New Algorithm for On-line Signature Verification Using Characteristic Global Features

Marcin Zalasiński

Abstract In this paper we propose a new algorithm for on-line signature verification using characteristic global features values. It is based on so-called global features which describe characteristic attributes of the signature, e.g. time of signing process, number of pen-ups, average velocity of the pen etc. Our method assumes evaluation of the global features for the individual and selection of the most characteristic ones, which are used during classification phase (verification of the signature). Classification is performed using specially designed flexible neuro-fuzzy one class classifier.

Keywords Behavioural biometrics · Dynamic signature verification · Global features of the signature · Flexible fuzzy one-class classifier

1 Introduction

On-line signature is a behavioural biometric attribute used for an identity verification. It is acquired using digital input device and it contains many information about dynamics of the signing process.

Approaches used to the dynamic signature verification can be divided into few main groups (see e.g. [1, 2]). In this paper we focus on the approach based on so-called global features, which are extracted from signature and used during training and classification phase. We use a set of global features proposed in [3]. It should be noted that the proposed fast algorithm is not dependent on the initial feature set, which can be reduced or extended.

In this paper we propose a new algorithm for on-line signature verification, which selects the most characteristic global features of the individual. The method determines for each user weights of importance of features. Next, it selects the most

M. Zalasiński (✉)
Institute of Computational Intelligence, Częstochowa University of Technology, Częstochowa, Poland
e-mail: marcin.zalasinski@iisi.pcz.pl

Z. Wilimowska et al. (eds.), *Information Systems Architecture and Technology: Proceedings of 36th International Conference on Information Systems Architecture and Technology – ISAT 2015 – Part IV*, Advances in Intelligent Systems and Computing 432, DOI 10.1007/978-3-319-28567-2_12

characteristic ones, which are used during classification process. Global feature selection is used to: (a) elimination of features that may have a negative impact on the verification accuracy, (b) simplification of verification process and increasing the interpretability of used fuzzy system, (c) obtaining additional information about specifics of template signatures of each user (which can be e.g. processed to obtain information about certain psychological characteristics). For the purposes of the proposed method, we have developed a new fuzzy one-class classifier, proposed by us earlier (see e.g. [1, 4]). It does not require supervised learning and so-called skilled forgeries (forged signatures) to proper work.

To test the proposed method we used the BioSecure Database (BMDB) distributed by the BioSecure Association (see [5]) which is admitted source of data used in this field.

This paper is organized into four sections. In Sect. 2 we present description of the new method for dynamic signature verification based on global features. In Sect. 3 simulation results are presented. Conclusions are drawn in Sect. 4.

2 Description of the New Method for Dynamic Signature Verification Based on Global Features

General description of the fast training phase for the user i (procedure `Training(i)`) can be described as follows (see Fig. 1). **Step 1.** Acquisition of J training signatures of user i. **Step 2.** Determination of matrix $\mathbf{G}_i$ of all considered global features, describing dynamics of signatures, for all available J training signatures of user i. **Step 3.** Determination of vector $\bar{\mathbf{g}}_i$ of average values for each global feature, obtained in Step 2 for J training signatures of user i. **Step 4.** Determination of weights of importance $w_{i,n}$ for global feature n of user i. **Step 5.** Selection of N' the most characteristic global features of the user i and creation of reduced matrix $\mathbf{G}'_i$ and reduced vector $\bar{\mathbf{g}}'_i$, which contain only information about selected features. **Step 6.** Selection of classifier parameters used in the test phase (procedure `Classifier Determination` $(i, \mathbf{G}'_i, \bar{\mathbf{g}}'_i)$). **Step 7.** Storing in a database the following information about user i: vector $\bar{\mathbf{g}}'_i$, parameters of classifier $maxd_{i,n}$ and $w'_{i,n}(n = 1, 2, \ldots, N')$. Detailed description the procedure `Training(i)` is presented below.

