Offline Handwritten Signature Verification using Radial Basis Function Neural Networks

George Azzopardi
St. Martin's Institute of IT
Email: geazzo@gmail.com

Kenneth P. Camilleri
Dept of Systems and Control Engineering
University of Malta
and
St. Martin's Institute of IT
Email: kpcami@eng.um.edu.mt

Abstract

This study investigates the effectiveness of Radial Basis Function Neural Networks (RBFNNs) for Offline Handwritten Signature Verification (OHSV). A signature database is collected using intrapersonal variations for evaluation. Global, grid and texture features are used as feature sets. A number of experiments were carried out to compare the effectiveness of each separate set and their combination. The system is extensively tested with random signature forgeries and the high recognition rates obtained demonstrate the effectiveness of the architecture. The best results are obtained when global and grid features are combined producing a feature vector of 592 elements. In this case a Mean Error Rate (MER) of 2.04% with a False Rejection Rate (FRR) of 1.58% and a False Acceptance Rate (FAR) of 2.5% are achieved, which are generally better than those reported in the literature.

Index Terms

Offlline Signature Verification, RBFNN, Neural Classifiers, Signature Features

1. Introduction

Personal verification and identification is an actively growing area of research and development. Different biometrics have been used to authenticate the identity of an individual, which can be categorised as physiological (face, iris, fingerprint, odour) and behavioural traits (signature, voice) among other characteristics [1].

This study concerns with handwritten signatures which are considered a behavioural biometric, and its acceptance is widespread socially and legally as a means of authentication [1].

1.1. Motivation

According to Gori and Scarselli [2], although Multi-Layer Perceptrons (MLPs) are very good at performing discriminative classification between patterns of well-defined classes, they are not adequate for applications requiring a reliable rejection. It was suggested that other architectures, such as autoassociator-based classifiers and RBFNNs are more suitable for handling outliers.

Baltzakis and Papamarkos [3] established the viability of a two-stage neural networks signature verification architecture based on a first-stage of MLPs followed by an RBFNN layer. In this study, the viability of a single stage RBFNN for OHSV is investigated.

2. Methodology

The methodology of this study involves data acquisition, preprocessing, feature extraction, signature comparison process, and performance evaluation which are discussed below.

2.1. Data Acquisition

A signature database of 2492 signatures is collected from 65 different signers which are scanned using a resolution of 300 dpi and stored as a BMP file type (no compression used). All signature sheets are manually cropped using a photo editor to separate them as individual images.

The data acquisition process involved:

1) Acquisition of a total of 40 signatures from each author; 25 on blank sheets and 15 in provided random sized rectangles
2) Acquiring the signatures on 5 different days (when possible); 5 on blank sheets and 3 in provided rectangles on each day
3) Using 8 different pens which vary in colour (black, blue, red and green) and type (ball point, normal pen and fountain pen)
4) Signers asked to use as much as intra personal variations as possible

The group of 65 persons contributing to this exercise comprises mainly of family members, friends and work colleagues having different background; education level, language, age and region. They represent a wide variety of signature styles, from completely incomprehensible line strokes to clear and tidy handwriting.

2.2. Pre-Processing

The pre-processing stage follows the four steps proposed in [3]: data area cropping, width normalization, binarization and skeletonization.

Noise reduction is not required since the signatures are acquired on white sheets.

2.2.1. Data Area Cropping. Initially, the original 24-bit colour image is segmented from the background to remove the white space surrounding the signature using the segmentation method of vertical and horizontal projections [4].

2.2.2. Width Normalization. The cropped image is scaled using bicubic interpolation to a constant width, keeping the aspect ratio fixed.

2.2.3. Binarization. The 24-bit color signature is converted to grayscale and then binarized using a histogram-based binarization.

2.2.4. Skeletonization. The algorithm proposed by [5] is used in order to reduce data storage without losing the structural information of the image as well as to facilitate the extraction of morphological features from digitised patterns.

2.3. Feature Extraction and Selection

The choice of powerful set of features is essential in optical recognition systems. The selected features must be suitable for the application of the applied classifier.

Feature extraction is divided into 3 sets of features including global, grid and texture features.

2.3.1. Global Features. Global features provide information about the entire structure of the signature. The proposed set of global features by [6] are extracted from the skeletonized signature in this study.

