

## Off-line Bangla Signature Verification

Srikanta Pal  
School of Information  
and Communication  
Technology, Griffith  
University, Gold Coast  
Australia, Email:  
srikanta.pal@griffithuni.  
edu.au

Vu Nguyen  
School of Information  
and Communication  
Technology, Griffith  
University, Gold Coast,  
Australia, Email:  
vu.nguyen2@griffithuni.  
edu.au

Michael Blumenstein  
School of Information  
and Communication  
Technology, Griffith  
University, Gold Coast,  
Australia, Email:  
m.blumenstein@griffith.  
edu.au

Umapada Pal  
Computer Vision and  
Pattern Recognition  
Unit, Indian Statistical  
Institute, Kolkata-  
700108, India. Email:  
umapada@isical.ac.in

**Abstract**— In the field of information security, biometric systems play an important role. Within biometrics, automatic signature identification and verification has been a strong research area because of the social and legal acceptance and extensive use of the written signature as an individual authentication. Signature verification is a process in which the questioned signature is examined in detail in order to determine whether it belongs to the claimed person or not. Despite substantial research in the field of signature verification involving Western signatures, very few works have been dedicated to non-Western signatures such as Chinese, Japanese, Arabic, or Persian etc. In this paper, the performance of an off-line signature verification system involving Bangla signatures, whose style is distinct from Western scripts, was investigated. The Gaussian Grid feature extraction technique was employed for feature extraction and Support Vector Machines (SVMs) were considered for classification. The Bangla signature database employed in the experiments consisted of 3000 forgeries and 2400 genuine signatures. An encouraging accuracy of 90.4% was obtained from the experiments.

**Keywords**- *Offline verification systems, signature verification, biometrics, authentication systems, Gaussian Grid feature, SVMs.*

### 1. INTRODUCTION

The field of biometrics is an important area of study as it offers many advantages over more commonly used authentication methods such as photo ID cards, magnetic strip cards etc. Nowadays, biometric technologies are increasingly and more frequently being used to ensure identity verification. Signatures often incorporate complex geometric patterns that make them a relatively secure means for authorization for high security environments. For historical reasons, the handwritten signature continues to be the most commonly accepted form of transaction confirmation, as well as being used in civil law contracts, acts of volition, or authenticating one's identity. Signature verification has been a topic of intensive research during the past several years due to the important role it plays in numerous areas, including in financial applications.

Signatures have been accepted as an official means to verify personal identity for legal purposes on such documents as cheques, credit cards, wills etc. The handwritten signature is therefore well established and accepted as a behavioral biometric. Considering the large number of signatures verified daily through visual inspection by people, the construction of a robust and accurate automatic signature verification system has many potential benefits for ensuring authenticity of signatures and reducing fraud and other crimes.

The goal of an automatic signature verification system is to be able to verify the identity of an individual, based on the analysis of his or her signature through a process that discriminates a genuine signature from a forgery. The verification of human signatures is particularly concerned with the improvement of the interface between human-beings and computers [2]. A signature verification system and the associated techniques used to solve the inherent problems of authentication can be divided into two classes [3]: (a) on-line methods [4] to measure temporal and sequential data by utilizing intelligent algorithms [5] and (b) off-line methods [6] that use an optical scanner to obtain handwriting data written on paper. Off-line signature verification deals with the verification of signatures, which appear in a static format [7]. On-line signature verification has been shown to achieve much higher verification rates than off-line verification [6], as a considerable amount of dynamic information is not readily available in the off-line mode. However, off-line systems have a significant advantage that they do not require special processing devices when the signatures are produced. Moreover, the off-line systems have many more practical application areas than their on-line counterparts.

Signatures represent a particular writing style and very often are a combination of symbols and strokes. So it is obviously necessary to deal with a signature as a complete image with a special distribution of pixels, representing a particular writing style and is not considered as a collection of letters and words [8]. It is often difficult for a human to instantly verify two signatures of the same person because signature samples from the same person are similar but not

identical, and signatures can change depending on elements such as mood, fatigue, and time. In addition, a person's signature may change radically during their lifetime. Great inconsistency can even be observed in signatures according to country, habits, psychological or mental state, physical and practical conditions [9].

