offline signature verification,
shape context

Marcin ADAMSKI[1], Khalid SAEED[2]

OFFLINE SIGNATURE VERIFICATION BASED ON SHAPE CONTEXTS USING SHARED AND USER-SPECIFIC THRESHOLDS

In this paper we present a system for offline signature verification based on Shape Context Descriptors. The system input are binarized images of handwritten signatures from GPDS database available for non-commercial research. During preprocessing each signature image is thinned using KMM algorithm in order to obtain 1-pixel wide skeleton. The feature vector is built from Shape Context Descriptors computed for selected points on skeletonized signature line. The verification process is based on the distance measure that uses Shape Context Descriptors. The presented system is evaluated using random and skilled forgeries with shared and user-specific thresholds.

1. INTRODUCTION

Handwritten signature is behavioral biometric that is used on everyday basis for authorization of formal documents and financial operations. As a biometric trait its also an active subject of research since many years. There are many methods and commercial systems that are focused on automatic verification of handwritten signatures, however, there are still no perfect solutions and one can find many fields for improvements [5]. Usefulness of a signature as a biometric trait can be assessed based on the characteristics proposed in [6]. Its advantages are high acceptability of usage and simple registration (high collectivity). The main drawbacks are the ease of creating a forgery that can fool the system (high circumvention), low distinctiveness. In a larger database different persons may have very similar signatures and low permanence. The signature shape changes over time. Some of those problems can be reduced by using information about dynamics of the signing process when such data is available [4], [5].

2. BIOMETRIC SIGNATURE SYSTEM

Biometric systems designed for handwritten signatures typically follow an architecture given in Fig. 1. During the first stage the signatures are registered and stored in digital form. After the acquisition, signature data is usually (but not always) preprocessed in order to extract signature line from the image or resample and filter dynamic data. During feature extraction a feature vector is constructed - a mathematical representation of the signature instance that contains significant information required for proper classification. The classification process is used for one of the two tasks: identification and verification. The aim of verification is to decide whether the given biometric sample is genuine or

[1]Faculty of Computer Science, Bialystok University of Technology, Bialystok, Poland.
[2]Faculty of Physics and Applied Computer Science, AGH University of Science and Technology, Cracow, Poland.

forged. During the identification the system finds the individual whose signature best matches the given sample.

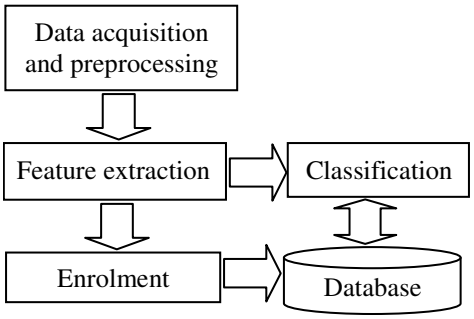

Fig. 1. Biometric signature system architecture.

Biometric signature systems are usually divided into two distinct categories based on how the acquisition process is conducted. When a signature is given on a piece of paper, a scanner or a camera device can be used to get its digital image. This kind of acquisition gives only static information about the signature shape. The samples obtained in this manner are called offline signatures. The main drawback of using offline data is that it is very difficult or sometimes impossible to detect forgeries. Signature shape can be easily imitated when a forger has the access to examples of authentic signatures and when comparing only the shapes of original and forged samples, it may be very difficult to distinguish them.

Handwritten signatures can be also registered using devices that capture the dynamics of the signing process (e.g., tablets and smartphones). In this case the registration includes not only the shape of the signature but also the dynamic data that describes the way the signature was written. It allows for better forgery detection due to the fact that dynamic features are much harder to imitate [4], [5]. Signatures acquired in this way are called online. However, in many applications where handwritten signatures are given on an piece of paper, only static information is available and such scenario is investigated in this work.