In the **Step 2** Matrix $\mathbf{G}_i$ is determined. It contains all considered global features of all J training signatures of user i and it has the following structure:

$$\mathbf{G}_i = \begin{bmatrix} g_{i,1,1} & g_{i,2,1} & \cdots & g_{i,N,1} \\ g_{i,1,2} & g_{i,2,2} & \cdots & g_{i,N,2} \\ & & \vdots & \\ g_{i,1,J} & g_{i,2,J} & \cdots & g_{i,N,J} \end{bmatrix} = \begin{bmatrix} \mathbf{g}_{i,1} \\ \mathbf{g}_{i,2} \\ \vdots \\ \mathbf{g}_{i,N} \end{bmatrix}^T, \tag{1}$$

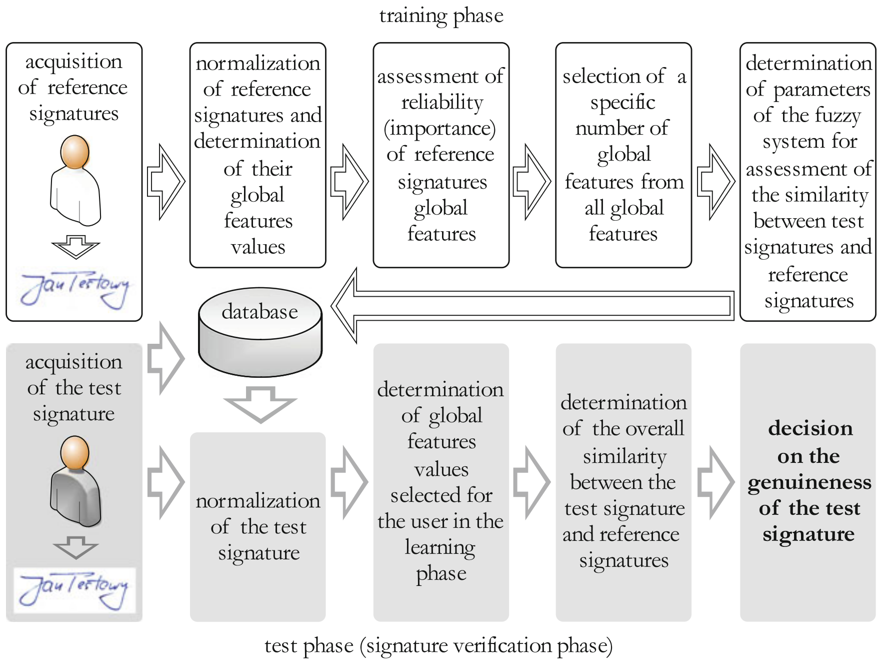

Fig. 1 Idea of the proposed algorithm for on-line signature verification based on selection of the most characteristic set of global features (realized individually for each user)

where $\mathbf{g}_{i,n} = [g_{i,n,1} \quad g_{i,n,2} \quad \ldots \quad g_{i,n,J}]$, $g_{i,n,j}$ is a value of the global feature n, $n = 1, 2, \ldots, N$, determined for the signature j, $j = 1, 2, \ldots, J$, created by the user i, $i = 1, 2, \ldots, I$, i is a number of the users, J is a number of the signatures created by the user in the acquisition phase, N is a number of the global features. As already mentioned, the detailed method of determining each of the considered features is described in [3].

