New Algorithm for Evolutionary Selection of the Dynamic Signature Global Features

Marcin Zalasiński, Krystian Łapa, and Krzysztof Cpałka

Częstochowa University of Technology,
Institute of Computational Intelligence, Poland
{marcin.zalasinski,krystian.lapa,krzysztof.cpalka}@iisi.pcz.pl

Abstract. Methods using dynamic signature for identity verification may be divided into three main categories: global methods, local function based methods and regional function based methods. Global methods base on a set of global parametric features, which are extracted from signature of user. Global feature extraction methods have been often presented in the literature. Another interesting task is selection of a features group which will be considered individually for each user during training and verification process. In this paper we propose a new approach to automatic evolutionary selection of the dynamic signature global features. Our method was tested with use of the SVC2004 public on-line signature database.

1 Introduction

Biometrics is the science of recognizing the identity of a person based on some kind of unique personal attributes. The attributes may be physical like a fingerprint, a face (see e.g. [14], [25]-[26]) or behavioural like a gait, a signature (see e.g. [8], [30]-[31]) etc. Dynamic signature is a kind of signature biometric attribute, which contains not only information about trajectory, but also information about the dynamics of the signature. Signature may be treated as a unique identifier of the user, used during identity verification process (see e.g. [10], [30]-[31]).

The dynamic signature verification methods may be categorized into three main groups (see e.g. [22]): global methods, local function based methods and regional function based methods. In global methods a set of features are extracted from the signature (e.g. signature total duration, number of pen-downs) and classification process is performed on the basis of these features. Local function based methods use the time functions of the signatures, which are matched and compared. One of the methods used for this purpose is Dynamic Time Warping (see e.g. [28]). In regional function based approaches some regional properties are estimated from the signatures and then they are used to train classifier.

In this paper we focus on the global approach. In the literature a few sets of features have been proposed (see [18], [23], [24]). Some features have been taken from the literature relating to identity verification based on signature but other have been created by authors on the basis of their experience in the signature

L. Rutkowski et al. (Eds.): ICAISC 2013, Part II, LNAI 7895, pp. 113–121, 2013.

acquisition procedure. In [9] the authors have collected all features from [18], [23], [24], some of them have been adapted and some new features have been add to this set. In this paper we base on the feature set proposed in [9]. It should be noted that the operation of our method is not dependent on the adopted feature set. The feature set can be practically arbitrarily reduced or extended.

Large global feature set may be reduced by selection of optimal features subset, which will be considered during classification phase. In this case global features may be ranked and only features with the highest rank value are used in classification process. Approaches to the rate of global features have been presented in [9] and [18]. In [18] authors propose approach based on computation of rank using mean value and standard deviation of the feature. Rank for each feature is calculated separately. In [9] the method based on the distance between the mean value of the feature from training signatures of the user and the feature value of all training signatures from all users is presented. This method performs the overall ranking of features, without detailing the importance of features for the user.

In this paper we propose a new method for selection of the dynamic signature global features. The method is based on a modified genetic algorithm which is used to select of features set used in the verification process. Idea of genetic algorithms has been taken from the natural process in which live organisms transfer their features to their offspring during creation of new population. The algorithms may be applied to solve various optimization problems. In the method presented in this paper features are selected individually for each signer. Our algorithm is characterized by a special way of determining the fitness function. It bases on classification with use of the Principal Component Analysis method (see e.g. [19]) and so called reconstruction error.

This paper is organized into four sections. In Section 2 we present novel algorithm for evolutionary selection of the dynamic signature global features. In Section 3 simulation results are presented. Conclusions are drawn in Section 4.

2 Idea of the New Method of Features Selection

Method proposed in this paper uses combination of genetic algorithm (see e.g. [2]-[7], [11], [16], [17]) and PCA method for selection of global features subset used during verification process. The selection process is performed individually for each user, so each subset of features is unique for the signer. Subset of features is chosen from all one hundred features presented in [9]. Examples of features which belong to the set of features in [9] are shown in Table 1.

