Off-line Signature Verification: A Circular Grid-Based Feature Extraction Approach

Marianela Parodi

<parodi@cifasis-conicet.gov.ar>

Juan Carlos Gómez

<jcgomez@fceia.unr.edu.ar>

Laboratory for System Dynamics and Signal Processing
FCEIA, Universidad Nacional de Rosario
CIFASIS, CONICET, Argentina
www.cifasis-conicet.gov.ar

Outline

Motivation for Off-line Signature Verification

- Today's society need for personal authentication has made automatic personal verification to be considered as a fundamental task in many daily applications.
- Signature verification is the most popular method of identity verification.
- Financial and administrative institutions recognize signatures as a legal means of verifying an individual's identity.
- No invasive methods of collecting the signature are needed.
- The use of signatures is familiar to people in their everyday's life.

Contributions

- A new feature extraction approach for off-line signature verification based on a circular grid is presented.
- Graphometric features used in the rectangular grid segmentation approach are adapted to this new grid geometry.
- A Support Vector Machine (SVM) based classifier scheme is used for classification tasks and a comparison between the rectangular and the circular grid approaches is performed.

XIII JCC

Circular Grid Feature Extraction Approach

A circular chart enclosing the signature is divided in N identical sectors, and graphometric features are computed for each sector. The circular grid is placed so that the center of the grid matches the center of mass of the binary image of the signature.

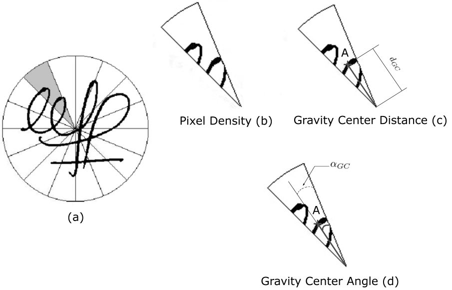


Fig. 1: Features extracted from segmented sectors with the circular grid approach: (a) Segmented sector being analyzed; (b) Pixel Density Distribution; (c) Gravity Center Distance; (d) Gravity Center Angle.

Circular Grid Feature Extraction Approach (cont.)

Some of the graphometric features used in rectangular grid segmentation are adapted to the new grid structure. Three static graphometric features are considered:

- **Pixel density distribution**

$$x_{PD_i} = \frac{\text{number of black pixels inside the sector}}{\text{total number of pixels inside the sector}} \quad i = 1, ..., N$$

- **Gravity center distance**

$$x_{DGC_i} = \frac{d_{GC_i}}{R} \quad i = 1, ..., N$$

- **Gravity center angle**

$$x_{AGC_i} = \frac{\alpha_{GC_i}}{\alpha_{max}}, \quad \text{being } \alpha_{max} = \frac{2\pi}{N} \qquad i = 1, ..., N$$

Circular Grid Feature Extraction Approach (cont.)

Finally, the feature vector x_{sign} is composed of the features calculated for each of the N angular sectors in which the signature image is divided, *i.e.*

$$x_{sign} = [x_{PD}^T, x_{DGC}^T, x_{AGC}^T]^T \quad ,$$

where

$$\begin{aligned} x_{PD} &= [x_{PD_1}, x_{PD_2}, \cdots, x_{PD_N}]^T, \\ x_{DGC} &= [x_{DGC_1}, x_{DGC_2}, \cdots, x_{DGC_N}]^T, \\ x_{AGC} &= [x_{AGC_1}, x_{AGC_2}, \cdots, x_{AGC_N}]^T. \end{aligned}$$

SVM-based Classifier

- SVM is a quite recent technique of statistical learning theory developed by Vapnik.
- In recent years, SVM-based classifiers have shown a promising performance in Automatic Signature Verification.

Separable Case

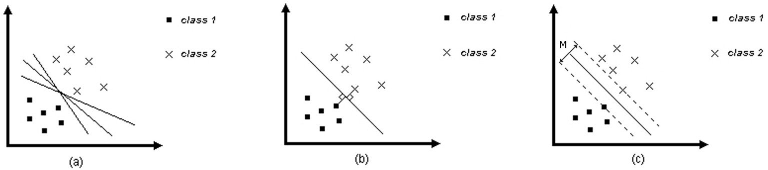


Fig. 2: Separable classification problem example: (a) Possible separating hyperplanes; (b) Selection of a unique hyperplane maximizing the distance between the nearest point of each class; (c) Optimal separating hyperplane that maximizes the margin.

SVM-based Classifier (cont.)

