in the literature as Persons R^2 coefficient [6]. This linear regression requires the use of strings of equal number of elements. In the case of signatures, this condition is not fulfilled. This arises from the variety of personal signatures submitted often under different conditions and with an excess or lack of elements constituting the same signature. One of the ways to match two various time strings technique is time Dynamic Time Warping - DTW [9]

Preliminary analysis of functioning the algorithm shows that the calculation complexity of algorithm depends both on the number of points that define the curve (and therefore the quality of the signature) and the control parameters of algorithm IPAN99. In the final outcome it affects the result of the signature identification, although in this case the relationship is difficult to define.

To determine the nature of those relationship preliminary studies was carried out. Their aim was to determine the effects of recognition as dependency of the parameters of algorithm IPAN, and therefore properties of function

$$Q_4 = f(d_{min}, d_{max}, \alpha_{max}, D_{IPAN})$$

which corresponds to R^2 values of algorithm evaluation.

4. TESTING

Practical tests on a method for testing the similarity of signatures were carried out at 150 signatures submitted by 15 people, 10 signatures each. Then all of the signatures were compared using scheme each to each. The total number of comparisons was 675 and for each measurement the Persons coefficient R^2 was calculated. All signatures of the persons have been so selected that the similarity between them was not less than 0.95. Then, the similarity of signatures were calculated taking into account various parameters IPAN99 algorithm. Algorithm parameters have been changed as follows: α_{max} – from 10° to 180° (with 10° increment), d_{min} – from 0 to 20 (with 1 increment), d_{max} – from 20 to 40 (with 1 increment), D_{IPAN} – from 45 to 95 (with 5 increment). For each set of features the similarities of signatures were rescheduled. The results grouped in 10 classes depending on the value of coefficient R^2 are presented in Table 1.

Table 1 The results grouped in 10 classes

	min	– max
Class 1	0.95	1.0
Class 2	0.9	0.94
Class 3	0.85	0.89
Class 4	0.8	0.84
Class 5	0.7	0.79
Class 6	0.6	0.69
Class 7	0.5	0.59
Class 8	0.4	0.49
Class 9	0.2	0.39
Class 10	0.0	0.19

The amount of each class was analyzed taking into account the various parameters of the algorithm. Results for the first three classes are presented in Tables 2-5.

Table 2 Influence of α_{max} parameter on the amount of each class.

	Parameter α_{max}				
	30	**40**	**50**	**60**	**70**
Class 1	10	146	29	17	8
Class 2	117	162	204	172	98
Class 3	125	126	138	125	193

Table 3 Influence of d_{min} parameter on the amount of each class.

	Parameter d_{min}				
	10	**11**	**12**	**13**	**14**
Class 1	23	6	98	113	67
Class 2	44	81	146	162	121
Class 3	154	216	165	173	169

Table 4 Influence of d_{max} parameter on the amount of each class.

	Parameter d_{max}				
	19	**23**	**24**	**25**	**26**
Class 1	52	90	71	32	29
Class 2	144	107	95	91	52
Class 3	70	69	80	80	48

Table 5 Influence of D_{IPAN} parameter on the amount of each class.

	Parameter D_{IPAN}				
	45	**50**	**55**	**60**	**65**
Class 1	0	4	61	158	39
Class 2	136	30	217	169	86
Class 3	87	49	86	73	14

When selecting parameters of IPAN99 algorithm one should examine the amount of Class 1. Class 1 includes the signatures whose the values of R^2 are in the range 0.95 ÷ 1.0, which is the same as results obtained without the algorithm IPAN99. By analyzing both results we can say that the parameters with great influence on IPAN99 algorithm are: α_{max}, d_{min} and D_{IPAN}. In the case of a parameter d_{max} it is difficult to determine the optimum value.

Few precise impact of d_{max} parameter allows (at least at first approach) to simplify our Q_4 function by reduction to three dimensions, and therefore to function

$$Q_3 = f(d_{min}, \alpha_{max}, D_{IPAN})$$

Our goal is to automate the process of recognition using algorithm IPAN99 by the selection of optimum values of control parameters. As a measure of the effectiveness we will use reduced function Q_3; its optimum will be determined by discrete version of the algorithm [1].

