

UNIVERSIDAD NACIONAL DE ROSARIO
FACULTAD DE CIENCIAS EXACTAS, INGENIERÍA Y AGRIMENSURA
LABORATORIO DE SISTEMAS DINÁMICOS Y PROCESAMIENTO DE LA INFORMACIÓN
CIFASIS, CONICET



Automatic Handwritten Signature Verification

Which features should be looked at?

Marianela Parodi

Supervisor: Prof. Juan Carlos Gómez
Evaluation Committee: Prof. Leandro Di Persia
Prof. Pablo Granitto
Prof. Angelo Marcelli
Prof. Giuseppe Pirlo

*A thesis submitted to Facultad de Ciencias Exactas,
Ingeniería y Agrimensura in partial fulfilment of the
requirements for the degree of*

Doctor en Ingeniería

ROSARIO, ARGENTINA

March 2015

©2014 by Marianela Parodi
Laboratorio de Sistemas Dinámicos y Procesamiento de la Información
Facultad de Ciencias Exactas, Ingeniería y Agrimensura
Universidad Nacional de Rosario
CIFASIS, CONICET
Rosario, Argentina

I hereby certify that the work embodied in this thesis is the result of original research and has not been submitted for a higher degree to any other University or Institution.

Marianela Parodi

*"A signature always reveals a man's character...
and sometimes, even his name".*

Even Esar

Acknowledgements

First of all, I would like to specially thank my supervisor Dr. Juan Carlos Gómez for his complete support during my research. This Thesis would not have been possible without his (infinite) patient and guidance. I would also like to thank him for giving me this great opportunity of joining his team and work together.

I would like to thank my lab colleagues for their kind support and friendship. It is always nice to know that you are around!

Finally, I would like to thank Ezequiel for being always by my side, my mom and dad for believing in me, and my sister and brother for being part of my life.

Abstract

The increasing need for personal authentication in many daily applications has made biometrics a fundamental research area. In particular, handwritten signatures have long been considered one of the most valuable biometric traits. Signatures are the most popular method for identity verification all over the world, and people are familiar with the use of signatures for identity verification purposes in their everyday life. In fact, signatures are widely used in several daily transactions, being recognized as a legal means of verifying an individual's identity by financial and administrative institutions. In addition, signature verification has the advantage of being a non-invasive biometric technique.

Two categories of signature verification systems can be distinguished taking into account the acquisition device, namely, offline systems, where only the static image of the signature is available, and online systems, where dynamic information acquired during the signing process, such as pen coordinates, pen pressure and pen inclination angles, is available. In this Thesis both, the offline and the online modalities, are addressed.

For the offline signature verification case, a new feature extraction approach based on a circular grid scheme is proposed. Graphometric features are adapted to be extracted resorting to this new grid geometry. In addition, the property of rotation invariance of the Fast Fourier Transform (FFT) is used in order to achieve robustness against rotation of the signatures, which is an important issue for offline signature verification.

For the case of online signature verification, contributions towards feature selection and extraction are done. In addition, the designed systems are tested on a challenging recently available public database containing signatures from different cultural origins, namely Western (Dutch) and Chinese signatures.

Two different approximations based on orthogonal series expansions of the time functions associated to the signing process are proposed for online signature feature extraction. One of them is based on Legendre polynomials, and the other one is based on wavelet decomposition. The coefficients in these orthogonal series expansions of the time functions are used as features to model the signatures. In addition, an in depth analysis of different combinations of several time functions is carried out in order to provide some insight on their actual discriminative power. Moreover, a novel consistency factor is proposed in order to quantify this

discriminative power.

On a subsequent step towards improving the performance of the online signature verification systems, a pre-classification stage based on global features is incorporated to the system. The idea is to quickly recognize and discard gross forgeries based on the pre-classification approach. It is expected that incorporating the pre-classification step would speed up and simplify the verification process.

To bridge the gap between Forensic Handwriting Expert (FHE) and Pattern Recognition (PR) communities, is currently one of the most important challenges in the field. In an attempt to make some contribution to this issue, the discriminative power of a set of features which seems to be relevant to signature analysis by FHEs is analysed and particularly compared to the discriminative power of automatically selected feature sets. This analysis is intended to help FHEs to further understand the signatures and the writer behaviour. Even the feasibility of developing a system which could complement the FHEs work is evaluated. For this purpose, a fusion between automatic selected features and FHE based features is proposed. To conclude this analysis, the feasibility of using only FHE based features for automatic online signature verification is evaluated.

Finally, an attempt to give an answer to the question of whether is it possible to combine all the proposed online features in order to achieve better verification results, is provided.

The obtained experimental results are technically sound, since the verification performances are comparable and, in some cases, even better than the ones in the state-of-the-art. In addition, interesting observations can be made on the basis of the analysis and discussions carried out which can make contributions towards some of the most important actual challenges in the field.

Contents

| | |
|----------------------------------------------------------------------|------------|
| Acknowledgements | vi |
| Abstract | vii |
| 1 Introduction | 1 |
| 1.1 Why Automatic Signature Verification? | 1 |
| 1.2 Automatic Signature Verification Systems | 3 |
| 1.3 Contributions | 4 |
| 1.4 Publications | 6 |
| 1.5 Thesis Organization | 8 |
| 2 State of the Art | 9 |
| 2.1 A Little bit of History | 9 |
| 2.2 The Current Challenges | 14 |
| 3 Offline Signature Verification: Searching for Robustness | 19 |
| 3.1 Introduction | 19 |
| 3.2 Feature Extraction | 21 |
| 3.2.1 Pre-processing | 21 |
| 3.2.2 Circular Grid Based Feature Extraction Approach | 21 |
| 3.3 Signature Database | 22 |
| 3.4 SVM-based Classifier Applied to Signature Verification | 23 |
| 3.5 Evaluation Protocol | 24 |
| 3.6 Results and Discussion | 25 |
| 3.7 Searching for Robustness against Signature Rotation | 28 |
| 3.7.1 Feature Extraction | 28 |
| 3.7.2 Evaluation protocol | 30 |
| 3.7.3 Results and Discussion | 31 |
| 3.8 Some Concluding Remarks | 33 |
| 3.9 Some Ideas for Further Work | 34 |
| 4 Online Signature Verification: Features | 35 |
| 4.1 Introduction | 35 |
| 4.2 Feature Extraction | 36 |
| 4.2.1 Time Functions and Preprocessing | 37 |