Consider the training set $\{x_k, y_k\}_{k=1}^{n}$, with input data $x_k \in \mathbb{R}^d$, output data $y_k \in \{-1, +1\}$ and suppose that all the training data satisfy the following constraints:

$$\begin{aligned} \omega^T x_k + b &\geq +1, \quad \text{for } y_k = +1 \\ \omega^T x_k + b &\leq -1, \quad \text{for } y_k = -1 \end{aligned}$$

Then the classifier takes the form

$$y_k[\omega^T x_k + b] - 1 \geq 0, \quad k = 1, ..., n.$$

where ω is normal to the hyperplane, $|b|/\|\omega\|_2$ is the perpendicular distance from the hyperplane to the origin and $\|\omega\|_2$ is the Euclidean norm of ω.

SVM-based Classifier (cont.)

The margin M, in this case, equals $2/\|\omega\|_2$ and the problem is solved by minimizing $\|\omega\|_2$ subject to the restrictions imposed by the data, *i.e.*, by solving the following optimization problem

$$\min_{\omega,b} \quad J_P(\omega) = \frac{1}{2}\omega^T\omega$$

$$\text{s.t.} \quad y_k[\omega^T x_k + b] \geq 1, \quad k = 1, ..., n.$$

SVM-based Classifier (cont.)

Non-Separable Case

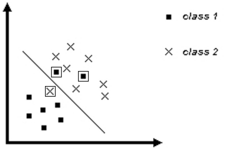


Fig. 3: Non-separable classification problem example.

In the non-separable case, one cannot avoid misclassifications. Then, slack variables have to be included in the formulation of the problem

$$y_k[\omega^T x_k + b] \geq 1 - \xi_k, \quad k = 1, ..., n.$$

SVM-based Classifier (cont.)

In this case, the optimization problem becomes

$$\min_{\omega,b,\xi} \quad J_P(\omega,\xi) = \tfrac{1}{2}\omega^T\omega + c\textstyle\sum_{k=1}^{n}\xi_k$$

$$\text{s.t.} \quad y_k[\omega^T x_k + b] \geq 1 - \xi_k, \qquad k = 1, \ldots, n$$

$$\xi_k \geq 0, \qquad k = 1, \ldots, n.$$

SVM-based Classifier (cont.)

Non-linear Case

The extension from the linear to the nonlinear case is straightforward. The linear separating hyperplane is calculated in a higher dimensional feature space where the input data lie after being mapped by a nonlinear mapping $\varphi(x)$. Then, the classifier in the case of nonlinear data is

$$y_k[\omega^T \varphi(x_k) + b] \geq 1 - \xi_k, \quad k = 1, ..., n.$$

No explicit construction of the nonlinear mapping $\varphi(x)$ is needed, by applying the so-called kernel trick. That is, by defining a Kernel as $K(x_k, x_\ell) = \varphi(x_k)^T \varphi(x_\ell)$ for $k, \ell = 1, ..., n$. The SVM solution can be found by solving the following optimization problem

$$\begin{array}{lll} \min\limits_{\omega,b,\xi} & J_P(\omega, \xi) = \frac{1}{2}\omega^T\omega + c\sum_{k=1}^{n} \xi_k & \\ \text{s.t.} & y_k[\omega^T \varphi(x_k) + b] \geq 1 - \xi_k, & k = 1, ..., n \\ & \xi_k \geq 0, & k = 1, ..., n. \end{array}$$

SVM-based Classifier (cont.)

The SVM classifier takes the following form

$$y(x) = sign[\sum_{k=1}^{n} \alpha_k y_k K(x, x_k) + b].$$

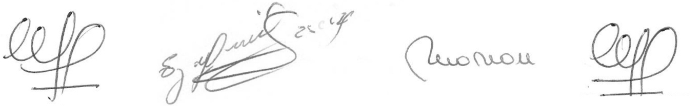
Different Kernels have been used in the literature to solve pattern recognition problems. Linear, Polynomial and Radial Basis Functions (RBF) Kernels are among the most popular in the bibliography

$$\begin{aligned} K_{linear}(x_k, x_\ell) &= x_k^T x_\ell, \\ K_{polynomial}(x_k, x_\ell) &= (1 + x_k^T x_\ell)^d, \\ K_{RBF}(x_k, x_\ell) &= exp(- \parallel x_k - x_\ell \parallel_2^2 / \sigma^2). \end{aligned}$$

Experiments and Results

The aim of a signature verification system is to accurately distinguish between two categories of signatures, namely, genuine and forged signatures.

- **Types of errors:**
  - False Rejection Rate (FRR)
  - False Aceptance Rate (FAR)
- **Types of forgeries:**
  - random forgery
  - simple forgery
  - skilled forgery



(a) (b) (c) (d)

Fig. 4: An original signature instance and its different types of forgeries. (a) Original signature; (b) Random forgery; (c) Simple forgery; (d) Skilled forgery.

Experiments and Results (cont.)