Block diagram of the proposed algorithm is presented in Fig. 1 a). The algorithm selects the subset of global features and creates classifier for the i-th user. First, selection of training signatures for i-th user is performed (**block 1**). The next random selection of the other users signatures is performed (**block 2**). This signatures are treated as forgeries. This approach commonly appears in the literature in the field of identity verification based on the dynamic signature (see e.g. [29]). In the third step, selection of the optimal global features

Table 1. Examples of features used in [9]

Feature description	**Symbol**
signature total duration	T_s
number of pen-ups	N(pen-ups)
standard deviation of y-trajectory acceleration	σ_{a_y}
standard deviation of y-trajectory velocity	σ_{v_y}
mean value of velocity divided by maximum value of velocity	$\bar{v}/v_{\max}$

subset for the i-th user is performed (**block 3**). For this purpose the algorithm *GeneticFeaturesSelection()* described in Section 2.1 and presented in Fig. 1 b) is used. Next, the algorithm returns the subset of genuine training signatures which are used during creation of the classifier (**block 4**) and subset of global features selected for the i-th user (**block 5**). The data from block 4 and block 5 are used during the verification phase.

2.1 Genetic Features Selection

Schema of genetic features selection is presented in Fig. 1 b). The algorithm selects the chromosome encoding the best subset of global features for the i-th user. In the first step the creation of an initial population is performed, which consists of the random selection of chromosomes (**block 1**). Each chromosome contains information about selected subset of features, each gene in the chromosome represents one global feature. Length of the chromosome corresponds to the number of all features. Each chromosome is represented by a binary sequence. If the value of the gene is "1", the feature which number is equal to the number of a gene's position is selected for the subset of the features represented by the chromosome. Otherwise if the value of gene is "0", the feature which number is equal to the number of a gene's position is not selected for the subset of the features represented by the chromosome. Next, the evaluation of the fitness of chromosomes in a population is performed (**block 2**). During this step value of the fitness function for each chromosome is calculated. It is realized in a separate algorithm, called *CalculateFf()*, presented in Fig. 2 a). In our algorithm the lower value of this function means the better the "quality" of the chromosome. In determination of fitness function value PCA method is used for computation of reconstruction errors. The errors decide about quality of subset of features encoded in the chromosomes for the i-th user. The detailed description of the fitness function used in our algorithm is presented in Section 2.2. Next, the stopping criterion is checked (**block 3**). If the stopping criterion is not met then the next step is selection of chromosomes (**block 4**). The selection of chromosomes consists in selecting, based on the calculated values of the fitness function. Selected chromosomes will take part in the creation of offspring until the next generation. The chromosomes having the best value of the fitness function have the most of the chances for the participation in the creation of new individuals. The selection can be performed using one of the methods

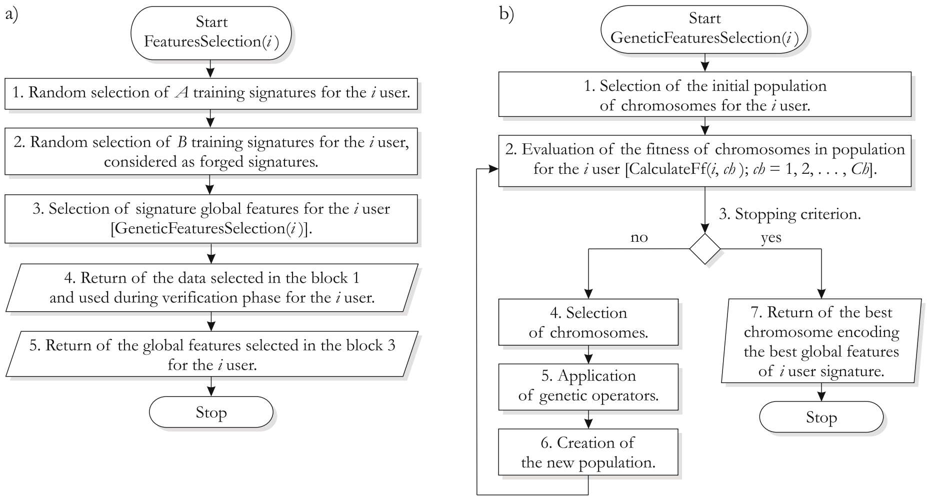

Fig. 1. Block diagram of: a) new method of features selection, b) genetic features selection