| | | |
|----------|----------------------------------------------------------------------------|-----------|
| 4.2.1.1 | Basic functions | 37 |
| 4.2.1.2 | Normalisation | 37 |
| 4.2.1.3 | Extended functions | 39 |
| 4.2.2 | Orthogonal Time Function Representation | 40 |
| 4.2.2.1 | Time Function Approximation via Legendre Polynomials | 40 |
| 4.2.2.2 | Time function approximation via wavelets | 42 |
| 4.3 | Signature Database | 44 |
| 4.4 | Feature Combination Experiments | 45 |
| 4.4.1 | Combination Experiments Using a Basic Features Set | 46 |
| 4.4.1.1 | Evaluation Protocol | 46 |
| 4.4.1.2 | Results and Discussion | 49 |
| 4.4.2 | Combination Experiments Using a More Complex Feature Set | 52 |
| 4.4.2.1 | Evaluation Protocol | 53 |
| 4.4.2.2 | Results and Discussion | 54 |
| 4.5 | Consistency Measure | 57 |
| 4.5.1 | Definition | 58 |
| 4.5.2 | Evaluation Protocol | 59 |
| 4.5.3 | Results and Discussion | 60 |
| 4.6 | Some Concluding Remarks | 66 |
| 4.7 | Some Ideas for Further Work | 68 |
| 5 | Online Signature Verification: Exploiting characteristics | 71 |
| 5.1 | Introduction | 71 |
| 5.2 | Pre-classification | 72 |
| 5.3 | Feature Extraction | 74 |
| 5.3.1 | Global based features | 74 |
| 5.3.2 | Time function based features | 75 |
| 5.4 | Evaluation protocol | 75 |
| 5.5 | Results and Discussion | 76 |
| 5.5.1 | Univariate case | 76 |
| 5.5.2 | Multivariate case | 78 |
| 5.6 | Some Concluding Remarks | 80 |
| 5.7 | Some Ideas for Future Work | 81 |
| 6 | Bridging the gap between FHE and PR | 83 |
| 6.1 | Introduction | 83 |
| 6.2 | FHE Time Function Based Features | 85 |
| 6.3 | FHE vs. Automatically Selected Features | 86 |
| 6.4 | On Combining FHE and Automatically Selected Features | 88 |
| 6.5 | Automatic signature verification based on FHE Features | 93 |
| 6.5.1 | FHE Based Global Features | 93 |

| | | |
|----------|------------------------------------------------------------|------------|
| 6.5.2 | Global and Time function FHE based features fusion | 94 |
| 6.5.3 | Evaluation Protocol | 95 |
| 6.5.4 | Results and Discussion | 95 |
| 6.6 | Some Concluding Remarks | 97 |
| 6.7 | Some Ideas for Future Work | 98 |
| 7 | When the whole is greater than the sum of the parts | 99 |
| 7.1 | Introduction | 99 |
| 7.2 | Feature Combination Approaches | 100 |
| 7.2.1 | Decision Level Fusion | 100 |
| 7.2.2 | Pre-classification | 101 |
| 7.3 | Evaluation protocol | 101 |
| 7.3.1 | Decision level fusion | 101 |
| 7.3.2 | Pre-classification | 102 |
| 7.4 | Results and Discussion | 102 |
| 7.5 | Some Concluding Remarks | 105 |
| 7.6 | Some Ideas for Future Work | 105 |
| 8 | Conclusions and Future Work | 107 |
| A | Notation | 111 |
| B | Datasets | 113 |
| B.1 | GPDS300Signature CORPUS Database | 113 |
| B.2 | SigComp2011 Database | 113 |
| C | Support Vector Machine | 115 |
| D | Random Forest | 119 |
| D.1 | Decision trees | 119 |
| D.2 | Ensemble methods | 119 |
| D.3 | Random Forest | 120 |
| | Bibliography | 123 |

List of Figures

| | | |
|------|--------------------------------------------------------------------------------------------------------------------------------------------|----|
| 1.1 | Automatic signature verification scheme | 3 |
| 1.2 | Offline and online versions of a handwritten signature | 4 |
| 2.1 | Genuine and forged signature samples | 12 |
| 3.1 | Circular grid based feature extraction | 22 |
| 3.2 | Typical genuine and forged offline signatures | 23 |
| 3.3 | FRR for circular and rectangular grid schemes | 26 |
| 3.4 | FAR (skilled forgeries) for circular and rectangular grid schemes | 27 |
| 3.5 | FAR (fictitious signatures) for circular and rectangular grid schemes | 27 |
| 3.6 | Rotation invariant vs. non rotation invariant performance | 31 |
| 3.7 | FRR and FAR for the rotation invariant feature extraction approach | 32 |
| 3.8 | Wrongly classified signatures | 33 |
| 4.1 | Legendre polynomials based time function approximation | 42 |
| 4.2 | Filter bank scheme for wavelet decomposition | 44 |
| 4.3 | Wavelet decomposition based time function approximation | 45 |
| 4.4 | Offline and online versions of a typical genuine and forged Dutch signature. | 46 |
| 4.5 | Offline and online versions of a typical genuine and forged signatures. | 47 |
| 4.6 | Verification errors of different feature combinations for the Dutch dataset | 49 |
| 4.7 | Verification errors of different feature combinations for the Chinese dataset | 50 |
| 4.8 | Intraclass and interclass distances distribution | 59 |
| 4.9 | Consistency factor for the Dutch dataset | 61 |
| 4.10 | Consistency factor for the Chinese dataset | 62 |
| 4.11 | Spearman's correlation coefficient between the consistency factors for the Dutch dataset | 64 |
| 4.12 | Spearman's correlation coefficient between the consistency factors for the Chinese dataset | 65 |
| 4.13 | Spearman's correlation coefficient: consistency (genuine signatures) vs. consistency (skilled forgeries) for the Dutch dataset | 66 |
| 4.14 | Spearman's correlation coefficient: consistency (genuine signatures) vs. consistency (skilled forgeries) for the Chinese dataset | 67 |

| | | |
|-----|---------------------------------------------------------------------------------------------------------------------|-----|
| 5.1 | Threshold computation from the global based feature distribution for genuine and forged signatures | 73 |
| 5.2 | Threshold computation from the global based feature vector distribution for genuine and forged signatures | 74 |
| 5.3 | Distribution of two dimensional global based feature vectors | 79 |
| 5.4 | Percentage of signatures pre-classified as forgeries | 81 |
| 6.1 | Wrongly classified signatures with FHEs based features | 88 |
| 6.2 | Fusion information schemes | 89 |
| 6.3 | Optimisation of the weighting parameter for the Dutch and Chinese datasets | 91 |
| 6.4 | Beta optimisation for the Dutch and Chinese datasets | 96 |
| C.1 | Separable classification problem example | 115 |
| C.2 | Non-separable classification problem example | 116 |

List of Tables

| | | |
|------|-------------------------------------------------------------------------------------------------------------------------------------|-----|
| 3.1 | Best results for circular and rectangular grid approaches | 28 |
| 3.2 | Verification performance comparison with state-of-the-art approaches | 33 |
| 4.1 | Best FIT using Legendre polynomials | 42 |
| 4.2 | Best FIT using wavelet decomposition | 44 |
| 4.3 | List of feature combinations | 48 |
| 4.4 | Best verification results using Support Vector Machines and Random Forests classifiers for the Dutch and Chinese datasets | 51 |
| 4.5 | Best verification results for the Dutch and Chinese datasets | 52 |
| 4.6 | Verification results for the Dutch Dataset | 55 |
| 4.7 | Verification results for the Chinese Dataset | 55 |
| 4.8 | Optimal feature sets for the Dutch and Chinese datasets | 56 |
| 4.9 | Verification results: Summary for the Dutch dataset | 68 |
| 4.10 | Verification results: Summary for the Chinese dataset | 69 |
| 5.1 | Univariate pre-classification: Verification results for the Dutch and Chinese datasets | 77 |
| 5.2 | Multivariate pre-classification: Verification results for the Dutch and Chinese datasets | 78 |
| 6.1 | FHEs based features vs. Automatically selected features | 86 |
| 6.2 | Feature and decision level fusion between FHEs and a reduced set of automatically selected features | 90 |
| 6.3 | Decision level fusion between FHEs and automatically selected features | 93 |
| 6.4 | Verification results using only FHE based features for the Dutch and Chinese datasets | 96 |
| 7.1 | Verification results for the Dutch and Chinese datasets | 103 |
| B.1 | Online Dutch and Chinese datasets | 114 |