Matrix $\mathbf{G}_i$ is used to determine value of the vector $\bar{\mathbf{g}}_i$ in the **Step 3**. Vector $\bar{\mathbf{g}}_i$ of average values of each global feature of all training signatures J of user i is described as follows:

$$\bar{\mathbf{g}}_i = [\bar{g}_{i,1}, \bar{g}_{i,2}, \ldots, \bar{g}_{i,N}], \tag{2}$$

where $\bar{g}_{i,n}$ is average value of nth global feature of training signatures of user i, computed using the following formula:

$$\bar{g}_{i,n} = \frac{1}{J} \sum_{j=1}^{J} g_{i,n,j}. \tag{3}$$

Next, in the **Step 4**, weights of importance of all considered global features are determined. Weight of nth global feature of the user i is computed on the basis of standard deviation of nth global feature of the user i and average value of distances between the feature and its mean. This process is described by the following formula:

$$w_{i,n} = 1 - \frac{\sqrt{\frac{1}{J}\sum_{j=1}^{J}\left(\bar{g}_{i,n} - g_{i,n,j}\right)^2}}{\frac{1}{J}\sum_{j=1}^{J}\left|\bar{g}_{i,n} - g_{i,n,j}\right|}. \quad (4)$$

After this process in the **Step 5** N' the most characteristic global features are selected (see Fig. 1). They are features whose weights values are the highest. Next, reduced matrix $\mathbf{G}'_i$ and reduced vector $\bar{\mathbf{g}}'_i$ are determined. They are created taking into account the only N' the most characteristic features. Moreover, weights of the most characteristic global features are denoted as $w'_{i,n}(n = 1, 2, \ldots, N')$.

Determination of the classifier (**Step 6**) and its parameters are described in next subsection.

2.1 Determination of Classifier

In the procedure `Classifier Determination`$(i, \mathbf{G}'_i, \bar{\mathbf{g}}'_i)$ described in this section the most characteristic global features are considered.

In the **Step 1** maximum distances $maxd_{i,n}$ between each characteristic global feature n and average value of the global feature for all J signatures of user i is computed using the following formula:

$$maxd_{i,n} = \max_{j=1,\ldots,J}\left\{\left|\bar{g}_{i,n} - g_{i,n,j}\right|\right\}. \quad (5)$$

Please note that distance $maxd_{i,n}$ is associated with the global feature n of the user i and determines instability of the signature in the context of the feature n. Value of the distance $maxd_{i,n}$ is also dependent on the variability of feature and it has an impact on the work of the signature classifier (see Fig. 2).

Next, a classifier is created (**Step 2**). We use flexible neuro-fuzzy system of the Mamdani type (see e.g. [6–9]). This system is based on the rules in the if-then form. The fuzzy rules contain fuzzy sets which represent the values, e.g. “low” and “high”, of the input and output linguistic variables. In our method the input linguistic variables are dependent on the similarity between the global features of test signature and average values of global features computed on the basis of training signatures. The system uses N' features. Output linguistic variables describe the reliability of the signature. In our method parameters of input fuzzy sets are

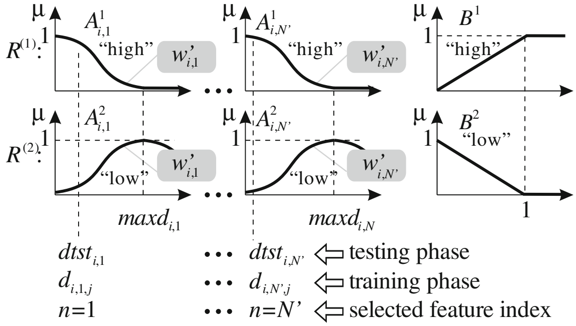

Fig. 2 Input and output fuzzy sets of the flexible neuro-fuzzy system of the Mamdani type for verification signature of user *i*

individually selected for each user. Please note that if training signatures are more similar to each other, the tolerance of our classifier is lower ($maxd_{i,n}$ takes smaller values).

The flexibility of the classifier results from the possibility of using in the classification the importance of global features, which are selected individually for each user. Taking into account the weights of importance of the global features is possible thanks to the use of proposed by us earlier (see e.g. [7, 10, 11]) aggregation operators named the weighted triangular norms.

