

Global Features for the Off-Line Signature Verification Problem

Vu Nguyen, Michael Blumenstein
School of Information & Communication Technology
Griffith University, Australia
 {Vu.Nguyen2, M.Blumenstein}@griffith.edu.au

Graham Leedham
School of Science & Technology
University of New England, Australia
 Graham.Leedham@une.edu.au

Abstract

Global features based on the boundary of a signature and its projections are described for enhancing the process of automated signature verification. The first global feature is derived from the total 'energy' a writer uses to create their signature. The second feature employs information from the vertical and horizontal projections of a signature, focusing on the proportion of the distance between key strokes in the image, and the height/width of the signature. The combination of these features with the Modified Direction Feature (MDF) and the ratio feature showed promising results for the off-line signature verification problem. When being trained using 12 genuine specimens and 400 random forgeries taken from a publicly available database, the Support Vector Machine (SVM) classifier obtained an average error rate (AER) of 17.25%. The false acceptance rate (FAR) for random forgeries was also kept as low as 0.08%.

1. Introduction

Handwritten signatures have long been accepted as an official means to verify personal identity for legal purposes on such documents as cheques, credit cards, contracts and wills. The handwritten signature is therefore well established and accepted as a behavioural 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. As a consequence, research into signature verification has been vigorously pursued for several decades, particularly with reference to off-line verification [1, 2].

Off-line verification refers to when the signature is only available as a static image, typically obtained after it has been written on paper using a variety of writing instruments, with no reference to the sequence and timing of the pen-

strokes, which created the signature. When the sequence of the pen-strokes is available the process is referred to as on-line signature verification.

The off-line signature verification problem is more challenging due to the lack of valuable behavioural information about how the person created the signature in terms of pen-point velocity and accelerations, writing pressure and stroke sequence. Access to this on-line information during the training phase, has been demonstrated to result in improved performance in off-line signature verification systems [3].

Like any other pattern recognition scheme, one crucial aspect of a signature verification system includes appropriate feature extraction procedures. Thus, new feature extraction techniques are being extensively explored. The survey by Weiping *et al.* [1] summarises some previously investigated features and approaches. In this paper we report our latest results in the pursuit of new global and local features for addressing the off-line signature verification problem.

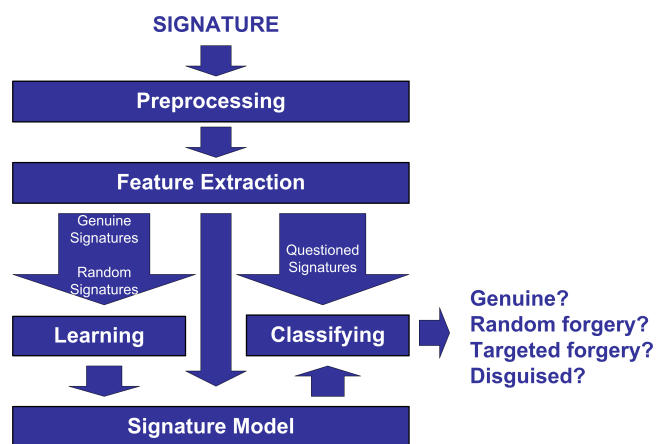


Figure 1. Automatic Offline Signature Verification System

2. Methodology

A general automated off-line signature verification system that employs only genuine specimens and other random signatures in the training phase is depicted in Figure 1. After being preprocessed through conversion to a portable bitmap (PBM) format, the boundary of the signature image is extracted. The Modified Direction Feature (MDF) and other global features are then extracted from the boundary in the feature extraction stage.

To perform learning and classification tasks for signature verification, machine learning techniques such as Artificial Neural Networks, Support Vector Machines, or a threshold decision method can be employed. The signature model is constructed in the learning phase using predefined parameters, which are invoked in the classification phase. It is hypothesised that there exist a set of parameters that is universal for every writer.

3. Feature Extraction

Features extraction plays an important role in a signature verification system, and the features extracted can be categorized as global or local features. Global features treat patterns as a whole whilst local features are extracted from a portion or a limited area of the pattern. Global features tend to be less sensitive to variation or noise whilst local features provide more detailed information.

A suitable combination of global and local features has previously been found to improve a classifier's ability to recognize forgeries and to tolerate intrapersonal variances [1]. In this research, the combination of the Modified Direction Feature with three other global features is reported.

