A Writer-Dependent Approach to Off-line Signature Verification

Valery Starovoitov
Laboratory of the System Identification
United Institute of Informatics Problems
Minsk, Belarus
valerys@newman.bas-net.by

Umidjon Akhundjanov
Laboratory of the System Identification
United Institute of Informatics Problems
Minsk, Belarus
umidjan_90@mail.ru

***Abstract*—Results of a new approach to off-line signature verification are presented. The approach is writer-dependent. To verify a signature, only $15 \geqslant N \geqslant 5$ genuine signatures of the person are used. The signature images are pre-processed and normalized into a contour representation. We then compute two new signature features: the distribution of LBP values and local curvature of contours in the binary signature image. For a signature submitted for analysis, *N* genuine signatures of this person are randomly selected and a one-class SVM classifier is developed. Accuracy of our approach in verification of all 2640 signatures from the public CEDAR database was 99.77%. All fake signatures were correctly recognized even with *N*=5 genuine signatures used to build the classifier.**

Keywords—signature, off-line verification, image processing, features, classifier, one-class SVM

I. INTRODUCTION

Signature recognition is a behavioral biometric technology. Signature verification is a procedure for determining its authenticity in the presence of a small number of genuine signatures of the same person. Most methods on this topic use machine learning to extract signature features and classify them into fake and genuine [1-3]. The problem is that in practice, when an expert establishes the authenticity of a person's signature; his fake signatures are not given! The number of authentic human signatures presented on paper available for analysis is also limited. Usually it does not exceed 5-15 samples.

Since the features of each person's signature are individual, the classifier needs to be trained on the signatures of one person. In the absence of fake signatures of this person, the task of signature verification is one-class and methods using a neural network approach are unsuitable for solving it. At the moment, the only machine learning method suitable for solving the problem of offline signature verification is development of a one-class SVM model using available samples of genuine signatures of the person [4].

The paper describes a procedure for signature image preprocessing, its image normalization, two new signature features and a new feature space for training the one-class SVM model and verification the given signature.

II. SIGNATURE IMAGE PREPROCESSING

A signature on paper can be made by pens of different thicknesses and different colors. As a result of our experiments, it was found that it is optimal to scan any signature with a resolution of 600 dpi as a color image. The image is then converted to grayscale, taking into account the dominant tones of the background and the signature itself. After this, the image is converted into binary representation by the Otsu method, since, as a rule, the background of the documents on which the signature is presented is uniform.

Morphological and median filters are applied to the binary signature image. The image is then rotated using the main principal component computed by PCA method so that the signature is oriented horizontally.

Since all digitized signatures have different sizes in pixel, we find a circumscribing rectangle in the image that describes the signature, cut it out and scale it to a certain size. In the experiments, we used a normalized signature template size of 300x150 pixels. Using a method of mathematical morphology, a contour representation of the signature image was constructed. All signatures are transformed to such binary contour representation of a fixed size, see an example in Fig.1.

Fig. 1. Example of an original signature image and its normalized contour representation.

III. NEW SIGNATURE FEATURES

The contour representation of the normalized signature is used to compute its features. As new features that describe the individual characteristics of a signature, it is proposed to use local binary patterns (LBP) and the local curvature of the signature contours, calculated in the neighborhood of each pixel of the normalized signature.

A. LBP signature code

After calculating all the values of LBP features of a binary signature image, their histogram is constructed (Fig. 2). It is convenient to write it as an array of 256 elements. From this array, the first and last elements corresponding to the cases when all eight neighboring pixels have white or black values are discarded, and the histogram is normalized. We call the set of 254 numbers as the LBP signature code. It is a multidimensional feature that describes the frequency distribution of local structures of the signature contour,

regardless of color, line thickness, original dimensions and orientation of the original signature represented on paper.

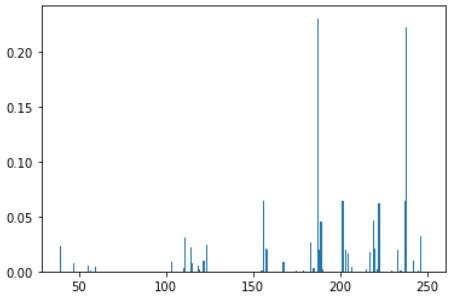


Fig. 2. Normalized histogram of the LBP values for the contours represented in Fig.1.

B. Local curvature code

Curvature is a widely used invariant feature for image classification. Flynn and Jain [5] report an empirical study of five curvature estimation methods. The main finding of their study is that the estimated curvature values are extremely sensitive to quantization noise and require multiple smoothing of the lines to obtain stable estimates.