available in the literature (see e.g. [27]). Next, application of genetic operators is performed (**block 5**). The operators are applied to the chromosomes selected in 4-th step. It leads to the creation of a new population constituting the offspring population derived from the parents population. In our algorithm standard operators for genetic algorithms are used - crossover and mutation (see e.g. [1], [20]). In the next step creation of a new population is performed (block 6). The chromosomes obtained as a result of the operation of genetic operators belong to a new population and this population becomes the so-called current population for a given generation of the genetic algorithm. Next, stopping criterion is checked (**block 3**). If the stopping condition is met, "best" chromosome, which has the best fitness function value, is presented (**block 7**). In our algorithm the "best" chromosome contains information about subset of global features of the i user which will be considered during verification phase.

2.2 Fitness Function Determination

Schema of the fitness function used in our algorithm is presented in Fig. 2 a). The algorithm calculates fitness of the chromosome. First, selection of C signatures from A training signatures of the i-th user is performed (**block 1**). Next, PCA classifier is created (**block 2**). The classifier is created on the basis of the signatures from the previous step and features encoded in the chromosome for which fitness is calculated. In the next step (**block 3**) reconstruction error for the remaining genuine training signatures $errg_{i,j,ch}$, $i = 1, 2, \ldots, I$, $j = 1, 2, \ldots, A - C$, $ch = 1, 2, \ldots, Ch$, is computed from the following equation:

$$errg_{i,j,ch} = \left\| \mathbf{f}_{i,j,ch} - \left(\mathbf{E}_{i,ch} \mathbf{E}_{i,ch}^{T} \right) \mathbf{f}_{i,j,ch} \right\|, \tag{1}$$

where $(\mathbf{f}_{i,j,ch})_{p\times 1}$, $i = 1, 2, \ldots, I$, $j = 1, 2, \ldots, A-C$, $ch = 1, 2, \ldots, Ch$, contains features encoded in ch-th chromosome for j-th signature of i-th user, $\mathbf{E}_{i,ch}$, $i = 1, 2, \ldots, I$, $ch = 1, 2, \ldots, Ch$, contains the eigenvectors of the covariance matrix created on the basis of the features from C training signatures selected in the block 1. In this step reconstruction error for the forged training signatures $errf_{i,j,ch}$, $i = 1, 2, \ldots, I$, $j = 1, 2, \ldots, B$, $ch = 1, 2, \ldots, Ch$ is also computed analogously. This operation is performed with use of the PCA classifier created in the block 2. Next, selection of maximum value of reconstruction error for genuine training signatures, $Rg_{\max}$, is performed (**block 4a**) and selection of minimum value of reconstruction error for forged training signatures, $Rf_{\min}$, is performed (**block 4b**). In the next step fitness function of the chromosome is determined (**block 5**). Value of this function is calculated from equation:

$$\mathrm{ff}(ch) = \frac{Rgmax_{i,ch}}{Rfmin_{i,ch}}, \tag{2}$$

where ch is chromosome for which fitness function is determined, $Rgmax_{i,ch}$, $i = 1, 2, \ldots, I$, $ch = 1, 2, \ldots, Ch$, is maximum value of reconstruction error for genuine training signatures of i-th user determined by the following equation:

$$Rgmax_{i,ch} = \max\left\{errg_{i,1,ch}, errg_{i,2,ch}, \ldots, errg_{i,M-O,ch}\right\} \tag{3}$$

and $Rfmin_{i,ch}$, $i = 1, 2, \ldots, I$, $ch = 1, 2, \ldots, Ch$, is minimum value of reconstruction error for forged training signatures of i-th user determined by the following equation:

$$Rfmin_{i,ch} = \min\left\{errf_{i,1,ch}, errf_{i,2,ch}, \ldots, errf_{i,N,ch}\right\}. \tag{4}$$

In the last step of this algorithm value of the fitness function for the chromosome is returned (**block 6**).