3.1. Modified Direction Feature

The Modified Direction Feature (MDF) feature extraction technique [4] employs the location of transitions from background to foreground pixels and the direction at transitions in the vertical and horizontal directions of the boundary representation of an object. At each transition, the Location of the Transition (LT) and the Direction Transition (DT) values are recorded.

The LT value is calculated by dividing the position value where a transition occurs by the distance across the entire image in a particular direction, whilst the DT is obtained by examining the stroke direction of an object's boundary at the position where the transition occurred. Finally, the LT and DT values are locally averaged to reduce the feature vector size. The local averaging process also makes the MDF feature extraction technique more tolerant to local noise and intrapersonal variations.

3.2. Features from Energy Information

The first global feature investigated in this work was derived from the total energy a writer uses to create their signature. It is hypothesised that the planned execution of the signature uses the same amount of energy, and this information can be extracted from the trajectory of a signature using its digital image. Although it is difficult to precisely estimate the amount of energy used between any two points on the signature's trajectory, it is of interest to determine whether information extracted from the energy of special segments (e.g. cusp to cusp, end to end) could increase verification accuracy. Recent work of Qiao *et al.* [3] showed that the signature trajectory could effectively be recovered using the information from on-line registration. As a consequence, it is expected that an energy-based feature could enhance this approach.

For simplicity, in this work, we only extract energy information from the boundary of the whole signature image. This energy information was decomposed into horizontal and vertical components. The features extracted are the values from the signature width divided by horizontal energy (e_h), signature height divided by vertical energy (e_v), and $\min(e_h, e_v)/\max(e_h, e_v)$. The following is the pseudo code for the extraction of these values:

```
function Energy(binary image im)
1:  $Sum_h \leftarrow 0$  {horizontal energy}
2:  $Sum_v \leftarrow 0$  {vertical energy}
3:  $Sum_L \leftarrow 0$  {energy from the left diagonal}
4:  $Sum_R \leftarrow 0$  {energy from the right diagonal}
5:  $h \leftarrow im_{height}$ 
6:  $w \leftarrow im_{width}$ 
7: Mark all black pixels of im as unvisited
8: for each black pixel p of im do
9:   if p is not visited then
10:    track(p)
11:   end if
12: end for
13:  $e_v \leftarrow h/Sum_v$ 
14:  $e_h \leftarrow w/Sum_h$ 
15:  $e_{hv} \leftarrow \min(Sum_v, Sum_h)/\max(Sum_v, Sum_h)$ 
16:  $e_{lr} \leftarrow \min(Sum_L, Sum_R)/\max(Sum_L, Sum_R)$ 
17:  $e_{hv} \leftarrow Sum_v + Sum_h$ 
18:  $e_{LR} \leftarrow Sum_L + Sum_R$ 
19:  $e_r \leftarrow \min(e_{hv}, e_{LR})/\max(e_{hv}, e_{LR})$ 
20: return  $\{e_v, e_h, e_{hv}, e_{lr}, e_r\}$ 
end function
```

```
procedure track(pixel p)
```

```
1: mark p as visited
2: for each neighbour pixel pNb of p do
3:   if pNb is not visited then
```

```

4:   {subscripts y and x mean the row and column of a
      pixel}
5:    $Sum_v \leftarrow Sum_v + |pNb_y - p_y|$ 
6:    $Sum_h \leftarrow Sum_h + |pNb_x - p_x|$ 
7:   if  $(pNb_y - p_y) * (pNb_x - p_x) = 1$  then
8:      $Sum_L \leftarrow Sum_L + 1$ 
9:   end if
10:  if  $(pNb_y - p_y) * (pNb_x - p_x) = -1$  then
11:     $Sum_R \leftarrow Sum_R + 1$ 
12:  end if
13:  track(pNb)
14:  break
15: end if
16: end for
end procedure

```

3.3. The Maxima Feature

Information from horizontal and vertical projections has been adopted by researchers for image alignment [5], and profiling signatures for off-line signature verification [6, 7]. From our observations, handwriting boundary projections contain valuable information about key strokes, and near straight curves, which are horizontal or vertical. It is suggested that the proportion of the distance between two main strokes, either both horizontally, or both vertically, and the height and width of a writer's signature remain relatively stable among signature specimens and thus can be used as a global feature in the offline signature verification problem.