1) *Signature Height* - The height of the signature (in pixels), after width normalization, is considered as a global characteristic.
2) *Height-to-Width Ratio* - The proportionality rate of the skeleton signature image. This is calculated by dividing the height with the width of the signature.
3) *Pure Width* - The width of the skeleton signature with horizontal blank spaces removed.
4) *Pure Height* - The height of the skeleton signature with vertical blank spaces removed.
5) *Image Area* - The number of black pixels in the skeleton signature.
6) *Maximum Horizontal Projection* - The skeleton signature image is scanned vertically and each time calculating the horizontal projection. The horizontal projection represents the number of black pixels in the current row. Then, the row containing the maximum number of black pixels is taken to represent the maximum horizontal projection.
7) *Maximum Vertical Projection* - Similarly to above, the maximum vertical projection represents the maximum number of black pixels in a column when scanning the skeleton signature image horizontally.
8) *Vertical Projection Peaks* - This represents the number of local maxima of the vertical projection histogram. The vertical projection histogram is the frequency of black pixels for each column of the skeleton signature image.
9) *Horizontal Projection Peaks* - Similarly to above, this represents the number of local maxima of the horizontal projection histogram.
10) *Vertical Centre of Gravity* - Vertical centre of gravity is a measurement indicating the vertical location of the signature image based on the horizontal projections P_h and is calculated as:

$$C_v = \frac{\sum_i i \times P_h[i]}{\sum_i P_h[i]} \qquad (1)$$

11) *Horizontal Centre of Gravity* - Similarly, horizontal centre of gravity is a measurement indicating the horizontal location of the signature image based on the vertical projections P_v and is calculated as:

$$C_h = \frac{\sum_i i \times P_v[i]}{\sum_i P_v[i]} \qquad (2)$$

12) *Baseline Shift* - It is the difference between the vertical centres of gravity of the left and right part of the skeleton signature image. It is calculated by splitting the signature image vertically into two halves and calculate the vertical centre of gravity for each half; C_L and C_R for left and right half respectively. Then the baseline shift is defined as $BS = C_L - C_R$.

13) *Global Slant Angle* - It is the overall direction of line strokes in the skeleton signature. The original signature is rotated from $-45°$ to $45°$ in steps of $5°$. For each rotation, the original signature is first pre-processed followed by counting the number of vertical 3-pixels connections from the rotated skeleton image. The global slant angle is the angle having the maximum number of vertical 3-pixels connections.

14) *Local Slant Angle* - It is the angle of dominant strokes in the skeleton image. The original image is rotated similarly as mentioned above and for each angle the vertical projection histogram is calculated and the highest 70 projections are summed up. The local slant angle is the angle having the maximum sum of the top 70 projections.

15) *Number of Edge Points* - According to [3] an edge point is a black pixel having only one 8-neighbour.

16) *Number of Cross Points* - A cross point is a connected component in which each pixel has at least three 8-neighbours. The below figure illustrates six different cross points.
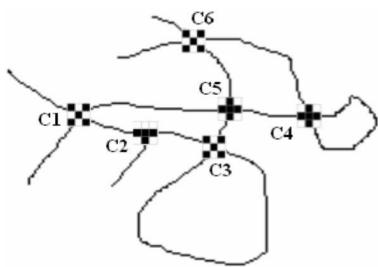
Figure 1. Cross Points

17) *Number of Closed Loops* - The number of closed regions in a skeletonised image and is computed as described in [3]. The following image shows a signature with seven closed loops.
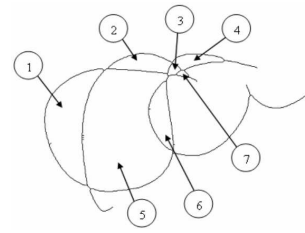
Figure 2. Closed Loops

2.4. Grid Features

As explained in [7] grid segmentation is a technique used for signature detail analysis. A virtual grid of 12×8 segments is superimposed on the skeleton image and the following features are calculated for each segment.
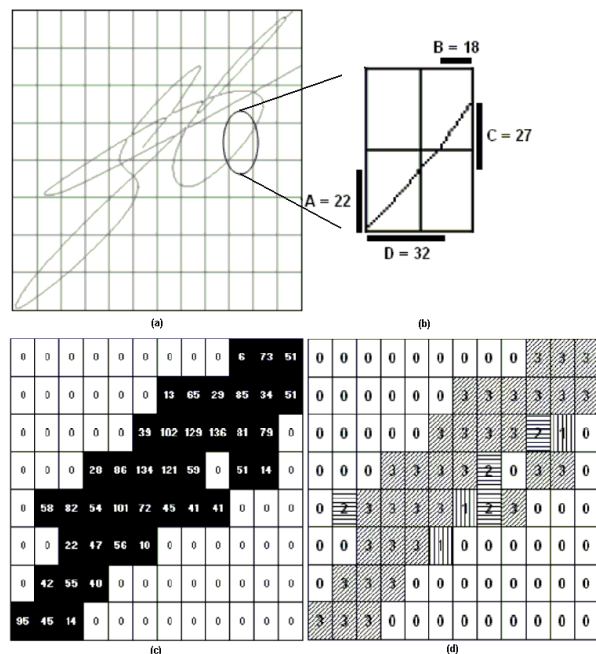
Figure 3. (a) Skeleton Signature; (b) Pixel Distribution; (c) Pixel Density; (d) Predominant Axial Slant