Introduction

"Definitely, further research is necessary to fully investigate and interpret the potential of handwritten signatures, which remain very distinct signs, unequivocally demonstrating the inspiration and complexity of human beings." (D. Impedovo and G. Pirlo. Automatic signature verification: The state of the art. IEEE Trans. on Systems, Man, and Cybernetics - Part C: Application and Reviews, 38(5):609635, 2008)

1.1 Why Automatic Signature Verification?

In recent years, the society's need for personal authentication has increased significantly, making biometrics a fundamental task in many daily applications. Biometrics refer to the recognition of individuals based on quantifiable data related to their characteristics and traits. Biometric systems can either verify or identify a person. In the verification case, the system has to determine if an individual's claimed identity is genuine or not, while in the identification case, the system has to establish the individual's identity. Two different types of biometric traits can be defined, namely, physiological and behavioural. The former considers biological characteristics of a person, such as face, hand geometry, fingerprint, retina and iris, while the latter is related to psychological aspects of a person, like voice, handwriting and handwritten signatures. Despite the variety of biometric features available, it is widely accepted that the usefulness of a biometric trait is strongly dependent on the specific application. Then, the choice of the biometric trait has to be made based not only on technical aspects but also on social and cultural ones. In [JBP99], it is stated that, as a general rule, a biometric characteristic should be robust, *i.e.*, should not present substantial changes over time; distinctive, *i.e.*, should show a great variation over different subjects; available, *i.e.*, people should have this characteristic; accessible, *i.e.*, easy to collect; and acceptable, *i.e.*, people should agree for this characteristic to be taken from them.

Handwritten signatures have long been considered one of the most valuable biometric traits, being in the present the most popular method for identity verification all over the world. This popularity is probably due to the fact that people are familiar with their use for identity verification purposes in their everyday life, but also for being recognized as a legal means of verifying an individual's identity by financial and administrative institutions. In addition, signature verification has the important advantage of being a non-invasive biometric technique. The popularity and advantages of using handwritten signatures as a biometric trait, have made signature verification to be long considered an important task in the field of biometrics [PL89], [LP94], [PS00], [IP08] and [IPP12]. Further, in recent years, handwritten signatures have been considered with a renewed interest due to the widespread use of electronic pen-input devices, such as digitiser tablets and personal digital assistants (PDAs) in many daily applications [IPP12].

Two categories of signature verification systems can be distinguished taking into account the acquisition procedure, namely, offline and online systems. For offline systems, the acquisition takes place once the signing process has finished, thus only a static image of the signature is available. An updated survey of this type of systems can be found in [PBP11a]. For online systems, the acquisition is performed during the signing process, making dynamic information acquired during this process available. In this case, the signature is parameterised by several discrete time functions such as x and y pen coordinates, pen pressure and, when available, pen azimuth and altitude inclination angles. See [ZWW11] for a survey of online systems. Offline systems have the advantage of being low cost, since no specific equipment is required to collect the signatures, making this acquisition process easier than the one needed in the online case. On the other hand, since dynamic information is more difficult to forge, it is believed that offline systems are less accurate than online ones.

The need to verify an identity occurs in many different situations such as personnel identification, contracts, bank checks, invoice authentication, and any kind of fraud detection. In general, the current application and the available data determine the use of the offline or the online approach to perform the signature verification. In particular, as already mentioned, the interest in the online approach has increased in recent years together with the development of electronic input devices such as digitiser tablets, PDAs, digital pens and ink pens, among others. Nevertheless, there are certain applications that demand the use of the offline approach like, for example, forensic signature comparisons, since Forensic Handwriting Experts (FHEs) only have the offline data available in their daily casework. The verification of signatures in bank checks has long been considered a fundamental application in the field of offline signature verification. Bank check images are very complex since they usually have a colourful background with several logos and preprinted guidelines. The extraction of signatures from the checks and their posterior process is a difficult task and the development of signature

verification systems with the accuracy required for banks and other financial and administrative institutions is an still open challenge [DIPS97], [YY97], [DNP98]. On the other hand, the ability to capture the signature and have it immediately available in a digital form for verification also opens up a range of new application areas. Basically, any system that uses a password or PIN, such as file and device access, secure physical entry systems and internet based transactions, can instead use an online signature for access, having the advantages of being more difficult to steal or guess than a password and also easier to remember for the user.

As suggested above, automatic handwritten signature verification has plenty of important applications in people's everyday life. This has made many researchers in the field of biometrics to give automatic signature verification special attention. As already mentioned, several works can be found in the literature, being [IP08] and [IPP12] among the most recent and complete surveys, summarizing the contributions in the field. Nevertheless, there are still several open issues to be addressed. This Thesis tries to give some insight into automatic handwritten signature verification and attempts to address some of its actual challenges.

1.2 Automatic Signature Verification Systems

The aim of an automatic handwritten signature verification system is to decide if the claimed identity for the input signature is genuine or not. Figure 1.1 schematically depicts an automatic handwritten signature verification system.

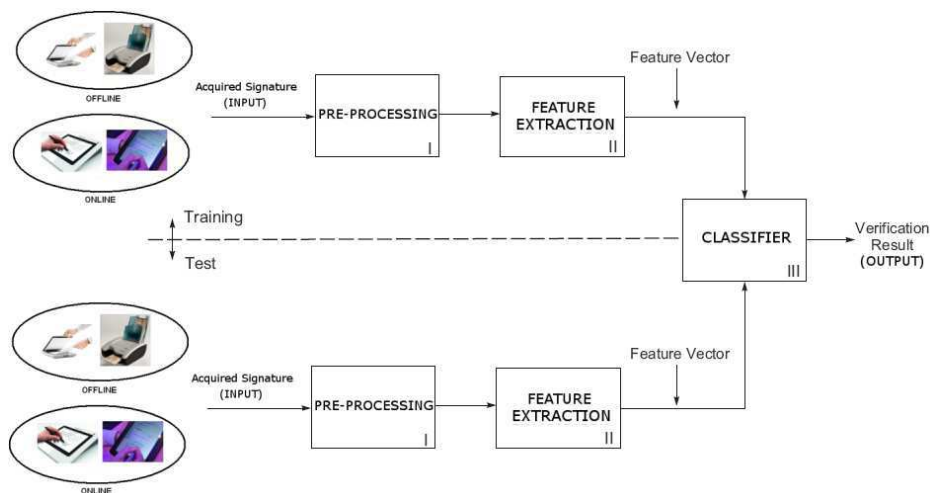


Figure 1.1: Automatic signature verification scheme.

The acquisition devices can be either scanners, for the case of offline systems, or digitiser tablets, for the case of online systems. Then, the measured data

(which is the output of the acquisition device and the input of the verification system) will be either a signature image or a set of discrete time functions such as pen coordinates x and y , pen pressure, and pen inclination angles, respectively. Figure 1.2 (left) shows an example of an offline signature, while Fig. 1.2 (right) shows an example of an online version of the same signature reconstructed from the x and y pen coordinates.