The formula of the 'maxima' feature is defined as:

$$\left(\frac{|col_{max}^1 - col_{max}^2|}{W}, \frac{|row_{max}^1 - row_{max}^2|}{H} \right) \quad (1)$$

Where W and H are the width and height of the signature image, row_{max}^1 , row_{max}^2 , col_{max}^1 , and col_{max}^2 are the indexes of the columns and rows that have the highest number of black pixels among the rows or columns.

3.4. Ratio Feature

The next global feature we used in conjunction with the MDF, the energy-based feature and the Maxima feature was the Ratio feature, which employed the width and height information of the rectangular box that encloses the signature pattern. The Ratio feature has been widely used by researchers in cursive handwritten character recognition and signature verification [4, 8]. In our previous work with neural networks [9], the Ratio feature (R1) was calculated using the following formula in order to generate a feature value ranging from 0 to 1:

$$Ratio = \frac{\arctan(width/height)}{\pi/2} \quad (2)$$

Table 1. Experimental Settings

	Phase	Genuine	Random	Targeted
Setting I	Training	12	400	-
	Testing	12	59	15
Setting II	Training	20	400	-
	Testing	4	59	15

However, our experiments showed that a more straightforward calculation of the Ratio feature would increase the accuracy. The alternative Ratio feature value is calculated by dividing the minimum by the maximum of the width and height values. The following is the alternative formula for the Ratio feature (R2):

$$Ratio = \frac{\min(width, height)}{\max(width, height)} \quad (3)$$

4. Experimental Settings

To enable result comparison with other work, the publicly available gpdsSIGNATURE [10] handwritten signature database was employed in our research. This large corpus consists of 160 signature sets with 24 genuine and 30 targeted forgeries in each set.

It is essential to select a reasonable number of genuine samples for the training process to construct a signature model that is tolerant to intrapersonal variances and effectively rejects imitations [11]. In each test, 12 genuine signatures and 400 random forgeries were employed for training. Whilst the random forgeries are easier to collect, the random forgeries were chosen from 100 randomly selected writers, four genuine signatures from each. To represent random forgeries for testing, 59 genuine signatures were chosen from the remaining 59 writers. Table 1 summarizes the sample configuration for the training and testing phases. For each signature set, the experiment is repeated 30 times in order to obtain a more stable result.

Support Vector Machines (SVMs) [12] were employed to construct the signature models. The choice of a suitable kernel for SVM training varies between classification problems and feature extraction techniques. Within the area of off-line signature verification, Ferrer *et al.* [10] and Lv *et al.* [13] obtained better results with the Radial Basis Function (RBF) kernel. Meanwhile, Justino *et al.* [14] achieved their best results with the linear kernel. Results from our previous research [9] suggested that the MDF and its variants performed best with the RBF kernel. Consequently, the RBF kernel was chosen and the experiments were conducted using SVM^{light} v6.01 [15].

5. Experimental Results

A signature verification system can be challenged by forgeries with varied levels of skill. It is much easier to compare the performance of different approaches against random forgeries than against targeted forgeries due to the unavailability of a standard signature corpus. As can be seen in Table 2, a false acceptance rate (FAR) of 0.08% for random forgeries was obtained with the proposed feature combination. This figure is comparable to the results of other researchers [5, 10, 14, 16, 17] and can be explained by the LT values of the MDF, which were effectively generalized by the local averaging process.

It is natural that a system has a substantially higher FAR for targeted forgeries than for simple forgeries, or random forgeries. However, some researchers observed an opposite trend regarding the proportion of random forgery errors as opposed to targeted forgery errors on other databases with different approaches [14, 17].

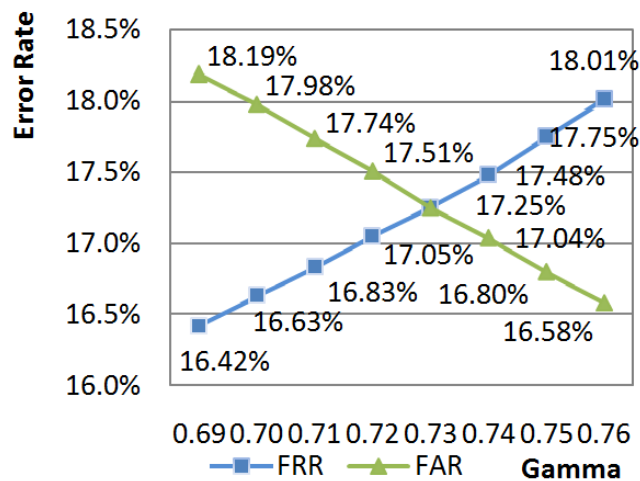


Figure 2. The FRR and FAR values obtained with different values of gamma when the RBF kernel was employed