Our system for the signature verification works on the basis of two fuzzy rules presented as follows:

$$\begin{cases} R^{(1)}: \left[\begin{array}{c} \text{IF}\left(dtst_{i,1}\text{is}A^1_{i,1}\right)\middle|w'_{i,1}\text{AND IF}\left(dtst_{i,2}\text{is}A^1_{i,2}\right)\middle|w'_{i,2}\text{AND}\ldots \\ \text{IF}\left(dtst_{i,N'}\text{is}A^1_{i,N'}\right)\middle|w'_{i,N'}\text{THEN}y_i\text{is}B^1 \end{array} \right] \\ R^{(2)}: \left[\begin{array}{c} \text{IF}\left(dtst_{i,1}\text{is}A^2_{i,1}\right)\middle|w'_{i,1}\text{AND IF}\left(dtst_{i,2}\text{is}A^2_{i,2}\right)\middle|w'_{i,2}\text{AND}\ldots \\ \text{IF}\left(dtst_{i,N'}\text{is}A^2_{i,N'}\right)\middle|w'_{i,N'}\text{THEN}y_i\text{is}B^2 \end{array} \right] \end{cases}, \tag{6}$$

where

- $dtst_{i,n}, i = 1,2,\ldots,I, n = 1,2,\ldots,N', j = 1,2,\ldots,J$, are input linguistic variables in the system for the signature verification.
- $A^1_{i,n}, A^2_{i,n}, i = 1,2,\ldots,I, n = 1,2,\ldots,N'$, are input fuzzy sets related to the global feature number n of the user i represent values "high" assumed by input linguistic variables. Analogously, fuzzy sets $A^2_{i,1}, A^2_{i,2},\ldots,A^2_{i,N'}$ represent values "low" assumed by input linguistic variables. Thus, each rule contains N' antecedents. In the fuzzy classifier of the signature used in the simulations we applied a Gaussian membership function (see Fig. 2) for all input fuzzy sets.
- $y_i, i = 1,2,\ldots,I$, is output linguistic variable interpreted as reliability of signature considered to be created by the ith signer.
- B^1, B^2 are output fuzzy sets shown in Fig. 2. Fuzzy set B^1 represents value "high" of output linguistic variable. Analogously, fuzzy set B^2 represents value

"low" of output linguistic variable. In the fuzzy classifier of the signature used in the simulations we applied the membership function of type γ (see e.g. [12]) in the rule 1. This membership function is defined as follows:

$$\mu_{B^1}(x) = \begin{cases} 0 & \text{for} \quad x \le a \\ \frac{x-a}{b-a} & \text{for} \quad a < x \le b \\ 1 & \text{for} \quad x > b \end{cases}. \tag{7}$$

In the rule 2 we applied the membership function of type L (see e.g. [12]). This membership function is defined as follows:

$$\mu_{B^2}(x) = \begin{cases} 1 & \text{for} \quad x \le a \\ \frac{b-x}{b-a} & \text{for} \quad a < x \le b \\ 0 & \text{for} \quad x > b \end{cases}. \tag{8}$$

In our system value of the parameter a for both rules from the rule base (6) is equal to 0 and value of the parameter b is equal to 1.

- $maxd_{i,n}, i = 1, 2, \ldots, I, n = 1, 2, \ldots, N'$, can be equated with the border values of features of individual users [see formula (5)].
- $w_{i,n}, i = 1, 2, \ldots, I, n = 1, 2, \ldots, N'$, are weights of importance related to the global feature number n of the user i [see formula (4)].

Please note that regardless of the set of features chosen individually for the user, the interpretation of the input and output fuzzy sets is uniform. Moreover, the way of the signature classification is interpretable (see [13]).