There has been substantial work in the area involving off-line verification for Western signatures. Justino *et al.* [10] proposed an off-line signature verification system based on Hidden Markov Models (HMMs) for the detection of random, casual, and skilled forgeries. Three features: a pixel density feature, a pixel distribution feature and an axial slant feature are extracted from a grid segmentation scheme. Ferrer *et al.* [2] presented an automatic offline signature verification system based on fixed-point arithmetic. They used the set of geometric features for offline automatic signature verification. The feature set was calculated using 16 bits fixed-point arithmetic and they tested it with different classifiers, such as Hidden Markov Models, Support Vector Machines, and the Euclidean distance classifier. Nguyen *et al.* [11] developed an off-line signature verification system using global features in a Support Vector Machine classifier. In their paper, the combination of the Modified Direction Feature (MDF) with three global features: features from Energy information, maxima feature, and ratio feature was reported. Schafer and Viriri [12] presented an off-line signature verification system based on the combination of feature sets. Some extracted features were: Aspect ratio, centroid feature, four surface features, six surface features, number of edge points, transition features etc. The verification of signatures was accomplished using the Euclidean distance classifier.

Signatures may be written in different languages and there is a need to undertake a systematic study in this area. Although most of the published work has been undertaken for Western signatures, only a few studies have been undertaken for Chinese, Japanese, Persian, and Arabic signatures [18]. Many systems involving off-line signature identification and verification have been developed but not a single technique has solely considered Bangla-script signatures. Moreover, to the best of the authors' knowledge there has been no published signature verification work on signatures written in Indian languages. The present work of Bangla signature verification would be considered as a novel contribution to the field of signature verification. Some signature samples of Bangla script are shown in Fig. 1.

The remainder of this paper is organized as follows. The different types of verification errors and forgeries are described in Section II. The Bangla signature database developed for the current research is described in Section III. Some notable properties of Bangla script are introduced in Section IV. Section V briefly describes the feature extraction technique employed. Details of the classifiers used are presented in Section VI. The experimental settings are presented in Section VII. Results and discussion are

given in Section VIII. Finally, conclusions and future work are discussed in Section IX.

টিকা মল্লিক	সুজিত মল্লিক
রুপা মল্লিক	বেলাল হোসেন
সৌম্য মল্লিক	সুজিত মল্লিক

Figure 1. Bangla signature samples

## II. TYPES OF FORGERIES

In general, off-line/on-line signature verification can be considered a two-class classification problem. Here the first class represents the genuine signature set, and the second class represents the forged signature set. Usually two types of errors are considered in a signature verification system: The False Rejection or Type-I error and the False Acceptance or Type-II error. These error types associate with two common types of error rates: False Rejection Rate (FRR) which is the percentage of genuine signatures misclassified as forgeries, and False Acceptance Rate (FAR) which is the percentage of forged signatures misclassified as genuine. According to Coetzer *et al.* [13], three basic types of forged signatures, which are often taken into account, are:

1. Random forgery. The forger has no access to the genuine signature (not even the author's name) and reproduces a random one. In many cases, the forgeries are the forgers' own genuine signature.
2. Simple forgery. The forger knows the author's name and the script, but has no access to a sample of the signature.
3. Skilled forgery. The forger has access to one or more samples of the genuine signature and is able to reproduce it.

## III. BANGLA SIGNATURE DATABASE

Although automatic signature verification has been an active research area for several decades, there has been no publicly available signature database for Bangla, the second most popular Indian script. Therefore, a Bangla signature corpus was created for the purpose of this work. The research in automatic signature verification has long been constrained by the unavailability of a standard database.

### A. Data Collection and Database Preparation

The signatures of Bengali script were considered for this signature verification approach. As there has been no public signature corpus available for Bengali script, it was necessary to create a database of Bengali signatures. The

signatures were collected from different parts of West Bengal, India. The majority of the signatures were contributed by students. The remaining signatures were collected from seniors of the locality. This Bangla signature database consists of 100 sets. The signature images employed consisted of 100 sets whereby the writer number ranges from B-S-Set-001 to B-S-Set-100 (Bangla-Signature-Set). In order to collect the genuine signatures corresponding to each individual, a collection form was designed. The form contained 24 boxes where the signatures could be written. From each individual, 24 genuine signatures were collected. A total number of 2400 genuine signatures from 100 individuals were collected. For each contributor, all genuine specimens were collected in a single day's writing session. In addition, only skilled forged signatures were collected for this proposed work. In order to produce the forgeries, the imitators were allowed to practice their forgeries as long as they wished with static images of genuine specimens. A total number of 3000 forged signatures were collected from the writers. Some genuine signature samples with their corresponding forgeries are displayed in Fig. 2.