The **Signature Database**[1] used includes **160 writers** with:

- 24 genuine signatures per writer
- 30 forged signatures per writer (simple and skilled forgeries)

An **SVM** model was trained for each writer
Training set:

- genuine samples: 14 samples per writer
- false samples (random forgeries): 5x159=795 samples per writer

Neither simple nor skilled forgeries were include in the training subset of false samples.

[1] Vargas, F., Ferrer, M.A., Travieso, C.M., Alonso, J.B.: Off-line Handwritten Signature GPDS-960 Corpus. In: IAPR 9th International Conference on Document Analysis and Recognition, Curitiba, Brazil, 764–768 (2007)

Experiments and Results (cont.)

Experiments with the circular grid approach were carried out with different number of divisions of the grid and different types of kernels:

- Different number of grid divisions $N = 8$, $N = 16$, $N = 32$, $N = 64$ and $N = 128$.
- Different types of kernels: linear, polynomial and RBF.

Experiments with the same number of divisions and the same types of kernels were carried out with the rectangular grid approach in order to compare both feature extraction techniques.

In order to obtain reliable results, Monte Carlo techniques were used. The experiments were carried out randomly resampling the dataset into training and testing sets for each one of the writers. The resampling process has been repeated 200 times.

Experiments and Results (cont.)

FRR

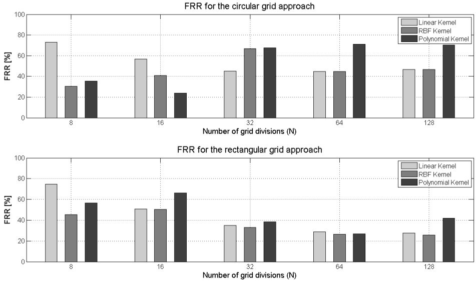


Fig. 5: FRR for different number of divisions and kernels, for the circular (top) and the rectangular (bottom) grid approaches.

FAR for simple and skilled forgeries

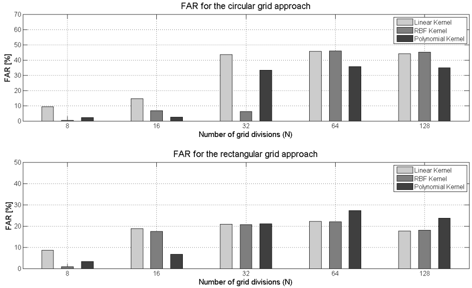


Fig. 6: FAR (simple and skilled forgeries) for different number of divisions and kernels, for the circular (top) and the rectangular (bottom) grid approaches.

FAR for random forgeries

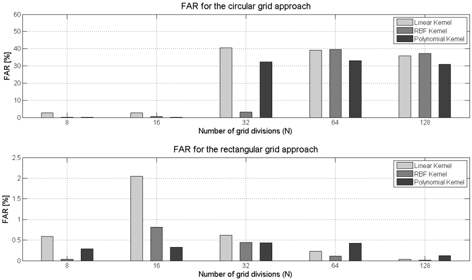


Fig. 7: FAR (random forgeries) for different number of divisions and kernels, for the circular (top) and the rectangular (bottom) grid approaches.

Experiments and Results (cont.)

Comparison between the best results reached with the circular and rectangular grid approaches:

	Circular Grid N=16, Poly kernel (degree 3) Feature Vect. Dim.=**48**	Rectangular Grid N=128, RBF kernel ($\sigma^2 = 100$) Feature Vect. Dim.=384
FRR	**23.9147%**	25.7678%
FAR (simple & skilled forgeries)	**2.4314%**	17.9528%
FAR (random forgeries)	0.0867%	**0.0321%**
EER	**8.8109%**	14.5842%

Table 1: Best results for circular and rectangular grid approaches.

Experiments and Results (cont.)

The statistical significance of the obtained results can be inferred from the *box plots* of the FRR, FAR for simple and skilled forgeries and FAR for random forgeries for each of the 160 writers in the database.

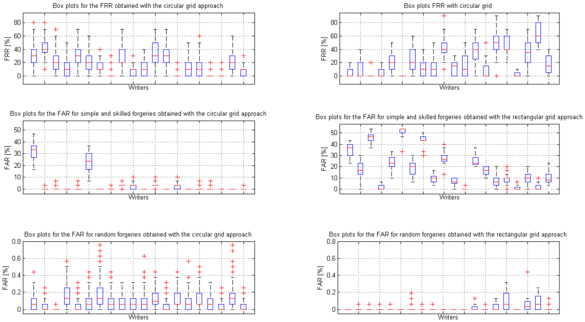


Fig. 8: Box plots for 20 writers in the database. Left column: circular grid. Right column: rectangular grid. Top: FRR. Middle: FAR for simple and skilled forgeries. Bottom: FAR for random forgeries.

Experiments and Results (cont.)

Comparison between the results obtained with the proposed approach and other approaches proposed in the literature.