2.2 *Identity Verification Phase*

Formal notation of the process of signature verification (`Signature Verification(`i`)`) is performed in the following way (see Fig. 1): **Step 1.** Acquisition of one test signature of the user which is considered as user i. **Step 2.** Download of information about average values of the most characteristic global features of the user i computed during training phase—$\bar{\mathbf{g}}'_i$ and classifier parameters of the user i from the database -$maxd_{i,n}, w'_{i,n}(n = 1, 2, \ldots, N')$. **Step 3.** Determination of values of global features which have been selected as the most characteristic for user i in training phase. **Step 4.** Verification of test signature using of one class flexible neuro-fuzzy classifier.

The test signature is considered to be true if the following assumption is satisfied:

$$\bar{y}_i = \frac{T^*\left\{\mu_{A^1_{i,1}}\left(dtst_{i,1}\right),\ldots,\mu_{A^1_{i,N'}}\left(dtst_{i,N'}\right); w'_{i,1},\ldots,w'_{i,N'}\right\}}{\left(\begin{array}{c} T^*\left\{\mu_{A^1_{i,1}}\left(dtst_{i,1}\right),\ldots,\mu_{A^1_{i,N'}}\left(dtst_{i,N'}\right); w'_{i,1},\ldots,w'_{i,N'}\right\}+ \\ T^*\left\{\mu_{A^2_{i,1}}\left(dtst_{i,1}\right),\ldots,\mu_{A^2_{i,N'}}\left(dtst_{i,N'}\right); w'_{i,1},\ldots,w'_{i,N'}\right\}\end{array}\right)} > cth_i, \qquad (9)$$

where

– $T^*\{\cdot\}$ is the algebraic weighted t-norm (see [7, 10, 14, 15]) in the form:

$$\begin{array}{l} T^*\left\{\begin{array}{c} a_1, a_2; \\ w_1, w_2 \end{array}\right\} = T\left\{\begin{array}{c} 1-w_1\cdot(1-a_1), \\ 1-w_2\cdot(1-a_2) \end{array}\right\} \\ \overset{\text{e.g.}}{=}(1-w_1\cdot(1-a_1))\cdot(1-w_2\cdot(1-a_2)) \end{array}, \qquad (10)$$

where t-norm $T^*\{\cdot\}$ is a generalization of the usual two-valued logical conjunction (studied in classical logic), w_1 and $w_2 \in [0,1]$ mean weights of importance of the arguments $a_1, a_2 \in [0,1]$. Please note that $T^*\{a_1, a_2; 1, 1\} = T\{a_1, a_2\}$ and $T^*\{a_1, a_2; 1, 0\} = a_1$.

- $\mu_A(\cdot)$ is a Gaussian membership function (see e.g. [12]).
- $\mu_{B^1}(\cdot)$ is a membership function of class L (see e.g. [12]).
- $\mu_{B^2}(\cdot)$ is a membership function of class γ (see e.g. [12]).
- $\bar{y}_i, i = 1, 2, \ldots, I$, is the value of the output signal of applied neuro-fuzzy system described by rules (6).
- $cth_i \in [0,1]$ is a coefficient determined experimentally for each user to eliminate disproportion between FAR and FRR error (see e.g. [16]).

Formula (9) was created by taking into account in the description of system simplification resulting from the spacing of fuzzy sets, shown in Fig. 2. The simplifications are as follows: $\mu_{B^1}(0) = 0, \mu_{B^1}(1) \approx 1, \mu_{B^2}(0) \approx 1, \mu_{B^2}(1) = 0$.

3 Simulations

Simulations were performed using the commercial BioSecure DS2 Signature database, which contains signatures of 210 users, and an authorial test environment. The signatures was acquired in two sessions using the digitizing graphic tablet. Each session contains 15 genuine signatures and 10 skilled forgeries per person. During training phase we used 5 randomly selected genuine signatures of each signer. During test phase we used 10 remaining genuine signatures and all 10 skilled forgeries of each signer. The process was performed five times, and the results were averaged. In the simulations we used a set of 85 features described in

Table 1 Comparison of the results for the dynamic signature verification methods for the database BioSecure