2.4.1. Pixels Density. This is the number of black pixels within each segment (*see* Fig. 3c).

2.4.2. Pixels Distribution. It represents the pixel geometric distribution in a cell. The black pixels are projected in four side-line cell sensors from the central axis of the cell. Each sensor provides a numerical value corresponding to the total of the projected pixels as shown in Fig. 3b.

2.4.3. Predominant Axial Slant. The predominant axial slant is a value representing the predominant inclination in each cell. For each cell the number of three pixels connections is calculated against the following templates.
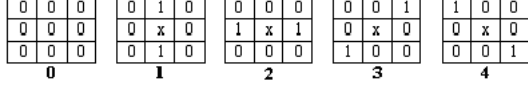
0	0	0		0	1	0		0	0	0		0	0	1		1	0	0
0	0	0		0	x	0		1	x	1		0	x	0		0	x	0
0	0	0		0	1	0		0	0	0		1	0	0		0	0	1
	0				**1**				**2**				**3**				**4**	

Figure 4. Predominant Axial Slant Template

The template which features most within the cell is the predominant axial slant (*see* Fig. 3d).

2.5. Texture Features

Similarly a 12×8 grid is used for texture analysis which is performed on the skeleton signature image. A 2×2 co-occurrence matrix is used to describe the transition of black and white pixels and is defined as [8]:

$$\mathrm{P}_{\vec{d}[i,j]} = \begin{bmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{bmatrix} \qquad (3)$$

where p_{00} is the number of times a pair of two white pixels, separated by $\vec{d}$, occur, p_{01} is the number of times a white pixel is followed by a black separated by $\vec{d}$, p_{10} is the number of times that a black pixel is followed by a white pixel separated by $\vec{d}$, and p_{11} is the number of times two black pixels separated by $\vec{d}$, occur.
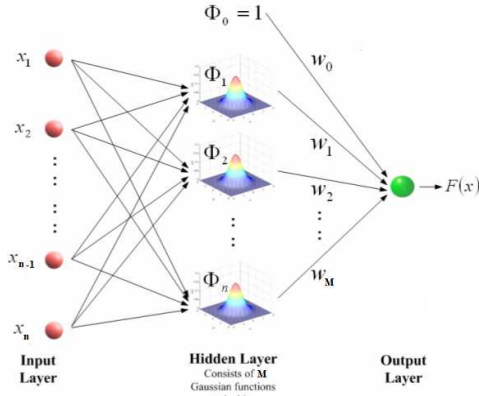
3. Classification

Figure 5. RBFNN Single Layer Architecture

An RBFNN single layer architecture based on Gaussian functions is used for every signature model (i.e. the set of signature features of the same signer) as shown in Fig. 5.

Define a Gaussian function as $\varphi_k(\vec{x}, \vec{\mu}) = \exp \frac{-\|\vec{x}-\vec{\mu}\|^2}{2\sigma^2}$. Taking each signature model as a single cluster, the centroid of model k is denoted by $\vec{\mu}_k$ and computed from the cluster feature vectors $\vec{x}_{ik}$. The respective model variance σ_k^2 is determined as:

$$\sigma_k^2 = \frac{1}{n_k} \sum_{i=1}^{n_k} (\|\vec{x}_{ik} - \vec{\mu}_k\|)^2 \qquad (4)$$

where $\|\vec{x}_{ik} - \vec{\mu}\|$ is the Euclidean distance and n_k is the number of data points in the cluster k. The RBF network for signature model k is defined as:

$$\begin{bmatrix} 1 & \varphi_{11} & \cdots & \varphi_{1M} \\ \vdots & \ddots & \vdots & \vdots \\ 1 & \varphi_{n_k 1} & \cdots & \varphi_{n_k M} \end{bmatrix} \begin{bmatrix} \omega_0 \\ \omega_1 \\ \vdots \\ \omega_M \end{bmatrix} = \begin{bmatrix} d_1 \\ \vdots \\ d_{n_k} \end{bmatrix}$$

This matrix is called the interpolation matrix, where M denotes the number of signature models. The first column represents the bias vector which is set to 1. The above can be conveniently changed to a vector equation as:

$$\Phi \vec{\omega} = \vec{d} \qquad (5)$$

where Φ is the interpolation matrix, $\vec{\omega}$ is the weight vector and $\vec{d}$ is the desired response vector.