There are only three variants of consecutive contour pixels invariant to rotation and giving different curvature values. The dark color in Fig. 3 shows the pixel where the circle touches the curve. It calculates the amount of curvature of a curve.

Fig. 3. Three variants for the positions of three consecutive contour pixels.

To increase the number of options for contour curvature values, we use five consecutive 8-connected contour pixels. In this case, the curvature in a pixel is not calculated exactly, it is approximated.

Note that rotations of the pixel configurations presented in Fig. 3 by 45° do not change the curvature values. The minimumal curvature is 0 in the case when all pixels lie on the same straight line.

Each person's signature in the normalized raster representation has a different number of pixels, so the histogram of the local curvature values must be normalized by dividing by the number of points at which the curvature was calculated (Fig.4).

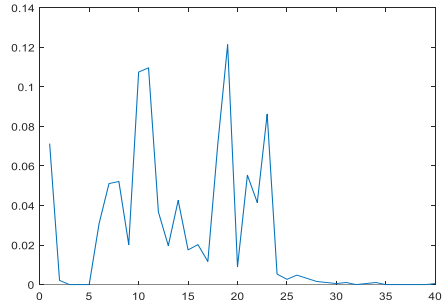


Fig. 4. Normalized histogram of the local curvature values for the contours represented in Fig.1.

Normalized histograms, presented as an array or a feature vector, we call the signature local curvature code. It describes the individual features of a person's digitized signature.

For every signature, two codes of local features are calculated. Codes of different signature images can be compared with each other and, by assessing their proximity, a conclusion can be made about the similarity of the original signatures.

IV. CLASSIFICATION

In our research, instead of the classical representation of an image of an object (in this study, a signature) as a point in a multidimensional space, we use the transformation of the image into a two-dimensional space. In it, images are presented in the form of proximity of pairs of signatures, while the proximity of signatures is assessed by rank correlation coefficients calculated for the codes of these signatures. It is possible to use other functions for calculating proximity, however, the correlation coefficient, unlike other functions, has a limited range of values [-1; +1] and is easy to calculate. This feature space describes the proximity images of all possible pairs of genuine human signatures. The number of patterns im this feature space depends on the N genuine signatures used in the verification procedure. These parameters are presented in Table I. One can see that bigger N, more patterns for classifier model learning.

TABLE I. NUMBER OF PATTERNS FOR DIFFERENT N OF GENUINE SIGNATURES USED IN VERIFICATION PROCEDURE

№	Number of patterns with verifiable signature, *N* subjects	Numbers of patterns of genuine signatures, *N*(*N*-1)/2 subject
1	5	10
2	7	21
3	9	36
4	11	55
5	13	78
6	15	105

Recall that generally used augmentation procedures such as image rotation, scaling, changing color and contrast do not change the number of the normalized signature representation, i.e. do not increase the number of patterns for training a classifier. At the same time, the transition from a signature pattern to patterns of the proximity of pairs of signatures allows us to increase the number of patterns for training, see Table 1. Classification in the original multidimensional feature space with a small number of signature patterns is less accurate and sensitive to the outliers, since their ranges of values are not equal. In the new feature space, this drawback is leveled out, since the correlation has the fixed and limited range of values, and the number of patterns of genuine objects increases. For example, with $N = 15$ instead of 15 patterns, we can learn the classifier model on 105 patterns.

When performing real verification examinations, there are no fake signatures, so we use a one-class classifier such as SVM [6]. After training its model on a class of authentic signatures of a person, two codes of the verified signature are compared with the codes of all authentic signatures used in training, i.e. with N patterns of the verifiable signature are constructed. Next, we analyze whether most of these N patterns fall into the class of genuine signatures or into outliers. The classification result is determined by the majority.

The experiments were carried out on the basis of the CEDAR dataset containing 1320 genuine and 1320 forged signatures. 24 signatures of each type for 55 people [7].

The one-class SVM model was retrained for every signature, i.e., trained on N randomly selected genuine signatures of the same person. The highest verification accuracy with $N = 15$ was 99.77% (Table II). Moreover, even with a minimal number of genuine signatures $N = 5$ used for model training, all fake signatures were correctly recognized. The results obtained are the best for the CEDAR database.