When the proposed system was challenged by targeted forgeries, the lowest average error rate of 17.25% was obtained with the combination of MDF and the global features. Figure 2 shows the false rejection rate (FRR) and false acceptance rate (FAR) values for different settings of sigma used in the RBF kernel of SVM. This result is comparable to the AER of 17.17% as reported by Zhang *et al.* [18] (although they employed a smaller database with 80 sets of signatures). It is worth mentioning that, to obtain our result, only 12 genuine signatures were used for the training process as compared to 20 genuine signatures in Zhang *et al.*'s research. In experiments using Setting II, our approach obtained a superior AER of 15.77%. However, training with a

larger number of genuine signatures reduces the feasibility of many systems.

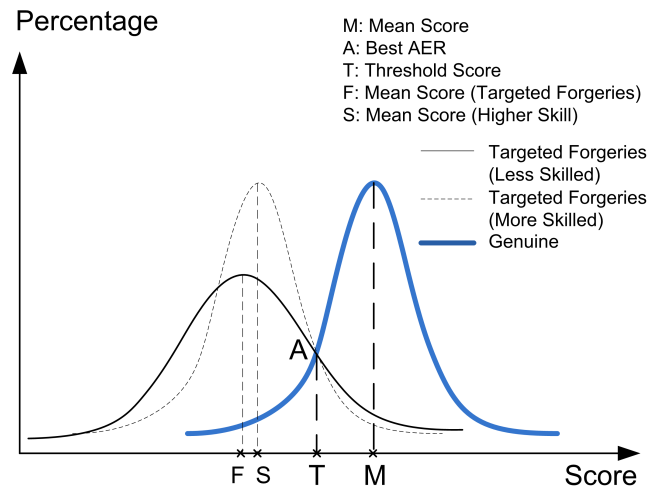


Figure 3. The distribution of score for genuine and targeted forgeries of varied skill

If it is hypothesised that the score for the genuineness of genuine signatures would follow a normal distribution regardless of which approach was employed, the distribution of the score curve for forged signatures would likely take on a similar shape. This is due to the fact that the genuine specimens were randomly selected for imitation. Considering the above assumptions, different approaches should either have their best AER values approximating their equal error rate (EER, where FAR equals FRR) or the database employed is less representative (Figure 3 - Less Skilled curve). Using the above arguments, the FRR and the targeted forgeries' FAR values obtained in this work support the premise that the quality of the gpdS SIGNATURE database is superior to other signature corpuses (Zhang *et al.* [18] exhibiting a 2.74% difference, Bansal *et al.* [19] with a 10.38% difference).

6. Conclusions

The performance of the MDF feature extraction technique in conjunction with three other simple global features has been investigated. With the proposed feature set, an automated off-line signature verification system obtained an AER of 17.25% when 12 genuine signatures were used in the training process.

The results obtained using the energy-based features described in this paper (in conjunction with other global and local features) encourage further investigation in automatic signature verification using the energy information extracted from the signature trajectory. Although it is not

Table 2. The error rates of feature combinations obtained using Setting I (FAR1 and FAR2 are false acceptance rates for random and targeted forgeries respectively)

Features	FRR	FAR1	FAR2	AER
MDF	18.14%	0.11%	18.32%	18.23%
MDF R1	18.42%	0.13%	18.29%	18.36%
MDF R1 M	17.99%	0.13%	18.82%	18.41%
MDF R2	18.56%	0.10%	17.49%	18.03%
MDF E	18.01%	0.09%	16.98%	17.50%
MDF E R1	18.09%	0.09%	16.90%	17.50%
MDF E R2	17.77%	0.08%	16.96%	17.37%
MDF E R2 M	17.25%	0.08%	17.25%	17.25%

easy to recover dynamic information directly from off-line images [3], energy information extracted from certain parts of a signature may help increase verification accuracy. As SVMs can still perform well when some of the features are missing [15], energy-based features could be extracted at the micro-level from an uncertain trajectory, rearranged, and could assist in lowering the verification error rate.

Future work will include (1) the automation of off-line handwritten signature trajectory recovery; (2) the extraction of energy information from different parts of the signature; (3) appropriately organizing the extracted information for use with SVMs.