Method	Average FAR (%)	Average FRR (%)	Average error (%)
Methods of other authors [20]	–	–	3.48–30.13
Evolutionary selection with PCA [21]	5.29	6.01	5.65
Evolutionary selection [17]	2.32	2.48	2.40
Our method	3.02	3.26	3.14

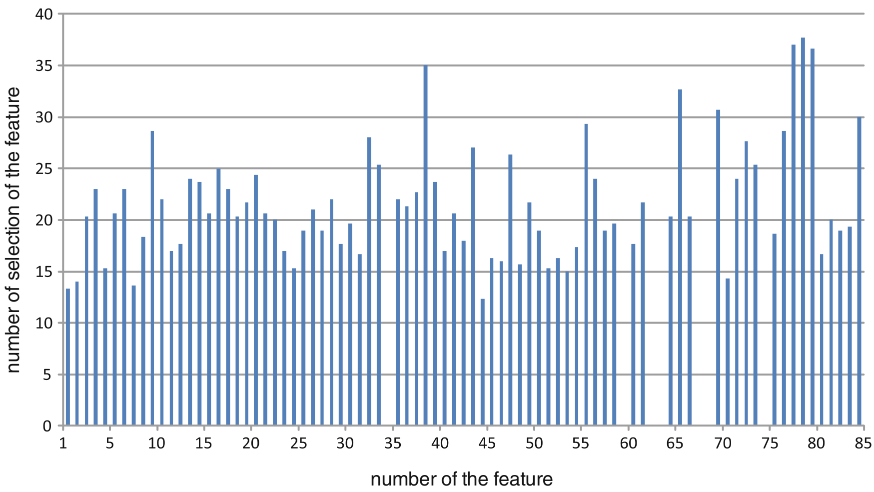

Fig. 3 Number of selection of global features averaged for one test session

[17]. A purpose of the algorithm was an automatic selection of 8 (about 10 %) the most characteristic features for the individual.

Table 1 contains a set of accuracies obtained by the proposed method, our previously developed methods and the ones proposed by other authors. The table contains values of FAR (False Acceptance Rate) and FRR (False Rejection Rate) errors which are commonly used in the literature to evaluate the effectiveness of identity verification methods (see e.g. [18, 19]).

It may be seen that the proposed method works with a very good accuracy for the BioSecure database taking into account all methods considered in the Table 1. It works a little worse than the evolutionary selection method but its complexity is lower and working time is shorter.

Moreover, in Fig. 3 we present a number of selection of global features averaged for one test session. The most often selected features are the ones denoted by numbers 79, 78 and 80 respectively. They are features associated with time moment of maximum jerk of the signature trajectory, time moment of maximum velocity of the signature and duration of writing process. Feature number 39 was also selected many times, it is associated with value of local maximum of signal x, value of signal x during first pen-down and so-called delta x, which is a measure of the signature range. It is worth to note that some features were not selected.

4 Conclusions

In this paper we propose a new method for the dynamic signature verification based on the most characteristic global features. The method evaluates each global feature of the individual and selects a number of his/her the most characteristic features which are used during verification process. It uses a dedicated flexible fuzzy one-class classifier. Accuracy of the method has been tested using authorial test environment implemented in C# and commercial BioSecure on-line signature database. The proposed algorithm worked with a very good accuracy in comparison to other methods and it is distinguished by a low complexity.

Acknowledgments The authors would like to thank the reviewers for very helpful suggestions and comments in the revision process. The project was financed by the National Science Centre (Poland) on the basis of the decision number DEC-2012/05/B/ST7/02138.