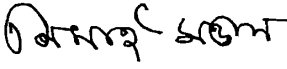
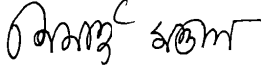
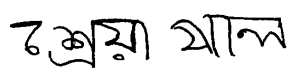


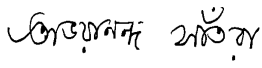
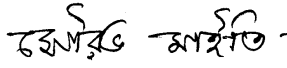
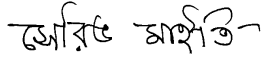
Bangla Signatures	
Genuine Signatures	Forgery Signatures
	
	
	
	

Figure 2. Genuine and forged signatures

### B. Pre-processing

The signatures to be processed by the system needed to be in a digital image format. It was necessary to scan all signature document pages. At the very beginning, the images were captured in 256 level grey scale at 300 dpi and stored in TIFF format (Tagged Image File Format) for the purpose of future processing. In the pre-processing step, a histogram-based threshold technique was applied for binarization. In this step, the digitized grey-level image is converted to a two-tone image. Then the signature images were extracted from the signature-collecting document forms. The extracted binary signature images were stored in

TIFF format. A typical scanned signature and its corresponding binary image are shown in Fig. 3 and Fig. 4, respectively.

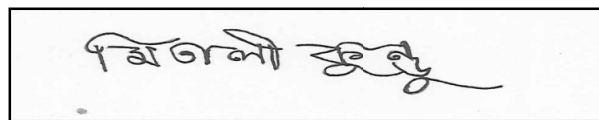


Figure 3. Scanned signature image

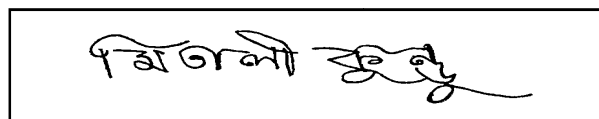


Figure 4. Binary signature image

## IV. PROPERTIES OF BANGLA SCRIPTS

Bangla is an oriental script descended from Brahmi script [14]. It is the second most popular official script in India. Other countries using Bangla include Bangladesh. It is estimated that there are about 300 million people using Bangla script. In Bangla script, the writing direction is from left to right and there is no concept of upper/lower case. Bangla script has fifty basic characters. These characters are presented in Fig. 5. There exists an upper guide-line in both printed and handwritten Bangla text.

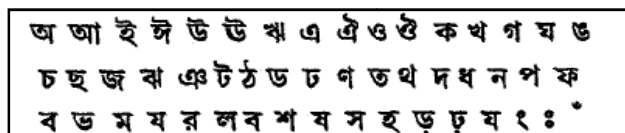


Figure 5. Basic characters of Bangla script

## V. FEATURE EXTRACTION

Feature extraction is a crucial step in any pattern recognition system. The Gaussian Grid feature extraction technique, which was previously employed for English signature verification, is described below.

### Gaussian Grid Feature:

The Gaussian grid feature is a relatively new grid-based feature extraction technique originally designed for the Western signature verification problem. In [15], this technique reportedly produced verification rates greater than the Modified Direction Feature (MDF) and the Gradient feature extraction techniques. The study also indicated that the high performance of the Gaussian grid feature was generated by appropriately applying the Gaussian filter on the matrices of accumulated direction information. The extraction of Gaussian grid features proceeds as follows: A 12x12 grid segments the noise cleaned contour

representation of the signature into cells of equal size. For each cell, the contours are traversed and the number of movements are counted for each of the four main directions (horizontal, vertical, left-diagonal, and right-diagonal) and are stored in four separate matrices. A 2D Gaussian filter is then applied on these matrices before the values are rescaled to be within the range of [0,1]. From these four matrices, two additional matrices of the same dimensions are created by manipulating pairs of matrices for perpendicular directions. The value of each cell is computed by dividing the minimum value by the maximum value of two corresponding cells of the pair of input matrices.

As mentioned above, the Gaussian Grid feature was initially developed for Western signatures, which are significantly different from Bengali signatures. By visual inspection, one could easily distinguish these two types of signatures. Bengali signatures tend to be wider and shorter. The existence of a baseline in Bengali signatures is also more evident than Western signatures. It is of interest to investigate the robustness of the Gaussian Grid feature, which was optimized for Western signatures on another style of signature like Bengali.