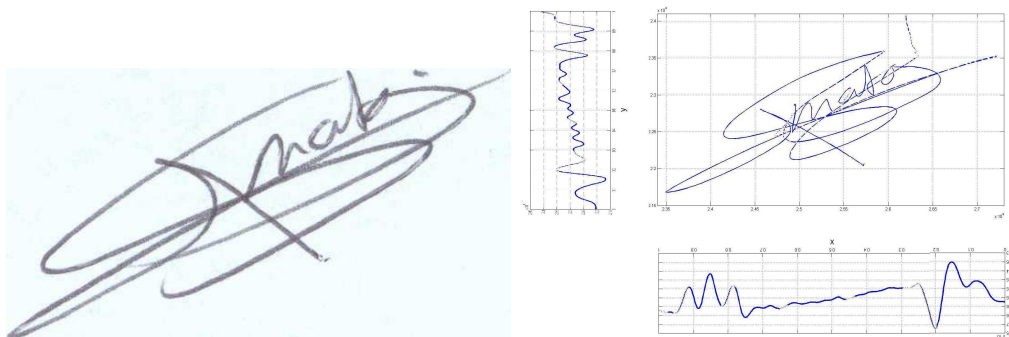


Figure 1.2: Offline (left) and online (right) versions of a handwritten signature.

After the acquisition of the input signature, the first step in the verification process consists in the pre-processing of the measured data (Block I) in order to enhance it. This stage is generally based on standard signal processing techniques. The next step in the verification process is the feature extraction (Block II), where several characteristics are extracted from the pre-processed data in order to construct a feature vector to represent the input signature. The more discriminative the selected characteristics are, the better the verification results will be. Feature extraction is considered to be the most important task of the whole process, and for this reason, special attention has been given to this task in the last years [IP08], [IPP12]. Although a large number of features and extraction techniques have already been proposed in the literature, it is a fact that the success of a set of features depends on several factors like the application, the cultural/social origin of the signatures, and the amount and quality of the available data, among others. Then, the feature extraction is still an open and interesting challenge in the field. Finally, the classification step (Block III) is also an important step in the verification process. Here, a classifier is employed to decide if the input signature corresponds to the claimed identity or not.

1.3 Contributions

As shown in Section 1.2, handwritten signature verification systems are composed of several phases. Researchers in the field have long discussed about the influence

of each one of them in the performance of the whole system. Despite the fact that it would not be sensible to give any definite conclusion, most researchers agree that the feature extraction phase determines the success or fail of the verification process. In line with this idea, most of the contributions of this Thesis are focused in the feature extraction stage of the signature verification system.

As already mentioned in Section 1.1, two categories of signature verification systems can be distinguished taking into account the acquisition device, namely, offline systems, where only the static image of the signature is available, and online systems, where dynamic information acquired during the signing process, such as x and y pen coordinates, pen pressure and pen inclination angles, is available. The interest in the online approach for signature verification has increased in recent years because of the widespread use of electronic pen-input devices, such as digitiser tablets and PDAs. In addition, it would be reasonable to expect that the incorporation of dynamic information acquired during the signing process would make signatures more difficult to forge and, in this way, the online verification systems more reliable than the offline ones. Nevertheless, there are certain applications that demand the use of the offline approach. For instance, FHEs often only have the offline data available in their daily casework. To perform a forensic signature comparison, it is then necessary to work with offline data, while online data can be used to perform biometric person verification/identification. In this Thesis both, the offline and the online modalities, are addressed. Different feature extraction approaches are proposed for each one of them. It is important to note, however, that most of the contributions are focused on the online approach since it is believed that, in a near future, this will be the option of choice in automatic signature verification systems.

Contributions towards offline signature verification are presented in [PG10a], [PG10b] and [PGB11]. A new feature extraction approach is proposed based on a circular grid scheme and the computation of graphometric features. In addition, the rotation invariance property is incorporated to this new feature extraction technique by using the Fast Fourier Transform (FFT).

Contributions towards feature extraction and selection for online signature verification are presented in [PGL12], [PGLA13a] and [PG14b]. New techniques for feature extraction based on orthogonal polynomials series are proposed in [PGL12] and [PGLA13a]. In particular, an approximation based on Legendre polynomials is presented in [PGL12], while an approximation based on wavelet decomposition is introduced in [PGLA13a]. In [PG14b] and [PGLA13a], an in depth analysis of different feature combinations is carried out in order to give some insight on their actual discriminative power. Finally, a consistency measure is proposed in [PG14b] to quantify the discriminative power of the features.

The idea of incorporating a pre-classification stage based on global features to an online signature verification system for the purposes of improving its performance, is introduced in [PG13]. It would be reasonable to expect that, in addition

of improving the system performance, the proposed approach would also have the advantage of simplifying and speeding up the verification process. In [PG14a], the univariate pre-classification proposed in [PG13] is extended to a multivariate case.

To bridge the gap between PR and FHE communities is a fundamental task. Contributions towards this direction are presented in [PGLA13b], [PGAL14] and [PGA14]. In [PGLA13b], the feasibility of developing a system which could complement the FHEs work is evaluated. For this purpose, the discriminative power of some features which seems to be relevant to signature analysis by FHEs is studied. In [PGAL14], a fusion between an automatic selected feature set and a FHE based feature set is proposed in order to improve the discriminative power of the system. Finally, the feasibility of using only FHE based features for automatic online signature verification is evaluated in [PGA14].

In [PGLA13a], [PG13] and [PGAL14], different features are proposed for online signature verification. The question arises whether it is possible to combine all of them in order to achieve better verification results. An attempt to give some answers to this question is also done in this Thesis.

1.4 Publications

The research results obtained within the framework of the present Thesis have been partially published in the following articles:

Book Chapters

[PG14a] Parodi, M. and Gómez, J.C., Using Global Features for Pre-classification in Online Signature Verification Systems, Chapter 4 in *Advances in Digital Handwritten Signature Processing - A Human Artefact for e-Society*, G. Pirlo, D. Impe-dovo and M. Fairhurst (Editors), pp. 39-52, World Scientific Publishing Co. and Imperial College Press, UK.

Journal Papers

[PGAL14] Parodi, M., Gómez, J.C., Alewijnse L. and Liwicki, M., Online Signature Verification: Automatic Feature Selection vs. FHE's Choice. *Journal on Forensic Document Examination*, Vol. 24 (2014).

[PG14b] Parodi, M. and Gómez, J.C., Legendre polynomials based feature extraction for online signature verification. Consistency analysis of feature combinations, *Pattern Recognition*, Elsevier, Vol. 47, No. 1 (2014), pp.128-140.

[PGLA13a] Parodi, M., Gómez, J.C., Liwicki, M., and Alewijnse L., Orthogonal Function Representation for Online Signature Verification. Which Features Should be Looked at?, *IET Biometrics*, Vol. 4, No. 2 (2013), pp.137-150.

[PG13] Parodi, M. and Gómez, J.C., Online Signature Verification: Improving performance through pre-classification based on global features. In A. Petrosino,

L. Maddalena, P. Pala (Eds.): ICIAP 2013 Workshops, LNCS 8158, pp. 69-76, Springer, Berlin, 2013.

[PG12] Parodi, M. and Gómez, J.C., Online Signature Verification Based on Legendre Series Representation. Consistency Analysis of Different Feature Combinations. In L. Alvarez et al. (Eds.): Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications, Lecture Notes in Computer Science, CIARP 2012, LNCS 7441, pp. 715-723, Springer, Berlin, 2012.

Refereed Conference Papers

[PGA14] Parodi, M., Gómez, J.C., and Alewijnse L., Automatic Online Signature Verification based only on FHE Features: an Oxymoron? In *Proceedings of the 14th International Conference in Frontiers of Handwriting Recognition (ICFHR 2014)*, Crete Island, Greece, pp. 73-78, September 1-4, 2014.