2.3 Identity Verification Phase

The process of identity verification based on the dynamic signature is presented in Fig. 2 b). First, test signature is acquired by the system (**block 1**). Next, PCA classifier is created (**block 2**). The classifier is created on the basis of features encoded in the best chromosome and C training signatures of the user which identity is verified. Next, verification is performed with use of the one-class PCA classifier (**block 3**). In the last step of the algorithm, the result of identity verification is presented (**block 4**). Identity is verified if the following equation is satisfied:

$$errtest_i \geq cth_i, \tag{5}$$

where $errtest_{i,ch}$ is reconstruction error of test signature computed on the basis of features encoded in the best chromosome from block 7 of *GenuineFeatures-Selection()* function, i is a number of user which identity is verified and cth_i is a parameter determined individually for each user during training phase. Value

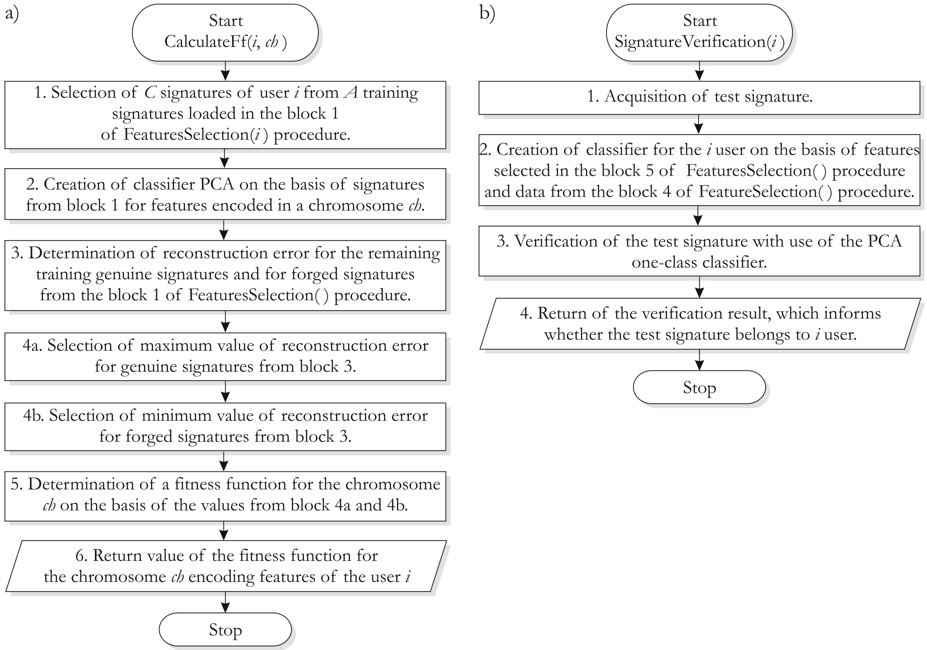

Fig. 2. Block diagram of: a) fitness function, b) signature verification

for this parameter is chosen to minimize the value of FAR and FRR errors and the difference between them.

3 Simulation Results

Simulations were performed using SVC 2004 public database (see [29]). During the simulation the following assumptions have been adopted:

- population contains 100 chromosomes,
- algorithm stops after the lapse of a determined number of 1000 generations,
- during selection of chromosomes tournament selection method is used,
- crossover is performed with probability equal to 0.8 at three points,
- mutation is performed for each gene with probability equal to 0.02
- number of all genuine training signatures $A = 5$, number of forged training signatures $B = 5$, number of genuine training signatures used for training PCA classifier in the *CalculateFf()* function $C = 2$.

The database contains 40 signers and for each signer 20 genuine and 20 forgery signatures. The test was performed five times, every time for all signers stored in the database. During test phase 10 genuine signatures (numbers 11-20) and 20 forgery signatures (numbers 21-40) of each signer were used. Simulations were performed in the authorial environment implemented in C#.

During simulation we tested two methods of verification based on global features. The first one was our method described in this paper. The second was the random subspace method proposed in [12] and used in [21]. The method generates B new training sets of features, builds classifiers on the basis of these sets, and finally combines them into a final decision rule. In our simulations PCA one class classifiers were used.