The weight values that minimise the error $\Phi \vec{\omega} - \vec{d}$ can be obtained using a pseudo-inverse technique as:

$$\vec{\omega} = (\Phi^T \Phi)^{-1} \Phi^T \vec{d} \qquad (6)$$

3.1. Normalization: Global Features

Due to the different units of the 17 global features explained in section 2.3.1, normalization is required to eliminate the units and to project the values in the range $[0, 1]$. Hence, the global features of each signature are represented by a 17×1 feature vector.

3.2. Vector Quantization

For both Grid and Texture features, Vector Quantization is used to convert the vectors into a symbol sequence. As suggested by [9] since the training database works with small training vectors (40 specimen) one codebook is used for all signers.

As explained in section 2.4, there are 6 grid features per segment namely 1 value for pixel density, 1 value for predominant axial slant and 4 values for pixel distribution, and 96 (12×8) segments in

the grid forming 6 96-element vectors which are organised in 6 separate codebooks.

On the other hand, texture features (*see* section 2.5), are composed of 8 texture features (4 co-occurrence matrices $\times$ 2 elements) per segment, organised as a single feature vector having $96 \times 8 = 768$ elements which are coded in a single codebook.

As suggested by [10], it is desirable for each symbol or codeword to be represented in the training set by at least two to five times the number of vector components used in clustering. Empirical tests indicated that a fixed number of 50 codewords is suitable to cluster the grid and texture feature vectors.

Finally, the selected codewords are normalised within the range $[0, 1]$.

4. Training and Testing Protocol

Further to the three sets of features and the sample acquisition process which involved the signer to sign either freely on a blank sheet or within a provided frame (*see* section 2.1), permits the system to be trained in several ways.

Each training/testing strategy was validated by splitting the available data into two parts, one part for training and the other part for testing. Since the sample size for each author is relatively high (40 signature samples per author), this validation was considered to be sufficient, and further cross-validation was not performed.

The training included a combination of the two types of signature samples (framed and non-framed) and the three groups of signature features. The system is evaluated with the following three scenarios:

1) Training and testing only samples without a frame (TNTN)
2) Training and Testing samples both with frame and without frame (TATA)
3) Training signatures without a frame and testing all signatures (TNTA)

After several pilot studies, it was decided that a ratio of 5:3 will be used to train and test a signature model in the first two strategies mentioned above. For instance, in the first strategy, a signature model containing 25 non-framed signature samples, the system is trained with 16 random samples and is tested with the remaining 9 genuine samples together with all non-framed signature samples of the other authors (random forgeries). For the second strategy, a signature model composed of

25 non-framed samples and 15 framed samples, 16 random signatures are randomly selected from the non-framed signature samples while 9 samples are randomly selected from the 15 framed samples. For the third strategy the system was tested for robustness, where from a signature model of 40 samples (25 non-framed, 15 framed), the system is only trained with 15 non-framed signature samples and tested with the remaining 25 genuine samples amongst other author's signature samples (random forgeries).

The proposed architecture is evaluated in terms of False Acceptance Rate (FAR), False Rejection Rate (FRR) together with Total Error Rate (TER) and Mean Error Rate (MER) with the following features; global features only, grid features only, texture features only, global and grid features, global and texture features, grid and texture features, and finally global, grid and texture features.

5. Results

The following Receiver Operating Characteristic (ROC) curves compare the performance of the above features and their combination using the TATA scenario.

Clearly, the system performed best when it is trained with the combined global and grid features where the operating point is at (0.034, 0.033).
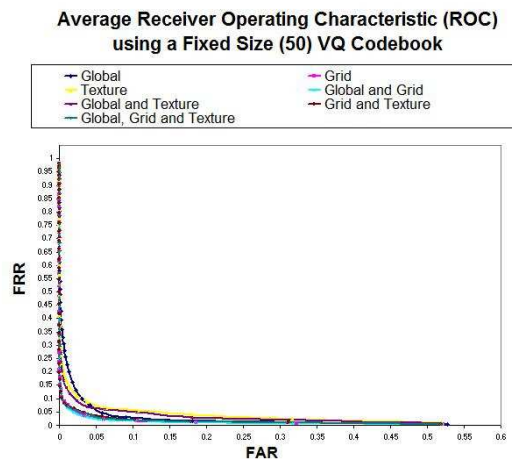

Figure 6. Average ROC

The worst performance is obtained with texture features alone with a TER of 11.83% with an FRR of 6.94%, an FAR of 4.89% and an MER of 5.915%.