Analyzing the obtained data allow to suppose that the optimal values of the parameters IPAN99 algorithm are in surroundings: α_{max} - 400, d_{min} - 13, d_{max} - 90, D_{IPAN} - 60.

5. CONCLUSIONS

Analysis of data from the tables suggests that the function Q_4 represents a compound surface and belongs to a class of multimodale functions[1,5]. Analysis of such functions and in particular to find their global maximum (or minimum) is very difficult, strongly growing with the increase in the number of dimensions.

There are many methods of optimization of unimodale functions, finding their minimum, but the assortment of methods draws from a variety of specific functions properties. Most of these methods [11] belong to a class of iterating methods using locale function properties (gradient, a decline, the locale minimum) to designate further proceedings; the stop condition determines the point of achieved optimum.

Unfortunately, none of these methods is not universal in the sense that they can find the optimum any unimodale function.

Even more difficult problem is finding the optimum of non-unimodale function. In the case of functions in discrete spaces, the classic algorithms – based on gradient methods– fail. The paper [1] presents an algorithm which is a modification Hooke-Jeeves algorithm [5], which can be adapted to the discrete space. Final version of this algorithm for the two-dimensional (but continuous) space showed good efficiency, both in terms of accuracy of optimum location, and time-consuming of calculations.

Our studies determine that there is the possibility of identifying signatures based on the reduced length of the signature characteristics vector. In subsequent stages of work we plan analysis signatures using other methods of finding of characteristic points as well as testing effectiveness of the IPAN99 algorithm.

The purpose of future work is also developing methods for optimization of multimodale functions.

BIBLIOGRAPHY

[1] BARGIELSKI M.J., Organizacja poszukiwania optymalnych parametrów modelu matematycznego, ZN Pol.Sl. Gliwice 1971, seria Automatyka, z. 17.

[2] DOROZ R., WIDUCH S., Signature characteristic points determination by means of the IPAN99 algorithm., Medical Informatics i Technologies, 2007, pp. 105-113.

[3] DOROZ R., WIDUCH S., Estimation of signature similarity by means of IPAN99 algorithm. XXXVIII International Autumn Colloquium ASIS 2006 Advanced Simulation of System. Ostrawa. Czech Republic. 2006 pp. 85-90

[4] FERRER M., ALONSO J., TRAVIESO C., Offline Geometric Parameters for Automatic Signature Verification Using Fixed-Point Arithmetic, 2005, Pattern Recognition and Artificial Intelligence pp. 993-997.

[5] HOOKE R., JEEVES T.A., Direct search solution of numerical and statistical problem. Journal of ACM 8,2 (Apr. 1961)

[6] LEI H. PALLA S., GOVINDARAJU V., ER2 : an Intuitive Similarity Measure for ON-line Signature VerificationFrontiers in Handwriting Recognition. 2004. IWFHR-9 2004. Ninth International Workshop, pp. 191-195.

[7] PARISSE C., Global Word Shape Processing in Off-Line Recognition of Handwritting. IEEE Transactions on Pattern Analysis and Machine Intelligence. 1996. Vol. 18, no. 5, pp. 460-464.

[8] 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.

[9] RATH T. M., MANMATHA R., Word Image Matching using Dynamic Time Warping, IEEE Conference on Computer Vision and Pattern Recognition, vol. 2, Madison, Wisconsin., 2003, pp. 521-527

[10] SCHEIDAT, T., VIELHAUER, C., DITTMANN, J., Distance-Level Fusion Strategies for Online Signature Verification, Multimedia and Expo, 2005. ICME 2005. IEEE International Conference onVolume, pp. 1294 – 1297.

[11] WILDE D.J., Optimum seeking methods. Prentice-Hall Inc., Engleqood Cliffs N.J, 1964.

[12] WIROTIUS M, RAMEL J.Y., VINCENT N., Distance and Matching for Authentication by On-Line Signature, Fourth IEEE Workshop on Automatic Identification Advanced Technologies (AutoID'05), pp. 230-235.