[TPG14] Terissi, L., Parodi, M. and Gómez, J.C., Lip reading using wavelet-based features and Random Forests classification. In *Proceedings of the 22nd International Conference on Pattern Recognition (ICPR 2014)*, Stockholm, Sweden, pp. 791-796, August 24-28, 2014.

[TSGP14] Terissi, L. Sad, G., Gómez, J.C. and Parodi, M., Noisy Speech Recognition based on Combined Audio-Visual Classifiers. In *Proceedings of the 3rd International Workshop on Multimodal pattern recognition of social signals in human computer interaction (MPRSS 2014)*, Stockholm, Sweden.

[PG13b] Parodi, M. and Gómez, J.C., Online Signature Verification: Improving performance through pre-classification based on global features. In *Proceedings of the Workshop on Emerging Aspects on Handwritten Signature Processing (EAHSP 2013)*, within the International Conference on Image Analysis and Processing (ICIAP), Naples, Italy, pp. 69-76, September 9-13, 2013.

[PGLA13b] Parodi, M., Gómez, J.C., Liwicki, M., and Alewijnse L., Wavelet Based Representation for Online Signature Verification. Automatic Feature Selection vs. FHE's Choice. In *Proceedings of the 16th International Graphonomics Society Conference (IGS2013)*, Nara, Japan, pp. 1-6, June 10-13, 2013.

[PGL12] Parodi, M., Gómez, J.C., and Liwicki, M., Online Signature Verification Based on Legendre Series Representation. Robustness Assessment of Different Feature Combinations. In *Proceedings of the 13th International Conference in Frontiers of Handwriting Recognition (ICFHR 2012)*, Bari, Italy, pp. 377-382, September 18-20, 2012.

N.B.: This paper was the recipient of the **GOOGLE Best Student Paper Award**.

[PG12b] Parodi, M. and Gómez, J.C., Online Signature Verification Based on Legendre Series Representation. Consistency Analysis of Different Feature Combinations. In *Proceedings of 17th Iberoamerican Conference on Pattern Recognition (CIARP 2012)*, Buenos Aires, Argentina, pp. 715-723, September 03-06, 2012.

[PGB11] Parodi, M., Gómez, J.C. and Belaïd, A., A Circular Grid-Based Rotation Invariant Feature Extraction Approach for Off-line Signature Verification. In *Proceedings of the 11th International Conference on Document Analysis and Recognition (ICDAR 2011)*, Beijing, China, pp. 1289-1293, September 18-21, 2011.

[PG10b] Parodi, M. and Gómez, J.C., Off-line Signature Verification: A Circular Grid-Based Feature Extraction Approach. In *Proceedings of the XXXVI Conferencia Latinoamericana de Informática (CLEI 2010)*, Asunción, Paraguay, October 18-22, 2010.

[PG10a] Parodi, M. and Gómez, J.C., A Circular Grid-Based Feature Extraction Approach for Off-line Signatura Verification. In *Proceedings of the XI Argentine Symposium on Artificial Intelligence (ASAI2010)*, 39th JAIIO, CABA, Argentina, pp. 163-174, August 30-31, 2010.

1.5 Thesis Organization

The Thesis is organized as follows. The automatic handwritten signature verification state-of-the-art is described in Chapter 2, where the principal challenges of the area are discussed. Chapter 3 is focused on the offline approach and the contributions made in this field are presented. The feature extraction and selection is a fundamental step in every verification system. This is discussed in Chapter 4 for online systems, and the contributions made in this area are presented. In Chapter 5, some ideas for exploiting the intrinsic characteristics of the selected features are discussed. Chapter 6 is devoted to one of the most important and actual challenges in the field: "Bridging the gap between FHE and Pattern Recognition (PR) communities". Chapter 7 shows how to put all parts together in order to improve the verification system. Finally, in Chapter 8 the conclusions are given and the future directions are discussed.

State of the Art

2.1 A Little bit of History

Automatic handwritten signature verification has been considered an important area in the field of biometrics as well as in the field of handwriting recognition since the 1970s. The first works concerning automatic signature verification were published in the mid 1970s [Ste75], [NR77], [HL77]. Studies on the offline [NR77] and online [HL77] modalities were presented at that time. These approaches were mostly focused on the description of the features and the implementation of some simple classifiers like the ones based on distances or k-Nearest Neighbours (kNNs). In the 1980s, the idea that the pixel's gray level is closely connected with the pen pressure was introduced in [AYF86]. In [AYF86], the main emphasis was given on high-pressure regions which resulted to be the dark pixels in the signature image corresponding to high-pressure points on the writing trace. This work was one of the firsts attempts to extract pseudo-dynamic information from a static image and gave offline signature verification new directions. Several works followed this pioneer one and made contributions in the same direction, such as [SP86], where image gradient analysis was proposed to pre-process the signatures, and [SD92], where the directional Probability Density Function (PDF) was proposed to perform the feature extraction.

As the computational power increased, more challenging problems could be faced. In the 1990s, larger databases could be compiled, larger amounts of data could be handled and more complex classifiers like Artificial Neural Networks (ANN) and Hidden Markov Models (HMM) could be used. All these works, together with other important contributions until the 2000s, are highlighted in [PL89], [LP94] and [PS00].

More recently, the constantly increasing demand of security systems in many daily applications and the development of new signature acquisition technologies, have given new perspectives to the field. Works such as [IP08] and [IPP12], summarize the most relevant contributions in the last years. In 2004, the First International Signature Verification Competition (SVC2004) was held within the

First International Conference on Biometric Authentication (ICBA 2004). This competition was the starting point of several competitions that began to be held within the most important conferences in the area: the International Conference on Document Analysis and Recognition (ICDAR) and the International Conference on Frontiers in Handwriting Recognition (ICFHR). The celebration of these competitions is of vital importance for the field since not only gives the researchers the opportunity to compare their systems' performances with other ones, but also to test their algorithms with benchmark databases that are then made available to the whole community. In addition, the databases presented during the competitions are usually quite challenging. For instance, in ICDAR 2009 [BvdHCKL09] and ICDAR 2011 [LMd⁺11] a Chinese database was made available, while in ICDAR 2013 [LMd⁺11] a Japanese database was introduced. Several tutorials and workshops have also been carried out in the last years as satellites of the mentioned and other important main conferences. These periodic meetings allows the community to discuss the new directions in the field, giving researchers the opportunity to exchange ideas and discuss their current work in a relaxed context.

Automatic handwritten signature verification systems are designed in order to discriminate genuine signatures from forged ones. Handwritten signatures are behavioural biometric traits, so the aspect of a signature and the way in which the writer signs are likely to depend on several factors, such as, for example, the signer's age, her home country, her physical, psychological and emotional state, and the conditions in which the signature process takes place. The ability of the system to reject the forged signatures and accept the genuine ones will then be subject to several different aspects, and the way in which these different factors are addressed along the verification process will determine its success or failure. The separability of the two classes of signatures depends on their inter and intra personal variability. Unfortunately, two instances of the same signature may differ significantly from each other, making the verification problem a challenging task. In addition, different type of signatures are employed to test the verification system. Researchers have long discussed about the categorization of these different types of signatures. In particular, there has long been a mismatch between the terminology used by researchers in the PR and in the FHE communities, leading to numerous misunderstandings and making a terminology unification strictly necessary. Recently, due to several efforts made in this direction, both communities have agreed in the use of a common terminology [ML12]. The most usual types of signatures are defined as follows, according to the adopted new terminology:

- Genuine signatures:
 - Regular: The specimen writer signs in the way she usually does.
 - Unnatural situation: The specimen writer signs under unusual condi-

tions that can alter the way in which she signs. These conditions can be intrinsic like diseases, or drugs, or extrinsic like writing on different surfaces.