References

1. Cpałka, K., Zalasiński, M.: On-line signature verification using vertical signature partitioning. Expert Syst. Appl. **41**, 4170–4180 (2014)
2. Cpałka, K., Zalasiński, M., Rutkowski, L.: New method for the on-line signature verification based on horizontal partitioning. Pattern Recogn. **47**, 2652–2661 (2014)
3. Fierrez-Aguilar, J., Nanni, L., Lopez-Penalba, J., Ortega-Garcia, J., Maltoni, D.: An On-line signature verification system based on fusion of local and global information. Lecture Notes in Computer Science, Audio- and Video-based Biometric Person Authentication vol. **3546**, pp. 523–532 (2005)
4. Zalasiński, M., Cpałka, K.: New approach for the on-line signature verification based on method of horizontal partitioning. Lect. Notes Comput. Sci. **7895**, 342–350 (2013)
5. Homepage of Association BioSecure. [Online] Available from: http://biosecure.it-sudparis.eu [Accessed: 3 June 2015]
6. Cpałka, K.: A new method for design and reduction of neuro-fuzzy classification systems. IEEE Trans. on Neural Networks **20**, 701–714 (2009)
7. Cpałka, K.: On evolutionary designing and learning of flexible neuro-fuzzy structures for nonlinear classification. Nonlinear Anal. Ser. A: Theor. Methods Appl. **71**, 1659–1672 (2009)
8. Cpałka, K., Łapa, K., Przybył, A., Zalasiński, M.: A new method for designing neuro-fuzzy systems for nonlinear modelling with interpretability aspects. Neurocomputing **135**, 203–217 (2014)
9. Cpałka, K., Rebrova, O., Nowicki, R., Rutkowski, L.: On design of flexible neuro-fuzzy systems for nonlinear modelling. Int. J. Gen. Syst. **42**, 706–720 (2013)
10. Rutkowski, L., Cpałka, K.: Flexible neuro-fuzzy systems. IEEE Trans. on Neural Networks **14**, 554–574 (2003)

11. Rutkowski, L., Przybył, A., Cpałka, K.: Novel online speed profile generation for industrial machine tool based on flexible neuro-fuzzy approximation. IEEE Trans. Ind. Electron. **59**, 1238–1247 (2012)
12. Rutkowski, L.: Computational Intelligence. Springer, Berlin (2008)
13. Gacto, M.J., Alcala, R., Herrera, F.: Interpretability of linguistic fuzzy rule-based systems: an overview of interpretability measures. Inf. Sci. **181**, 4340–4360 (2011)
14. L. Rutkowski, K. Cpałka, K., Designing and learning of adjustable quasi triangular norms with applications to neuro-fuzzy systems. IEEE Trans. Fuzzy Syst. **13**, 140–151 (2005)
15. Zalasiński, M., Cpałka, K.: Novel algorithm for the on-line signature verification. Lect. Notes Comput. Sci. **7268**, 362–367 (2012)
16. Yeung, D.Y., Chang, H., Xiong, Y., George, S., Kashi, R., Matsumoto, T., Rigoll, G.: SVC2004: first international signature verification competition. Lect. Notes Comput. Sci. **3072**, 16–22 (2004)
17. Zalasiński, M., Cpałka, K., Hayashi, Y.: New method for dynamic signature verification based on global features. Lect. Notes Comput. Sci. **8468**, 231–245 (2014)
18. Faundez-Zanuy, M.: On-line signature recognition based on VQ-DTW. Pattern Recogn. **40**, 981–992 (2007)
19. Kholmatov, A., Yanikoglu, B.: Identity authentication using improved online signature verification method. Pattern Recogn. Lett. **26**, 2400–2408 (2005)
20. Houmani, N., Garcia-Salicetti, S., Mayoue, A., Dorizzi, B.: BioSecure signature evaluation campaign 2009 (BSEC'2009): Results (2009). [Online] Available from: http://biometrics.it-sudparis.eu/BSEC2009/downloads/BSEC2009/_results.pdf Accessed 3 June 2015
21. Zalasiński, M., Łapa, K., Cpałka, K.: New algorithm for evolutionary selection of the dynamic signature global features. Lect. Notes Comput. Sci. **7895**, 113–121 (2013)