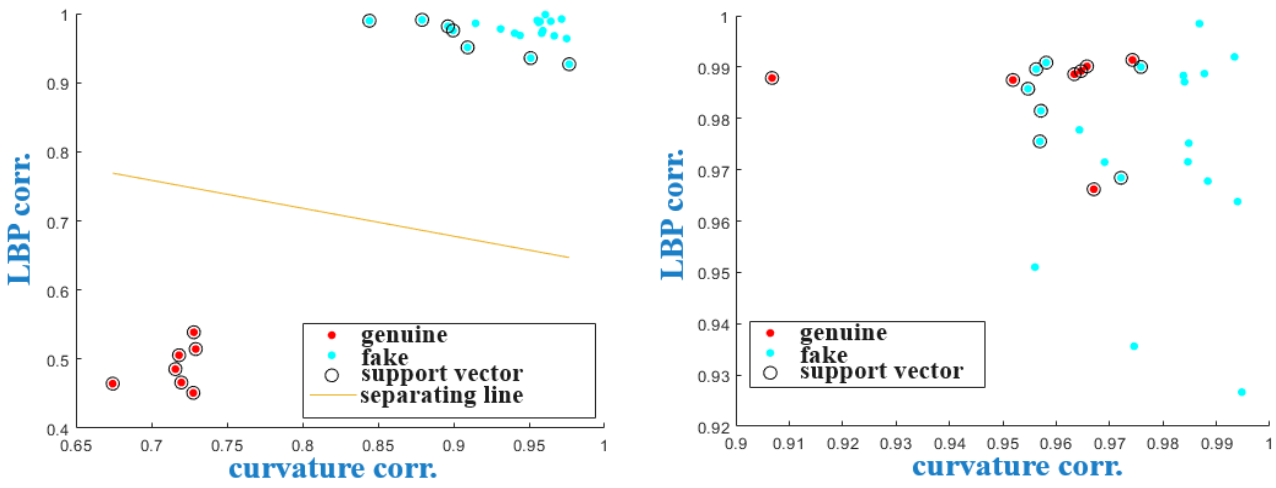
On the top of Figures 5 and 7 examples of forgeries and genuine signatures of one person used for verification are presented. Below are the genuine signatures of this person used to training the classifier. In Fig. 6 and 8 are presented patterns of proximity pairs of genuine signatures in green and pairs (verifiable, genuine) in red. In Fig. 6a) and 8a), red patterns can be separated from the class of patterns of genuine signatures presented in green, but in Fig. 6b) and 8b) – no. This means the top right signatures in Fig. 6a) and 8a) are original.

TABLE II. VERIFICATION RESULTS OF 2640 HANDWRITTEN SIGNATURES FROM CEDAR DATABASE WITH DIFFERENT N

N	TP	TN	FP	FN	*Accuracy, %*
5	1131	1320	189	0	92,84
7	1237	1320	83	0	96,86
9	1279	1320	41	0	98,45
11	1293	1320	27	0	98,98
13	1307	1320	13	0	99,51
15	**1314**	**1320**	**6**	**0**	**99,77**

The results obtained are superior to known results obtained using writer-dependent classifiers. The best published results for verifying signatures from CEDAR database based on writer-dependent classifiers were obtained in [4] with N = 12 and accuracy was equal to 94.4%. The average EER errors were 8.70%, 7.83%, and 5.60% when using N = 4, 8, and 12 genuine signatures to train the one-

Fig. 5. Top row – 2 images for verification from the CEDAR database: on the left – a fake signature, on the right – a genuine one; bottom row – 7 genuine signatures of the same person used for learning



a) b)

Fig. 6. An example of signature verification when compared with seven genuine ones: (a) the Ver_f.png signature may be separated from the genuine signatures and it is recognized as fake; (b) the signature Ver_t.png cannot be separated from other signatures and it is recognized as genuine

Fig. 7. Top row – 2 images for verification from the CEDAR database: on the left – a fake signature, on the right – a genuine one; bottom row – 7 genuine signatures of the same person used for learning