- Forgeries: A non-specimen writer tries to imitate the genuine signature from a known sample of it. There are different kinds of forgeries, any of them can be traced or freehanded. The most common types of forgeries are:
 - Simple: A non-practised simulation of the genuine signature.
 - Skilled: The forger is allowed to practise the known sample. If possible, the practised time has to be specified.
- Disguise: The specimen writer tries to make the signature to look like a forgery.
- Fictitious: The non-specimen writer does not have access to a sample of the genuine signature she is trying to imitate, then she makes up a signature. This type of signatures were formerly called random signatures.

It is important to note that it is not always possible to test a signature verification system against all the different kinds of attacks that can appear in practical situations (skilled forgeries, disguised signatures, fictitious signatures, etc.). Despite the fact that many databases have been made publicly available in the last years, the lack of benchmark databases continues being a key problem in the area. In addition, the available databases do not usually contain all the types of signatures. In particular, disguised signatures have first appeared in a publicly available database in 2010 [LvdHFM10], and only in their offline version. Disguised signatures together with skilled forgeries are one of the most frequent attacks in forensics daily casework. As an attempt to develop useful tools for FHEs, PR researchers are trying to work, when possible, not only on skilled forgeries but also on this type of signatures. Of course, the first step to take in this direction should be the compilation of databases containing disguised signatures for both, offline and online cases. Figure 2.1 shows an instance of a genuine signature (a) and the most common types of attacks in the PR literature, namely, fictitious signatures (b and c) and a skilled forgery (d).

As already introduced in Section 1.2, the automatic signature verification process consists of three main steps: pre-processing (Block I in Fig. 1.1), feature extraction (Block II in Fig. 1.1) and classification (Block III in Fig. 1.1). In the last decades of development, several techniques have been proposed concerning each one of these different steps involved in the verification process.

The pre-processing has two main goals, namely, enhancement of the input data and segmentation. The enhancement of the input data, is generally based on standard signal processing techniques. Of course, the applied techniques will depend on the nature of the data, *i.e.*, offline or online signatures. For the offline

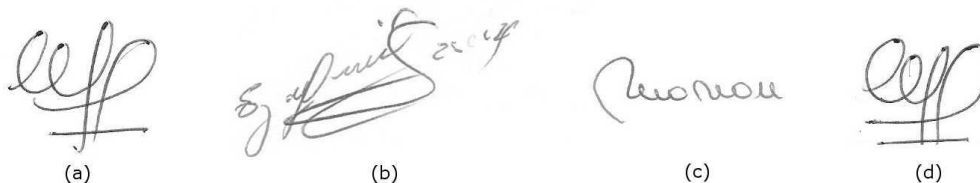


Figure 2.1: A genuine signature instance and its different types of forgeries. (a) Genuine signature; (b and c) Fictitious signatures; (d) Skilled forgery.

case, signature extraction based, in general, on global thresholding [Ots94], noise removal based on typical image filters, such as, median or morphological filters, signature size normalisation, binarization and thinning [GW08], [GWE09] are typically used. For the online case, where less pre-processing is usually needed, filtering, noise reduction and smoothing are used. These pre-processing tasks are usually based on mathematical transforms like Fast Fourier Transform (FFT) [KTN96] or Gaussian filters [JGC02]. In addition, signature normalisation regarding position, size, orientation and time duration is also important [IGHG04].

The segmentation is used to divide the signature in basic units that are representative of some aspect of the signature. The segmentation technique will also depend on the acquisition method. Structural descriptors based on the identification of connected components [CDIP93], tree structure [AYF90], statistic of directional data [SP86], among others, are used for the offline case. For the online case, pen up / pen down movements [IP92], velocity signal analysis [KHH03] and perceptually relevant points [BP93] are some of the used techniques. Also Dynamic Time Warping (DTW) [YO77] has been widely used since it allows to segment two or more signatures into the same number of corresponding elements [MC96], [WRV04], [MMLL07], [CS07].

As already said, feature extraction is probably, and according to many researchers in the area, the most important step in the whole verification process. The goal of the feature extraction step is to find a suitable description for the signature. The idea is to extract some characteristics from the pre-processed data in order to construct a feature vector which can represent the input signature and that allows to discriminate it from other signatures. The more discriminative the selected characteristics are, the better the verification results will be. Due to the importance of selecting powerful features, researchers in the field have analysed lots of different characteristics. For the offline case, geometric based parameters, such as area, height, width and height-to-width ratio of the signatures are typically computed [AYF90], [SPB94]. Other geometric descriptors like the ones based on countour [NP94], orientation [SPB94], direction [SD92], [DSG94], [ZNA05] slant [AYF90] and curvature [JGC02], [KSX04], are also widely used.

Grid based features, that are local features extracted from cells in which the signatures are divided, are also among the most popular in the literature [SP86], [JS01], [JYBS00], [SJBS04], [JBS05], [OJFS05] [BZZE13], and have been used since [NR77]. In recent years, descriptors like SIFT (Scale-Invariant Feature Transform) [SS10], [Sha14], and SURF (Speeded Up Robust Features) [GS13] [GMH12] have begun to be used. For the online case, local as well as global features computed from the dynamic signals acquired during to the signing process are widely used [RKD05], [GFRO05], [LG05]. Local features like pen position, velocity, acceleration, pressure and inclination have long been used and their consistency has long been studied [MM07], [HGD09]. Global features like total signature time duration, average velocity, number of pen lifts, among others, are widely used [LBA96], [KRD05]. Statistical models of the signatures have also been long and widely used [LP94]. In addition, coefficients obtained from mathematical transforms like FFT [YK09] and Wavelet decomposition [CDW⁺12] are also used to represent the signatures.

Finally, the classification stage is also a crucial step in the verification process. Several classification techniques have been used for signature verification purposes, being ANN [BC97], [HY97], [Lee95] and HMM [YWP95], [GD98], [JYBS00], [JS01] among the most popular ones. In addition, classifiers based on Support Vector Machines (SVM), first introduced in [Vap95], have been widely used [FAT05], [FGSD02], [JBS05], [PFV11]. In recent years, researchers have also shown an increasing interest in the use of classifiers like AdaBoost [HMM05], [HC13] and Random Forests (RF) [ASVE13]. To combine classifiers has also gained popularity in recent years and several works has been presented in this direction [IPB12], [Alh12].