	FRR	FAR (simple and skilled forgeries)	FAR (random forgeries)	EER
Proposed approach	23.9147%	2.4314%	0.0867%	8.8109%
Ferrer et. al. [4]	14.1%	12.6%		13.35%
Vargas et. al. [13]	10.01%	14.66%		12.33%

Table 2: Results obtained with the proposed approach and other approaches proposed in the literature.

Conclusions

- A new feature extraction approach based on a circular grid has been presented for off-line signature verification.
- A comparison between the circular and the rectangular grid based feature extraction approaches has been performed over a SVM-based classification scheme.
- The classification results, quantified by the FRR and the FAR for simple and skilled, and random forgeries, using the proposed features have shown improvements with respect to the ones based on features extracted from rectangular grids.
- The low FAR obtained indicates an improvement in the capability of the system to highlight the interpersonal variability.

Future Work

- In order to reduce the FRR, that is to increment the verification process capability to absorb the intrapersonal variability, possible strategies are:
  - Introduce new graphometric features (specially dynamic ones).
  - Modify the ratio between genuine and false samples used for training. It is likely that reducing the number of random forgeries used to train the false class, will result in an improvement in the FRR value.
- Increase the database in order to have more data available to perform the statistic tests.

References

[1] Impedovo, D., Pirlo, G.: Automatic Signature Verification: The State of the Art. IEEE Transactions on Systems, Man and Cybernetics-part C: Applications and Reviews. vol. 38., no. 5. 609–635 (2008)
[2] Justino, E., El Yacoubi, A., Bortolozzi, F., Sabourin, R.: An Off-Line Signature Verification System Using HMM and Graphometric Features. In: 4_{th} IAPR International Workshop on Document Analysis Systems, Rio de Janeiro, Brazil, 211–222 (2000)
[3] Justino, Edson J. R., Bortolozzi, F., Sabourin, R.: Off-line Signature Verification Using HMM for Random, Simple and Skilled Forgeries. In: Internat. Conf. on Document Analysis and Recognition. vol. 1. Seattle, USA, 105–110 (2001)
[4] Ferrer, M.A., Alonso, J.B., Travieso, C.M.: Offline geometric parameters for automatic signature verification using fixed-point arithmetic. IEEE Transactions on Pattern Analysis and Machine Intelligence. vol. 27. 993–997 (2005)
[5] Justino, Edson J.R., Bortolozzi, F., Sabourin, R.: A Comparison of SVM and HMM Classifiers in the Offline Signature Verification. Pattern Recognition Letters 26, 1377–1385 (2005)
[6] Özgündüz, E., Sentürk T., Karsligil, M. E.: Off-line Signature Verification and Recognition by Support Vector Machine. In: European Signal Processing Conference. Turkey (2005)
[7] Oliveira L. S., Justino, E., Freitas, C., and Sabourin, R.: The Graphology Applied to Signature Verification. In: 12th Conference of the International Graphonomics Society, 286–290 (2005)
[8] Santos, C., Justino, E., Bortolozzi, F., Sabourin, R.: An Off-Line Signature Verification Method based on the Questioned Document Expert's Approach and a Neural Network Classifier. In: Proceedings of the 9th Int'l Workshop on Frontiers in Handwriting Recognition (IWFHR-9) (2004)
[9] Vapnik, V.: The Nature of Statistical Learning Theory. Springer-Verlag, NY (1995)
[10] Vapnik, V.: Statistical Learning Theory. Wiley, NY (1998)
[11] Vargas, F., Ferrer, M.A., Travieso, C.M., Alonso, J.B.: Off-line Handwritten Signature GPDS-960 Corpus. In: IAPR 9th International Conference on Document Analysis and Recognition, Curitiba, Brazil, 764–768 (2007)
[12] Canu, S., Grandvalet, Y., Guigue, V., Rakotomamonjy, A.: SVM and Kernel Methods Matlab Toolbox. Perception Systèmes et Information, INSA de Rouen, Rouen, France (2005)

References (cont.)

[13] Vargas, F., Ferrer, M.A., Travieso, C.M., Alonso, J.B.: Off-line Signature Verification based on high pressure polar distribution. In: 11th International Conference on Frontiers in Handwriting Recognition (ICFHR08), Montreal, Canada, (2008)

Thanks a lot !!!

Off-line Signature Verification: A Circular Grid-Based Feature Extraction Approach

Marianela Parodi

<parodi@cifasis-conicet.gov.ar>

Juan Carlos Gómez

<jcgomez@fceia.unr.edu.ar>

Laboratory for System Dynamics and Signal Processing
FCEIA, Universidad Nacional de Rosario
CIFASIS, CONICET, Argentina
www.cifasis-conicet.gov.ar

CIFASIS
CONICET
UNR | UPCAM