3. PROPOSED SYSTEM

The first stage in every biometric system is the acquisition of the input data measuring the biometric trait. In the case of offline signatures, the input data is usually obtained by scanning paper documents that contain the analyzed signatures. In our work we used signature images from GPDS database [9] which were initially extracted from signature forms and segmented by applying fixed threshold. During the preprocessing stage the thinning procedure was applied to reduce the amount of data to be processed in the later stages of the system. For this task we used KMM algorithm presented in work [8]. Examples of original and thinned images of signatures obtained by this method can be seen in Fig. 2.

Fig. 2. An example of a signature and its thinned version.

For further reduction of data we used the sampling technique. During the procedure of sampling, we leave N equally spaced points from the thinned signature line, where N is chosen arbitrarily. The sampling algorithm iteratively deletes the signature pixels until only the required number of N pixels remains in the image. The signature pixels selected for deletion in each iteration have the smallest

distance from their neighboring pixel. Examples of sampled versions of two thinned signatures can be seen in Fig. 3.

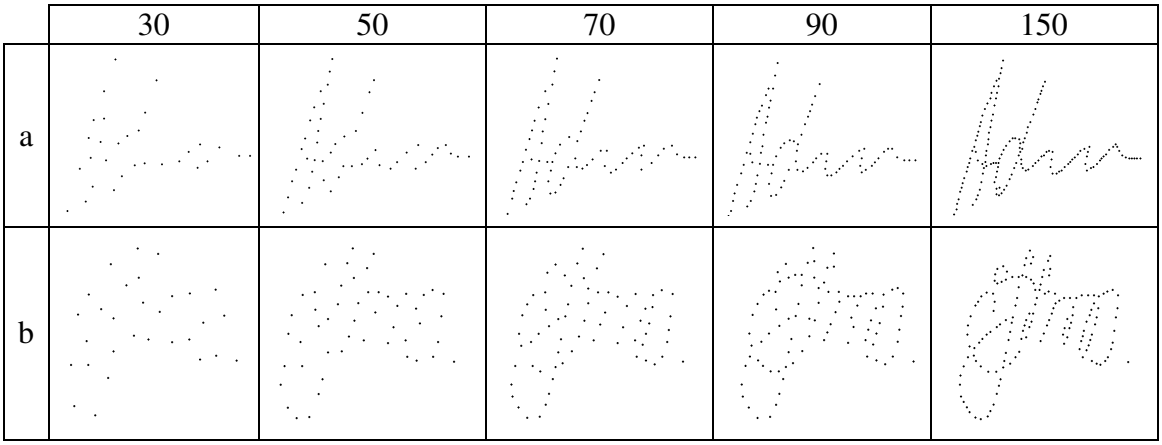

Fig. 3. Examples of the points selected form thinned signature lines for different values of N.

The number of the selected points N=150 was constant across all the signatures in the database and was based on the results of experiments conducted in authors' previous study [1]. The signature features description and the distance measure used for comparison of signatures was based on Shape Context method [3]. The Shape Context approach allows for measuring shape similarity between two graphical objects. Each of the two objects, whose shapes are compared, is represented by a set of points (1).

$$\begin{aligned} A &= \{a_1, a_2, \cdots, a_i, \cdots, a_N\} \\ B &= \{b_1, b_2, \cdots, b_i, \cdots, b_N\} \end{aligned} \tag{1}$$

For each point on the first object image (A) a corresponding point on the second object image (B) is found. In order to find the corresponding pairs each point is described by a Shape Context Descriptor. This descriptor contains information about the configuration of the entire shape relative to the point being described (Fig. 4). It is computed as a coarse histogram representing the distribution of points comprising the object relative to the reference point.

The histograms can be calculated by counting the number of points in each bin (2).

$$h_k(a_i) = \#\{a_j : a_j \in bin(k), j = 1...N \wedge i \neq j\} \tag{2}$$

where $bin(k)$ is the k-th bin of histogram describing the distribution of points in A relative to point a_i, K is the number of bins in the histogram (in our experiments we used K=60).