## VI. CLASSIFIERS DETAILS

SVM Classifier:

In our experiments, we have used Support Vector Machines (SVM) as classifiers. SVMs have been originally defined for two-class problems and they look for the optimal hyper plane, which maximizes the distance and the margin between the nearest examples of both classes, namely support vectors (SVs). Given a training database of  $M$  data:  $\{x_m | m=1, \dots, M\}$ , the linear SVM classifier is then defined as:

$$f(x) = \sum_j \alpha_j x_j \cdot x + b$$

where  $\{x_j\}$  are the set of support vectors and the parameters  $\alpha_j$  and  $b$  have been determined by solving a quadratic problem [16]. The linear SVM can be extended to various non-linear variants; details can be found in [16, 17]. In our experiments, the RBF kernel SVM outperformed other non-linear SVM kernels, hence we are reporting our recognition results based on the RBF kernel only.

## VII. EXPERIMENTAL SETTINGS

In the field of signature verification, there is lack of a publicly available signature database. The quality of available databases also varies, as there has been no standard collection protocol. Besides, it is very costly to create a large corpus with different types of forgeries, especially skilled forgeries.

For the experiments in the proposed research, our own database described in Section IV was used. For each signature set, an SVM was trained with 12 randomly chosen

genuine signatures. The negative samples for training (random forgeries) were the genuine signatures of 63 other signature sets. Four signatures were taken from each set. In total, there were  $63 \times 4 = 252$  random forgeries employed for training. For testing, the remaining genuine signatures and 15 randomly selected skilled forgeries of the signature set being considered were employed. The random forgeries for testing were the 36 genuine signatures from the remaining signature sets, which were not involved in the training phase. From each set, one genuine signature was randomly selected. The training and testing process was repeated 30 times and the error rates of the experiments were averaged before being reported so that the results obtained were more reliable. The number of samples for training and testing for experimentation with each signature set are shown in Table 1.

TABLE 1. NUMBER OF SIGNATURE SAMPLES USED FOR EXPERIMENTATION

	Genuine Signature	Random Forgeries	Skilled Forgeries
<b>Training</b>	12	252	n/a
<b>Testing</b>	12	36	15

## VIII. RESULTS AND DISCUSSION

For training and testing of the system, 5400 Bangla signatures were employed.

Using the Gaussian Grid feature, an AER as low as 9.6% was obtained. At this operational point, the FRR, FAR for random forgeries (FAR1), and FAR for skilled forgeries (FAR2) were 10.77%, 0.05%, and 8.42%, respectively. The AER obtained in this research is 4.33% lower than the AER obtained with the GPDS-160 database reported in [9] despite the smaller number of random forgeries (252 vs. 400) employed in the training process.

TABLE 2. RESULTS BASED ON DIFFERENT TYPES OF FORGED SIGNATURES

<b>Sigma</b>	<b>FRR (%)</b>	<b>FAR1</b>	<b>FAR2</b>
<b>2.9e-2</b>	<b>9.96</b>	<b>0.06</b>	<b>9.40</b>
<b>3.0e-2</b>	<b>10.14</b>	<b>0.06</b>	<b>9.17</b>
<b>3.1e-2</b>	<b>10.35</b>	<b>0.06</b>	<b>9.19</b>
<b>3.2e-2</b>	<b>10.55</b>	<b>0.05</b>	<b>8.64</b>
<b>3.3e-2</b>	<b>10.77</b>	<b>0.05</b>	<b>8.42</b>
<b>3.4e-2</b>	<b>11.03</b>	<b>0.05</b>	<b>8.19</b>
<b>3.5e-2</b>	<b>11.27</b>	<b>0.05</b>	<b>7.99</b>
<b>3.6e-2</b>	<b>11.57</b>	<b>0.04</b>	<b>7.80</b>
<b>3.7e-2</b>	<b>11.87</b>	<b>0.04</b>	<b>7.57</b>

Although the FRR and the FAR2 rates changed rapidly with sigma, the AER rate was more stable. The error rate obtained from our experiments is shown in Table 2. These encouraging results demonstrate the robustness of the Gaussian Grid feature extraction technique for signature verification. The results obtained in our experiments are shown by a graph in Fig. 6.