Results of the simulations are presented in the Table 1. 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. [13]-[15]).

Table 2. Results of simulation performed by our system

Method	**Average FAR**	**Average FRR**	**Average error**
Random subspace method [12]	25.75 %	24.60 %	25.18 %
Our method	**23.87 %**	**22.65 %**	**23.26 %**

Conclusions of the simulations can be summarized as follows:

- The accuracy of our method is higher in comparison to the method described in [12].
- Our method, in contrast to the method given in [12], in the process of classification uses a single set of attributes selected during training phase.
- The advantage of the proposed method is that it allows to characterize the individual user.
- It also allows to extract the features which are often selected as reliable in the context of all users.

4 Conclusions

In this paper a new method for evolutionary selection of the dynamic signature global features is presented. The method assumes selection of the subset of global features from a large set of the features. This process is performed using genetic algorithm with PCA. The features are selected individually for each user. The achieved accuracy of signature verification in comparison with the other method proves correctness of the assumptions. In the future we will seek to determine weights of importance for each feature in context of the user which will be also considered during the verification process.

Acknowledgment. The project was financed by the National Science Center on the basis of the decision number DEC-2011/01/N/ST6/06964.

References

1. Arabas, J.: Lectures on Evolutionary Algorithms. Scientific-Technical Publishing House WNT (2001) (in polish)
2. Cpałka, K., Rutkowski, L.: A new method for designing and reduction of neuro-fuzzy systems. In: IEEE Int. Conference on Fuzzy Systems, pp. 1851–1857 (2006)
3. Cpałka, K.: A method for designing flexible neuro-fuzzy systems. In: Rutkowski, L., Tadeusiewicz, R., Zadeh, L.A., Żurada, J.M. (eds.) ICAISC 2006. LNCS (LNAI), vol. 4029, pp. 212–219. Springer, Heidelberg (2006)
4. Cpałka, K., Rutkowski, L.: Neuro-fuzzy structures for pattern classification. WSEAS Trans. on Computers 697–688 (2005)
5. Cpałka, K., Rutkowski, L.: Neuro-fuzzy systems derived from quasi-triangular norms. In: IEEE International Conference on Fuzzy Systems, vol. 2, pp. 1031–1036 (2004)
6. Cpałka, K.: On evolutionary designing and learning of flexible neuro-fuzzy structures for nonlinear classification. Nonlinear Analysis series A: Theory, Methods & Applications 71 (2009)
7. Dziwiński, P., Rutkowska, D.: Ant focused crawling algorithm. In: Rutkowski, L., Tadeusiewicz, R., Zadeh, L.A., Zurada, J.M. (eds.) ICAISC 2008. LNCS (LNAI), vol. 5097, pp. 1018–1028. Springer, Heidelberg (2008)
8. Ekinci, M., Aykut, M.: Human Gait Recognition Based on Kernel PCA Using Projections. Journal of Computer Science and Technology 22, 867–876 (2007)
9. Fiérrez-Aguilar, J., Nanni, L., Lopez-Peñalba, J., Ortega-Garcia, J., Maltoni, D.: An On-Line Signature Verification System Based on Fusion of Local and Global Information. In: Kanade, T., Jain, A., Ratha, N.K. (eds.) AVBPA 2005. LNCS, vol. 3546, pp. 523–532. Springer, Heidelberg (2005)
10. Fierrez, J., Ortega-Garcia, J.: On-line signature verification. In: Jain, A.K., Ross, A., Flynn, P. (eds.) Handbook of Biometrics (2008)
11. Gabryel, M., Rutkowski, L.: Evolutionary methods for designing neuro-fuzzy modular systems combined by bagging algorithm. In: Rutkowski, L., Tadeusiewicz, R., Zadeh, L.A., Zurada, J.M. (eds.) ICAISC 2008. LNCS (LNAI), vol. 5097, pp. 398–404. Springer, Heidelberg (2008)
12. Ho, T.K.: The random subspace method for constructing decision forests. IEEE Trans. Pattern Anal. Mach. Intell. 20(8), 832–844 (1998)
13. Jain, A.K., Griess, F.D., Connell, S.D.: On-line signature verification. Pattern Recognition 35, 2963–2972 (2002)
14. Jie, Y., Fang, Y.Y., Renjie, Z., Qifa, S.: Fingerprint minutiae matching algorithm for real time system. Pattern Recognition 39, 143–146 (2006)
15. Kholmatov, A., Yanikoglu, B.: Identity authentication using improved online signature verification method. Pattern Recognition Letters 26, 2400–2408 (2005)
16. Korytkowski, M., Gabryel, M., Rutkowski, L., Drozda, S.: Evolutionary methods to create interpretable modular system. In: Rutkowski, L., Tadeusiewicz, R., Zadeh, L.A., Zurada, J.M. (eds.) ICAISC 2008. LNCS (LNAI), vol. 5097, pp. 405–413. Springer, Heidelberg (2008)
17. Li, X., Er, M.J., Lim, B.S., et al.: Fuzzy Regression Modeling for Tool Performance Prediction and Degradation Detection. International Journal of Neural Systems 20(5), 405–419 (2010)
18. Lee, L.L., Berger, T., Aviczer, E.: Reliable on-line human signature verification systems. IEEE Trans. on Pattern Anal. and Machine Intell. 18, 643–647 (1996)