The cost of matching two histograms forming a pair - one describing a point from object A and the other from B, can be based on χ^2 test statistics and is given by (3).

$$d_h(a_i, b_j) = \frac{1}{2} \sum_{k=1}^{K} \frac{[h_k(a_i) - h_k(b_j)]^2}{h_k(a_i) - h_k(b_j)} \tag{3}$$

The total cost of matching two objects A and B that is used in this work is the sum of minimal distances between the points a_i and b_j given in (4).

$$D_h(A_h, B_h) = \frac{1}{N} \sum_{a_i \in A} \min_{b_j \in B} d_h(a_i, b_j) \tag{4}$$

In work [2] authors introduced extended shape context descriptors that incorporate more information on original shape of an object image. In this method, during the computation of the histogram that describes the distribution of points relative to a particular point on a signature line, a complete skeletonized image is used (Fig. 5). Hence, each histogram contains more precise information than its basic

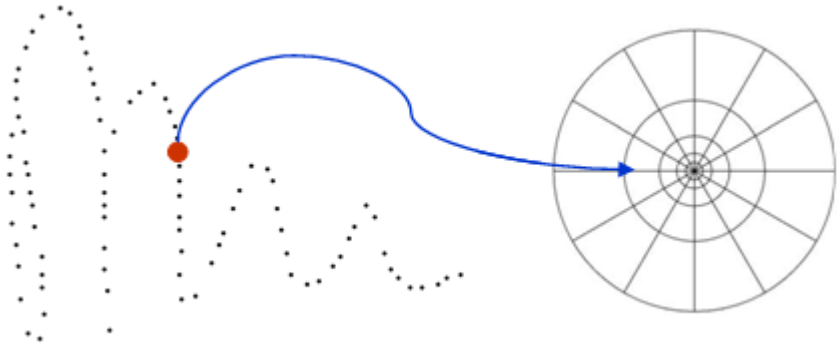

Fig. 4. Description of signature point using histogram based on a sampled signature skeleton [2].

version. The number (N) and position of reference points stays the same as in the original algorithm and therefore the cost of comparing two signatures on the basis of Shape Context measure is the same. Due to the fact that the shape contexts of each signature is calculated only once and can be stored as a reference for future comparisons, extended context descriptors allow for improvement of a signature image without a penalty in the system responsiveness.

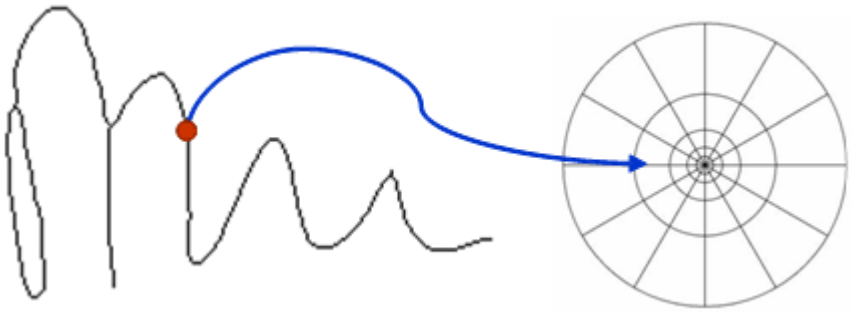

Fig. 5. Description of signature point using histogram based on complete signature skeleton without sampling [2].

4. EXPERIMENTS

In this section we present several experiments on handwritten signature verification carried out in order to assess the effectiveness of the proposed system. For the experiments 8 genuine signatures and 8 skilled forgeries of 40 individuals were selected at random from the GPDS database giving a total number of 640 signatures. Half of the genuine signature set for every individual was used as a reference set whilst the rest of the signatures were left for testing.

In the verification task the questioned signature is compared with the references of a particular person to assess its authenticity. The decision of the system is based on the threshold value T. When the minimum distance between the questioned signature and reference samples of a particular person exceeds the threshold value, the signature is rejected, otherwise it is accepted as a genuine example.