When the system is tested for robustness, that is trained with non-framed signatures samples and

tested with both non-framed and framed samples, the best results are also obtained when the global and grid features were combined. In this case, the system achieves a TER of 6.31% with an FRR of 3.4%, an FAR of 2.91% and an MER of 3.155%. These results reflect the robustness of the system and may suggest that the provided frames may had affected the proportionality of the signatures.

In the above experiments the lowest FRR was achieved when the system is evaluated with global and grid features, in which the system was trained and tested with both framed and non-framed signatures. In this case, an FRR of 1.3% is achieved where the system rejects just 13 genuine signature samples out of 1003. On the other hand, the lowest FAR of 1.8% is achieved when the system is evaluated with all features; that is global, grid and texture features.

5.1. Comparison of Results

Since no international public signature database exists, different signature databases are used for different studies resulting in a difficult comparison of performance. Any comparison must be carried out with the mentioned restriction in mind.

The study reported in [3] used a combination of global, grid and texture features resulting in a TER of 12.81% with an FRR of 3% and an FAR of 9.81%. When adopting the same three groups of features, our system achieves better results where a TER of 4.79% is obtained with an FRR of 2.99% and an FAR of 1.8%.

Edson et al [11], achieved an MER of 2.135% when grid features, comprised of pixel density, pixel distribution and axial slant, were used in an HMM classifier. When adopting the same features in our study, the system achieves an MER of 2.295% which is only slightly worse to the results obtained by [11]. However, our proposed system achieves better results when the grid features are combined with global features where an MER of 2.07% is achieved.

6. Conclusion and Future Work

This study has shown that an RBFNN is a suitable architecture for OHSV and its preformance compares well to results reported in the literature. Best results were obtained when global and grid features are combined in a vector of 592 features.

Future work may include testing the system with simple and skilled foregeries as well as using

an adaptive technique to calculate the required number of codebooks for vector quantization. It would also be interesting to investigate the effect of feature vector dimension reduction techniques, such as principal component analysis.

References

[1] L. Likforman-Sulem, S. Garcia-Salicetti, J. Dittmann, J. Ortega-Garcia, N. Pavesic, G. Gluhchev, S. Ribaric, and B. Sankur, "Report on the hand and other modalities stated of the art," *Biometrics for Secure Authentication*, 2005.

[2] M. Gori and F. Scarselli, "Are multilayer perceptrons adequate for pattern recognition and verification?," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 11, pp. 1121–1132, 1998.

[3] H. Baltzakis and N. Papamarkos, "A new signature verification technique based on a two-stage neural network classifier," *Engineering Applications of Artificial Intelligence*, vol. 14, no. 1, pp. 95–103, 2001.

[4] C. Gonzalez and P. Wintz, *Digital Image Processing*, Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 2nd edition, 1987.

[5] R. W. Zhou, C. Quek, and G. S. Ng, "A novel single-pass thinning algorithm and an effective set of performance criteria," *Pattern Recognition*, vol. 16, no. 12, pp. 1267–1275, 1995.

[6] Y. Qi and B. R. Hunt, "Signature verification using global and grid features," *Pattern Recognition*, vol. 27, no. 12, pp. 1621–1629, 1994.

[7] E. J. R. Justino, F. Bortolozzi, and R. Sabourin, "The interpersonal and intrapersonal variability influences on off-line signature verification using hmm," in *Proceedings of the 15th Brazilian Symposium on Computer Graphics and Image Processing*, 2002, pp. 197–202.

[8] S. Haykin, *Neural Networks: A Comprehensive Foundation*, Prentice Hall PTR, Upper Saddle River, NJ, USA, 1994.

[9] E. J. R. Justino, A. El Yacoubi, and R. Sabourin, "An off-line signature verification system using hmm and graphometric features," in *DAS 2000, 4th IAPR International Workshop on Document Analysis Systems*, 2000, pp. 211–222.

[10] A.J. Elms, "The representation and recognition of text using hidden markov models," in *Ph.D.*, 1996.

[11] E. J. R. Justino, F. Bortolozzi, and R. Sabourin, "Off-line signature verification using hmm for random, simple and skilled forgeries," in *Proceedings of 6th International Conference On Document Analysis and Recognition*, 2001, pp. 1031–1034.