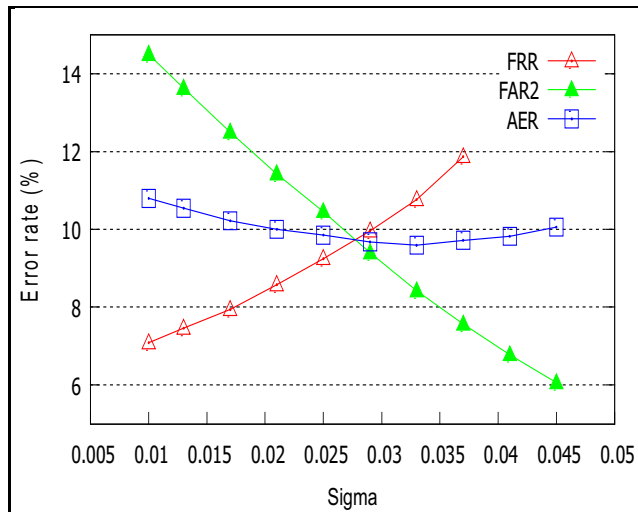


Figure 6. Error rates based on the value of sigma

## IX. CONCLUSIONS AND FUTURE WORK

This paper presents an investigation of the performance of a signature verification system involving Bangla off-line signatures. The Gaussian grid feature and SVM classifiers were employed and encouraging results were obtained. To the best of our knowledge, a Bangla signature database has never been used for the task of signature verification. The proposed off-line verification scheme is the first investigation for Bangla signatures in the field of off-line signature verification. In the near future, we plan to extend our work considering more datasets of Bangla off-line signatures.

## ACKNOWLEDGEMENT

The experimental results reported in this work were obtained using the Griffith University HPC Cluster – V20z.

## REFERENCES

[1] S. Chen, and S. Srihari, "Use of Exterior Contour and Shape Features in Off-line Signature Verification", Proceedings of 8th ICDAR, pp. 1280-1284, 2005.  
 [2] M.A. Ferrer, J.B. Alonso and C. M. Travieso, "Off-line Geometric Parameters for Automatic Signature Verification Using Fixed-Point Arithmetic" IEEE PAMI, Vol.27, no.6, pp. 993-997, 2005.  
 [3] S. Madabusi, V. Srinivas, S. Bhaskaran and M. Balasubramanian, "On-line and off-line signature verification using relative slope algorithm", International Workshop on Measurement Systems for Homeland Security, pp. 11-15, 2005.

[4] R. Plamondon and G. Lorette, "Automatic signature verification and writer identification - the state of the art", Pattern Recognition, 22, pp. 107-131, 1989.  
 [5] A. Kholmatov, and B. Yanikoglu, "Identity Authentication using improved online signature verification method", PRL, pp. 2400-2408, 2005.  
 [6] M. Kalera, S. Srihari, and A. Xu. "Offline signature verification and identification using distance statistics", IJPRAI, pp.1339-1360, 2004.  
 [7] I. Pottier and G. Burel, "Identification and Authentication of Handwritten Signatures with a Connectionist Approach", In Proc. IEEE Conf. On Neural Networks, pp. 2948-2951, 1994.  
 [8] B. Fang, C.H. Leung, Y.Y. Tang, K.W. Tse, P.C.K. Kwok and Y.K. Wong, "Off-line signature verification by the tracking of feature and stroke positions", Pattern Recognition, 36, pp. 91-101, 2003.  
 [9] Katrin Franke, "Analysis of Authentic Signatures and Forgeries" In Proc. IWCF, pp 150-164, 2009.  
 [10] E. Justino, E. Bortolozzi, R Sabourin, "Off-line Signature Verification Using HMM for Random, Simple and Skilled Forgeries." Proceedings of 7th ICDAR, pp. 105-110, 2001.  
 [11] V. Nguyen, M. Blumenstein, G. Leedham, "Global Features for the Off-Line Signature Verification Problem", Proceedings of 10th ICDAR, pp.1300-1304, 2009.  
 [12] B.Schafer, S.Viriri, "An Off-Line Signature Verification System", (ICSIPA- 2009), pp.95-100, 2009.  
 [13] J. Coetzer, B. Herbst, and J. D. Preez, "Off-line signature verification using the discrete radon transform and a hidden markov model", EURASIP Journal on Applied Signal Processing, 4, 559-571, 2004.  
 [14] B. B. Chaudhuri and U. Pal, "An OCR system to read two Indian language scripts: Bangla and Devnagari (Hindi)", Proceedings of 4th ICDAR, pp. 1011-1015, 1997.  
 [15] V. Nguyen and M. Blumenstein, "An application of the 2D Gaussian filter for signature verification problem", Proceedings of the 13th ICDAR, pp. 339-343, 2011.  
 [16] V.Vapnik, "The Nature of Statistical Learning Theory", Springer Verlag, 1995.  
 [17] C. Burges, "A Tutorial on support Vector machines for pattern recognition", Data mining and knowledge discovery, pp.1-43, 1998.  
 [18] H. Lv, W. Wang, C. Wang, Q. Zhuo, "Off-line chinese signature verification based on support vector machines," PRL, pp. 2390-2399, 2005.