During the verification task the system is assessed using genuine and forged samples. In this work we used two types of forgeries:

- random forgeries – genuine signatures of one person are presented as imitations of signatures of another person,
- skilled forgeries – imitations created by forgers who have access to genuine examples and can spend as much time as required to train how to imitate original signature shape.

4.1. VERIFICATION USING SHARED THRESHOLD

Experiments presented in this section were carried out using a single value of threshold for all signers. This constraint does not allow to adjust for differences in variability of the signatures given by particular individuals. However, this approach can find its applications, when variability cannot be reliably assessed due to small number of reference samples. The results obtained for shared threshold value were presented in authors' previous work [2] and are given here for comparison with new research where user-specific threshold is applied. Table 2 shows the average EER obtained using random forgeries and different selections of reference and test signatures. Experiments were carried out with both basic and extend contexts.

Table 1. Results of signature verification using random forgeries [2].

Equal Error Rate	
basic SC %	**extended SC %**
4.9	4.4

FAR and FRR curves computed for extended Shape Context in this experiment are shown in Fig. 6.

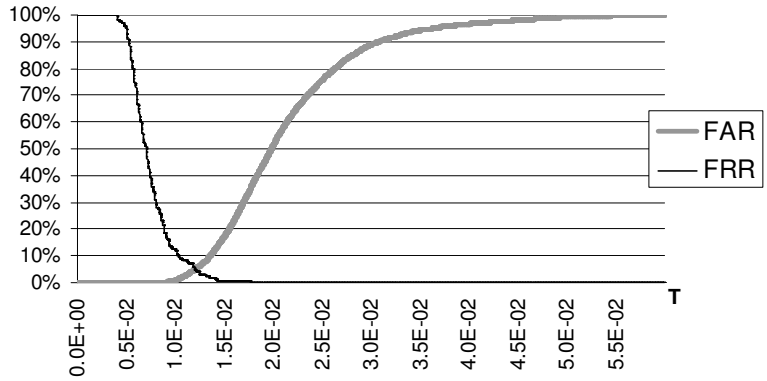

Fig. 6. Example of FAR and FRR curves obtained with extended SC for random forgeries [2].

The shared threshold approach was also investigated for skilled forgeries that are much harder to differentiate from genuine samples than random imitations. In some cases it may be almost impossible to detect such forgeries using only static information that is available in offline systems. The results for skilled forgeries are shown in Table 2.

Table 2. Results of signature verification using skilled forgeries [2].

Equal Error Rate	
basic SC %	**extended SC %**
22.4	20.6

The results in Table 2 are much worse than the case of random forgeries, however, similar errors are obtained by other offline systems (FRR=25%, FAR=26% in [7], EER= 19.51% in [9]) using GPDS database. Fig. 7 shows FAR and FFR curves computed to the extended Shape Context for skilled forgeries.

4.2. VERIFICATION USING USER-SPECIFIC THRESHOLD

In the experiments presented so far the threshold value used for making decision whether to accept or reject the given signature was the same for all subjects. Due to the fact that signatures of different

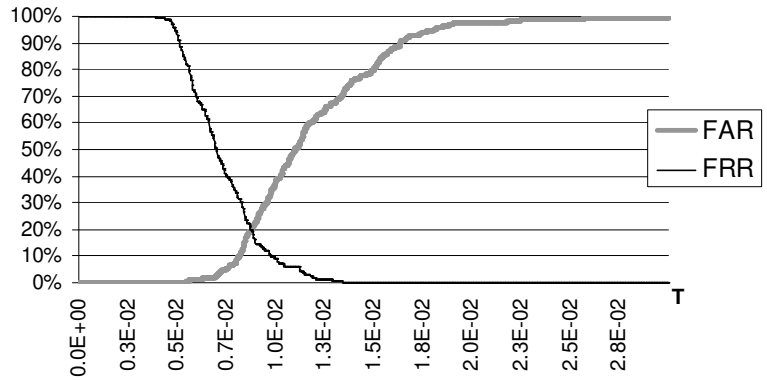

Fig. 7. Example of FAR and FRR curves obtained with extended SC for skilled forgeries [2].