References

- [1] H. Weiping, Y. Xiufen, and W. Kejun, "A survey of off-line signature verification," in *Intl. Conf. on Intelligent Mechatronics and Automation*, pp. 536–541, Aug. 26–31 2004.
- [2] R. Plamondon and S. Srihari, "Online and off-line handwriting recognition: a comprehensive survey," *PAMI, IEEE Trans. on*, vol. 22, pp. 63–84, Jan 2000.
- [3] Y. Qiao, J. Liu, and X. Tang, "Offline signature verification using online handwriting registration," in *CVPR, IEEE Conf. on*, pp. 1–8, June 2007.
- [4] M. Blumenstein, X. Y. Liu, and B. Verma, "A modified direction feature for cursive character recognition," in *Proc. IEEE Intl. Joint Conf. on Neural Networks IJCNN*, vol. 4, pp. 2983–2987, 25–29 July 2004.
- [5] A. Piyush Shanker and A. N. Rajagopalan, "Off-line signature verification using DTW," *Pattern Recognition Letters*, vol. 28, no. 12, pp. 1407–1414, 2007.
- [6] 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*, vol. 36, no. 1, pp. 91–101, 2003.
- [7] T. Wei and Q. Yizheng, "Off-line chinese signature verification based on optimal matching of projection profiles," in *Proc. of the 6th World Congress on Intelligent Control and Automation WCICA*, vol. 2, (Dalian, China), pp. 10240–10244, June 21 - 23, 2006.
- [8] F. Camastra, "A SVM-based cursive character recognizer," *Pattern Recognition*, vol. 40, pp. 3721–3727, 2007.
- [9] V. Nguyen, M. Blumenstein, V. Muthukkumarasamy, and G. Leedham, "Off-line signature verification using enhanced modified direction features in conjunction with neural classifiers and support vector machines," in *Proc. 9th Intl Conf on Document Analysis and Recognition ICDAR*, vol. 2, pp. 734–738, 23–26 Sept. 2007.
- [10] M. A. Ferrer, J. B. Alonso, and C. M. Travieso, "Offline geometric parameters for automatic signature verification using fixed-point arithmetic," *PAMI, IEEE Trans. on*, vol. 27, no. 6, pp. 993–997, 2005.
- [11] R. Plamondon and G. Lorette, "Automatic signature verification and writer identification - the state of the art," *Pattern Recognition*, vol. 22, no. 2, pp. 107–131, 1989.
- [12] V. N. Vapnik, *Statistical Learning Theory*. John Wiley & Sons, 1998.
- [13] H. Lv, W. Wang, C. Wang, and Q. Zhuo, "Off-line chinese signature verification based on support vector machines," *Pattern Recognition Letters*, vol. 26, no. 15, pp. 2390–2399, 2005.
- [14] E. J. R. Justino, F. Bortolozzi, and R. Sabourin, "A comparison of SVM and HMM classifiers in the off-line signature verification," *Pattern Recognition Letters*, vol. 26, no. 9, pp. 1377–1385, 2005.
- [15] T. Joachims, *Advances in Kernel Methods - Support Vector Learning*, ch. Making large-Scale SVM Learning Practical. MIT-Press, 1999.
- [16] J. F. Vargas, M. A. Ferrer, C. M. Travieso, and J. B. Alonso, "Off-line handwritten signature GPDS-960 corpus," in *Proc. Ninth International Conference on Document Analysis and Recognition ICDAR*, vol. 2, pp. 764–768, 23–26 Sept. 2007.
- [17] L. S. Oliveira, E. Justino, and R. Sabourin, "Off-line signature verification using writer-independent approach," in *Proc. International Joint Conference on Neural Networks IJCNN*, pp. 2539–2544, 12–17 Aug. 2007.
- [18] Z. Tai-Ping, F. Bin, X. Bin, H.-X. Chen, M. Chen, and Y.-Y. Tang, "Signature envelope curvature descriptor for offline signature verification," in *Intl. Conf. on Wavelet Analysis and Pattern Recognition ICWAPR*, vol. 3, pp. 1262–1266, 2007.
- [19] A. Bansal, D. Garg, and A. Gupta, "A pattern matching classifier for offline signature verification," in *Proc. First International Conference on Emerging Trends in Engineering and Technology ICETET*, pp. 1160–1163, 16–18 July 2008.