There exist several ways in which the performance of a signature verification system can be evaluated. Since the signature verification problem is a binary problem, typical errors used for binary classifiers have historically been used to evaluate the performance of signature verification systems. The False Rejection Rate (FRR) and the False Acceptance Rate (FAR) have been the most widely used. The FRR concerns to the rejection of genuine signatures, while the FAR concerns to the acceptance of forged signatures. Additionally, it can be defined the Equal Error Rate (EER), which is the system error rate when the FRR equals the FAR, and it is usually considered as a measure of the overall error of the verification system. Based on the FRR and the FAR, the Receiver Operating Characteristic (ROC) curves can be constructed. These curves indicate the capacity of the system to discriminate between genuine and false signatures. They show the true positives (TPR: True Positive Rate) vs. false positives (FPR: False Positive Rate), as a threshold is modified. The TPR and the FPR are related with the FRR and the FAR through $TPR = 1 - FRR$ and $FPR = FAR$. The Detection Error Tradeoff (DET) curves are also widely used. Moreover, they are often preferred over the ROC curves since they are more linear than ROC curves

and use most of the image area to highlight the differences of importance in the critical operating region.

In the last years, the importance of the computation of the log-likelihood ratios [MDK⁺97] to evaluate the performance of signature verification systems has been pointed out, specially, since the competition held within ICDAR 2011 [LMd⁺11] where the use of these measurements to evaluate the performance of a signature verification system was introduced. In particular, the cost of the log-likelihood ratio \hat{C}_{ur} and its minimal value \hat{C}_{ur}^{min} are usually computed. Having available the log-likelihood ratio information makes FHEs give an opinion on the strength of the evidence [GFRO05] although they are not in the position to make a leap of faith and judge about guilt or no guilt.

2.2 The Current Challenges

Despite the fact of being an active field of research in the last decades, there are still lots of open challenges in the area of automatic handwritten signature verification. In addition, in recent years due to the increasing needs of security in many daily applications, new challenges have emerged.

The analysis of the actual challenges of the field is a fundamental step in any research in order to be aware of the field needs. There exist interesting open challenges regarding several aspects of the signature verification problem. In [IPP12] an updated survey on new advances and open challenges in the field can be found.

The development of the digitising devices, ranging from the traditional table-based tablets to the recent handy digitiser tablets, PDAs, and input devices for mobile computing, makes device interoperability, that is, the capability of verification systems to adapt to data obtained from different devices a new problem that needs to be addressed [AFFAOG05], [MDFKG14], [MDFOG14], [KFGMD13]. This constitutes one of the most important issues regarding the acquisition of the signatures.

Regarding the pre-processing of the signals (images in the offline case and time functions in the online case), most of the applied techniques have been in use during the last decades and they are quite reliable. Nevertheless, it is worth to mention that one of the goals of the pre-processing is to guarantee the robustness of the system against translation, scaling and rotation of the signatures. The translation and scale invariance property of the system is usually easy to achieve. On the other hand, the robustness against rotation is not always easy to be achieved, and despite the fact that there are some works in the literature addressing this problem [CNPM04a], [WFTZ07a] it remains being an open field of research.

Many researchers agree that the feature selection and extraction is the most important part of the whole verification system. There are several decisions to

make in this stage, and the success of the system depends on them. Which features are the best ones to perform the verification is the principal question. It is not enough to have plenty of features to describe a signature, what is more important is to interpret those features and to evaluate the discriminative power they have. Nowadays, there exist powerful tools that can process large amounts of data, then classifying signatures described by lots of features is not a difficult task. Nevertheless, the tendency is not to use as many features as it can be processed but to use the minimum amount of features capable of accurately representing the signatures. Several studies have been carried out concerning which are the best features to model the signatures.

For the offline case, global and local features have long been studied and compared [PLD14], [MLD11]. Global features are less sensitive to noise and signature variations [HYW04], so they are a good choice when dealing with genuine signatures but they are not when dealing with skilled forgeries. That is why they are often used in combination with local features, which are believed to be more robust at the cost of a more complex and time consuming extraction technique [MLD11].

For the online case, the discussion is focused on which are the best time functions to represent the signatures. During SVC2004, the results using only pen coordinates outperformed those adding pen pressure and inclination angles [YCX⁺04]. Since then, several works have been presented concerning the best set of features to model the signatures. In [KY05b], the authors state that using only pen coordinates leads to better results than incorporating the pen pressure. In [HGD09], it is concluded that pen pressure is the most unreliable feature, pen inclination angles are too unstable, and pen coordinates are the most robust time functions in the presence of a long term time variability between training and testing data acquisition sessions. On the other hand, some works showed improvements when combining pen coordinates with pen pressure and inclination angles [MM07]. The conflicting results observed in the literature make the discussion about which features should be used still open.

The classification stage involves many critical aspects that range from techniques for signature matching and the design of the evaluation protocols to the availability of public databases and the strategy used for their compilation. Regarding the classification, HMM [YWP95], [GD98], [JYBS00], [JS01] and ANN [BC97], [HY97], [Lee95] have long been used, as already mentioned in Section 2.1. Recently, SVM have also been used by researchers in the field [FAT05], [FGSD02], [JBS05], [PFV11]. Even more recently, ensemble classifiers like Boosting, Bagging and RF are used for classification [HMM05], [HC13], [ASVE13]. To combine multiple classifiers has also been proposed in recent works [IPB12], [Alh12], constituting an interesting alternative for the classification step. Fortunately, there exist several classification approaches in the PR literature, the challenge for developers is to adapt them to their particular verification systems.

Regarding the evaluation protocol, PR researchers have long agree in computing the EER from ROC curves constructed based on the FAR and FRR. Recently, together with the growing efforts of the PR and FHEs communities to bridge the gap between them, the analysis of the log likelihood ratios [Bd06] to evaluate the verification performance has been proposed. Using these measurements to evaluate the performance of a signature verification system was proposed in AFHA 2011 Workshop¹, where the importance of computing the likelihood ratios was highlighted since they allow FHEs to give an opinion on the strength of the evidence [GFRO05], although they would not be in the position to make a leap of faith and judge about guilt or no guilt.

Benchmark databases are of crucial importance since they give researchers a reliable and standard framework where they can test their systems. The lack of such databases, as already mentioned in Section 2.1, is still an important problem in the field. Although not enough, several databases are currently available to the research community. The ones presented in [SJBS04] and [YCX⁺04] are among the most popular ones for offline and online verification, respectively. Unfortunately, these databases already have 10 years now and are hardly used since it is usually preferred to work with more recent databases. In [KY09], a publicly available online database is introduced and, even more interesting, benchmark evaluation protocols are proposed to evaluate the verification systems over the presented database. Several databases have been presented during signature verification competitions, as it is the case of the one in [YCX⁺04]. During the competitions held within the last three editions of ICDAR (in 2009, 2011 and 2013), databases containing offline as well as online Dutch, Chinese and Japanese signatures have been presented, respectively. During the competitions held within ICFHR 2010 [LvdHFM10] and 2012 [LMA⁺12] offline databases containing disguised signatures have been presented. Databases containing synthetic signatures have also been presented in last years [JPFO12], [JFOP12]. Even more recently, a new database containing signatures captured using a Samsung Galaxy Note device has been made publicly available [MM14].

Regarding the availability of data, the lack of sufficient reference data to characterize a given signature class, as is generally the case of many practical applications, is a critical point. To address this problem, it is necessary to evaluate different types of signatures, their complexity [AFFFOG07], [GPFMD11], and stability [PIP⁺13], [IPP12], [PFM13]. Different approaches have been proposed:

- Estimate the statistical significance of small-size training sets.
- Generation of additional training samples from the existing ones [JPFO12], [JFOP12].

¹First International Workshop on Automated Forensic Handwriting Analysis (AFHA 2011) held within ICDAR 2011, Beijing, China

- Selection of the optimal subset of reference signatures, among the specimens available. The stability analysis is a technique that can be used to perform this selection [PFM13].