people vary in complexity and stability, using a single threshold value is not optimal and may lead to an increase in the system error. In order to adjust for variability of signatures given by particular individuals we carried out additional experiments where the threshold value was computed separately for each signer. Table 3 presents the average EER for random forgeries obtained when user-specific threshold was applied. As can be seen from Table 3, threshold specific for each user allows to distinguish random forgery from genuine samples with much lower error than a single threshold value (Table 1).

Table 3. Results of signature verification using random forgeries with user-specific threshold.

Equal Error Rate	
basic SC %	**extended SC %**
0.6	0.5

Table 4 presents the average EER for skilled forgeries obtained when a user-specific threshold was applied. As can be seen from Table 4 the errors are much lower compared to the results achieved when single threshold value was utilized (Table 2). In this case, extended SC also gave better results than its basic version.

Table 4. Results of signature verification using skilled forgeries with user-specific threshold.

Equal Error Rate	
basic SC %	**extended SC %**
9.3	8.4

5. CONCLUSIONS

In this paper we presented the Shape Context Descriptor and its extended version applied to signature verification using shared and user-specific models. The main contribution of this work is investigation of Shape Context Descriptors for building user-specific models and comparison to shared model evaluated under the same experimental conditions. The results obtained from this investigation encourage to further work on user-specific models and Shape Context Descriptor. Other improvements may be related to performance of signature matching - a simple pruning based on global features might eliminate the need of computing complete descriptors in cases where signature shapes are significantly different.

ACKNOWLEDGMENTS

This work is partially supported by the Rector of Bialystok University of Technology (grant no. S/WI/2/2013).

BIBLIOGRAPHY

[1] ADAMSKI M., SAEED K., Offline Signature Identification and Verification using Noniterative Shape Context Algorithm, Journal of Medical Informatics and Technologies, 2009, Vol. 13, pp. 47-58.
[2] ADAMSKI M., SAEED K., TABĘDZKI M., RYBNIK M., Signature System Based on Extended Shape Context Descriptors, International Conference on Biometrics and Kansei Engineering (ICBAKE), 2013, p. 267-272.
[3] BELONGIE S., MALIK J., PUZICHA J., Shape Matching and Object Recognition Using Shape Contexts," IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, Vol. 24, pp. 509-522.
[4] DOROZ R., PORWIK P., PARA T., WROBEL K., Dynamic signature recognition based on velocity changes of some features, International Journal of Biometrics, 2008, Vol. 1, No. 1, pp. 47-62.
[5] IMPEDOVO D., PIRLO G. , Automatic Signature Verification: The State of the Art, IEEE Transactions on Systems, Man and Cybernetics, Part C: Applications and Reviews, 2008, Vol. 38, pp. 609-635.
[6] JAIN A. K., ROSS A., PRABHAKAR S. , An Introduction to Biometric Recognition, IEEE Transactions on Circuits and Systems for Video Technology, 2004, Vol. 14, pp. 4-20.
[7] PIRLO G., IMPEDOVO D., On the measurement of local stability of handwriting: An application to static signature verification, in IEEE Workshop on Biometric Measurements and Systems for Security and Medical Applications (BIOMS), 2010, pp. 41-44.
[8] SAEED K., TABĘDZKI M., RYBNIK M., ADAMSKI M., K3M - A Universal Algorithm for Image Skeletonization and a Review of Thinning Techniques, International Journal of Applied Mathematics and Computer Science, 2010, Vol. 20, pp. 317-335.
[9] VARGAS J. F., FERRER M. A., TRAVIESO C. M., ALONSO J. B., Off-line Handwritten Signature GPDS-960 Corpus, in Ninth International Conference on Document Analysis and Recognition (ICDAR), 2007, pp. 764-768.