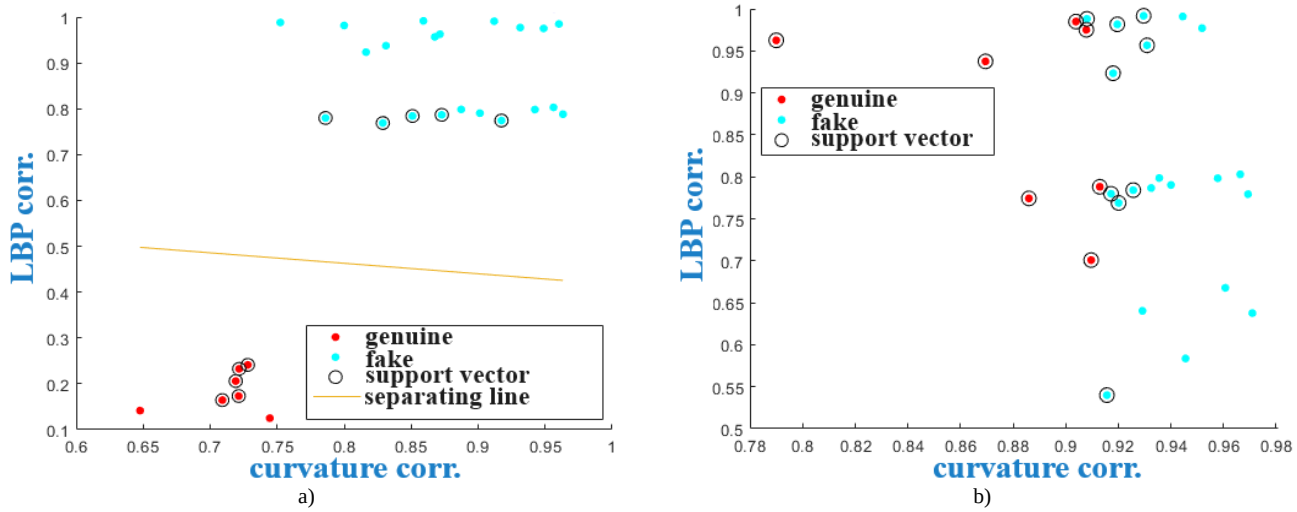


Fig. 8. An example of signature verification when compared with seven genuine ones: (a) the Ver_f.png signature may be separated from the genuine signatures and it is recognized as fake; (b) the signature Ver_t.png cannot be separated from other signatures and it is recognized as genuine

class model. Note, that in [4] the results were obtained not for the all database signatures, but for 45,4% of the signatures presented in it. The rest 55% were used for classifier training.

Ghosh in [8] reported about a Neural Network classification model trained on 12 genuine signatures of each person presented in the CEDAR database. He verified the remaining genuine and all forged signatures from the database. Thus, for 75% of the rest images from the database, he obtained accuracy of 99.94%. Note that he used individual signature models of 55 people. If his system is tested on the signatures of people not included in this database, the result will be unpredictable.

V. Conclusion

The paper presents results of our research devoted to solving the problem of verifying signatures scanned from paper documents, i.e. the off-line signature verification.

We proposed two fundamentally new features that describe the individual characteristics of a person's signature. They describe the normalized frequency distribution of LBP and the normalized frequency distribution of local curvature values calculated from a binary signature contour representation.

To implement the verification procedure, we propose to use a new feature space in which patterns of the proximity of every signature pair codes are presented. Usage the proximity patterns of signature pairs allows us to reduce dimension of the feature space, but increase the number of patterns for training a classifier.

We have developed a writer dependent one-class SVM classifier. On Cedar database trained on 15 randomly selected genuine signatures it have demonstrated accuracy 99.77% when all 2640 signatures were verified. An individual classifier was built for every signature. Moreover, all fake signatures were correctly recognized when the classifier was trained on $N \geq 5$ genuine signatures of one person.

References

[1] L.G. Hafemann, R. Sabourin, L/S/ Oliveira, "Offline handwritten signature verification—literature review," 7-th Int. Conf. on Image Processing Theory, Tools and Applications, 2017 Nov 28, pp. 1-8.

[2] M. Stauffer, P. Maergner, A. Fischer, K. Riesen, "A survey of state of the art methods employed in the offline signature verification process," New trends in business information systems and technology, 2020, pp.17-30.

[3] H. Kaur, M. Kumar, "Signature identification and verification techniques: state-of-the-art work", Journal of Ambient Intelligence and Humanized Computing, 2023, Vol.14, No.2, pp.1027-1045.

[4] Y. Guerbai, Y. Chibani, B. Hadjadji, "The effective use of the one-class SVM classifier for handwritten signature verification based on writer-independent parameters," Pattern Recognition, 2015, Vol. 48, No. 1, pp.103–113.

[5] P.J. Flynn, A.K. Jain, "On reliable curvature estimation,", Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1989, Vol. 88, pp. 5–9.

[6] S. Alam, *et al*,. "One-class support vector classifiers: A survey," Knowledge-Based Systems, 2020, Vol. 196, pp. 105754.

[7] M.K. Kalera, S. Srihari, A. Xu, "Offline signature verification and identification using distance statistics," Int. J. of Pattern Recognition and Artificial Intelligence, 2004, Vol. 18, No.7, pp. 1339–1360.

[8] R.A Ghosh, "Recurrent Neural Network based deep learning model for offline signature verification and recognition system," Expert Systems with Applications, 2021, Vol. 168, pp. 114249.