19. Mazhelis, O.: One-class classifiers: a review and analysis of suitability in the context of mobile-masquerader detection. South African Computer Journal 36, 29–48 (2006)
20. Michalewicz, Z.: Genetic Algorithms + Data Structures = Evolution Programs. Springer (1998)
21. Nanni, L.: Experimental comparison of one-class classifiers for online signature verification. Neurocomputing 69, 869–873 (2006)
22. Nanni, L., Maiorana, E., Lumini, A., Campisi, P.: Combining local, regional and global matchers for a template protected on-line signature verification system. Expert Systems with Applications (2009)
23. Nelson, W., Kishon, E.: Use of dynamic features for signature verification. In: Proc. of the IEEE Intl. Conf. on Systems, Man, and Cyber., vol. 1, pp. 201–205 (1991)
24. Nelson, W., Turin, W., Hastie, T.: Statistical methods for on-line signature verification. Intl. Journal of Pattern Recognition and Artificial Intell. 8, 749–770 (1994)
25. Pabiasz, S., Starczewski, J.T.: Face reconstruction for 3D systems. In: Selected Topics in Computer Science Applications, pp. 54–63. EXIT (2011)
26. Pabiasz, S., Starczewski, J.T.: Meshes vs. depth maps in face recognition systems. In: Rutkowski, L., Korytkowski, M., Scherer, R., Tadeusiewicz, R., Zadeh, L.A., Zurada, J.M. (eds.) ICAISC 2012, Part I. LNCS (LNAI), vol. 7267, pp. 567–573. Springer, Heidelberg (2012)
27. Rutkowski, L.: Computational intelligence. Springer (2007)
28. Tormene, P., Giorgino, T., Quaglini, S., Stefanelli, M.: Matching incomplete time series with dynamic time warping: an algorithm and an application to post-stroke rehabilitation. Artificial Intelligence in Medicine 45, 11–34 (2009)
29. Yeung, D.-Y., Chang, H., Xiong, Y., George, S., Kashi, R., Matsumoto, T., Rigoll, G.: SVC2004: First International Signature Verification Competition. In: Zhang, D., Jain, A.K. (eds.) ICBA 2004. LNCS, vol. 3072, pp. 16–22. Springer, Heidelberg (2004)
30. Zalasiński, M., Cpałka, K.: A new method of on-line signature verification using a flexible fuzzy one-class classifier. In: Selected Topics in Computer Science Applications, pp. 38–53. EXIT (2011)
31. Zalasiński, M., Cpałka, K.: Novel algorithm for the on-line signature verification. In: Rutkowski, L., Korytkowski, M., Scherer, R., Tadeusiewicz, R., Zadeh, L.A., Zurada, J.M. (eds.) ICAISC 2012, Part II. LNCS (LNAI), vol. 7268, pp. 362–367. Springer, Heidelberg (2012)