An important factor that deserves more investigation, is the influence of the cultural origin of the signatures in the performance of the verification systems. Despite the fact that researchers are currently making valuable contributions in this area, there are still not many works in the literature that consider non-Western signatures such as Chinese, Japanese, Arabic, etc.. In [PBP11b], an updated survey of non-English and non-Latin signature verification systems can be found. Non-Western signatures do have different shapes and the writing style is different to the Western one. For instance, the Chinese handwriting style consists of one or more multi-trace ideograms, and Chinese characters usually convey their meaning through pictorial resemblance to a physical object. Among the literature of non-Western signature verification, more attention has been given to Chinese signatures than to Japanese, Arabic, Persian or Indian ones. Offline [LWWZ05], [JC09], and online [PFV11], [ZTLL03] verification systems have been presented in the literature for Chinese signature verification. Furthermore, as already mentioned in Section 2.1, during the Signature Verification Competition for Online and Offline Skilled Forgeries (SigComp2011) held within ICDAR 2011 [LMd⁺11], a new publicly available database containing Chinese signatures was introduced, encouraging researches to work on this type of data. On the other hand, Japanese and Arabic signatures, among others, have not been investigated so much. Japanese signatures consist of component characters spaced from each other. There is not much work done on this type of data [YY98], [Ued03], and most of these works are focused in offline data. Fortunately, in [MLA⁺13b] a new publicly available database containing Japanese signatures has been presented. Arabic script has the particularity of being written from right to left with most of the letters within a word directly connected to the adjacent letters. Although a lot of research has been carried out on Arabic handwriting recognition, not much work has been carried out on Arabic signature verification. In [IG00], an offline verification system for Arabic signatures is presented. Researchers in the field believe that any verification system should be able to deal with different writing styles, in order to be widely accepted. As pointed out in [PBP11b], there are still a lot of research to do in this area.

In this Thesis, some of the open challenges in the field of automatic handwritten signature verification are addressed and some contributions are made towards them.

Offline Signature Verification: Searching for Robustness

In this Chapter, a new rotation invariant feature extraction approach based on the computation of graphometric features resorting to a circular grid scheme and the use of the FFT properties for offline signature verification is introduced.

3.1 Introduction

The choice of a suitable set of features to represent the signatures is a fundamental step in a signature verification process. Several methods have been proposed in the offline signature verification literature to perform the selection and the extraction of features from the signature image. Generally, two categories of features can be distinguished, namely, global and local features. Global features refer to features that are representative of the whole signature image, while local features are those extracted from particular parts of the signature image. Grid segmentation schemes have been widely used in the literature to compute local features. In addition, particular attention has been devoted to the extraction of graphometric features, that are features inspired by the ones used in graphology, resorting to these grid schemes, as it is demonstrated in [OJFS05], [JYBS00], [JS01], [JBS05] and [SJBS04]. In [OJFS05], a wide set of graphometric features is derived from the ones used in graphology. Graphometric features can be divided into static and pseudo-dynamic features. Static features are related basically to the occupation of the graphical space, while the pseudo-dynamic features are directly related to the strokes of the signature. In [JYBS00] and [JS01], three features are extracted from each cell of the rectangular grid. Two static features, namely, the pixel density (number of pixels inside the cell) and the pixel distribution computed as in [SPB94], and the axial slant (predominant slant inside the cell) as a pseudo-dynamic feature. In [JBS05], the pixel density and the gravity centre (gravity centre distance in each cell) are computed as static features, while the

stroke curvature (curvature angle of the bigger stroke inside de cell) and the slant are computed as pseudo-dynamic features. In [SJBS04], four static features (calibre, proportionality, white spaces and base behaviour) and three pseudo-dynamic features (apparent pressure, curvature and progression) are used.

Although rectangular grid schemes have been widely used in the literature, it would be reasonable to expect that rectangular grids would not be always suitable for feature computation. In particular, it is usually the case that rectangular grid schemes lead to empty sectors. The ideal gridding technique would be to compute a bounding ellipsoid of the signature and to divide it into sectors. In this case, empty sectors would be avoided, but no regular sectors would result, which would probably be difficult to deal with. A bounding circular grid, instead, would allow the division in regular sectors, while avoiding the presence of empty sectors. In addition, it would be reasonable to expect that a circular grid scheme would be more adaptable to a rotation invariant feature extraction approach than the traditional rectangular grid schemes. Robustness against rotation of the signatures has long been one of the major challenges when dealing with off-line signature verification systems. In order to achieve invariance against rotation, it is a usual practice in the literature to perform a skew correction during the pre-processing phase [KJZ13], [DG14]. Nevertheless, the real challenge is to develop a feature extraction that can guarantee robustness against rotation. This problem has not been extensively dealt with in the literature.

In this Chapter, the use of a circular grid scheme is then proposed and graphometric features are adapted to be computed resorting to the this new grid geometry. The benefits of using this grid scheme over using the traditional rectangular ones are studied by comparing their corresponding verification results. Finally, once the advantages of the use of the circular grids have been analysed, they are taken into account together with the properties of the FFT, in order to develop a rotation invariant feature extraction technique.

The Chapter is organized as follows. Section 3.2 introduces the circular grid based feature extraction approach. In Section 3.3, the offline database used for the experiments is described. Section 3.4 is focused on the use of SVM based classifiers for signature verification. The evaluation protocol is presented in Section 3.5, and the results of this first part of the Chapter are presented in Section 3.6. Finally, Section 3.7 presents the modifications to the feature extraction proposed in Section 3.2 in order to guarantee robustness against rotation of the signatures. Some concluding remarks are given in Section 3.8, while some future directions are discussed in Section 3.9.

3.2 Feature Extraction

3.2.1 Pre-processing

The signatures in the database are acquired using a standard scanner and they are stored as 8-bit gray scale images. The first step in the pre-processing is to binarise the image. In addition, since dimensions of signatures belonging to different writers, or even the same writer, may differ, a width normalisation of the signature image is performed. This normalisation maintains the original height-to-width ratio of the signature image.

To perform the proposed feature extraction, it is necessary to compute the centre of mass of the signature in order to place the centre of the grid, and the principal axis of the signature in order to construct the circular chart to enclose the signature. These geometric values are also computed in this pre-processing step.

3.2.2 Circular Grid Based 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 centre of the grid matches the geometric centre of the binary image of the signature as shown in Fig. 3.1a. In this way, the probability of having empty grid sectors is reduced and also the invariance against translation of the signature is guaranteed. The invariance against scaling of the signature, is achieved by the width normalisation mentioned above which keeps the original height-to-width ratio of the signature image.

Three static graphometric features are considered, namely, pixel density distribution x_{PD} , gravity centre distance x_{DGC} and gravity centre angle x_{AGC} , defined as follows:

$$x_{PD_i} = \frac{\text{Num. of black pixels inside the sector}}{\text{total Num. of pixels inside the sector}}, \quad (3.1)$$

$$x_{DGC_i} = \frac{d_{GC_i}}{R}, \quad (3.2)$$

$$x_{AGC_i} = \frac{\alpha_{GC_i}}{\alpha_{max}}, \quad \text{being } \alpha_{max} = \frac{2\pi}{N}, \quad (3.3)$$

with $i = 0, \dots, N-1$, respectively. Here, d_{GC} is the distance between the gravity centre (point A in Fig. 3.1(b)-(c)) and the centre of the grid, R is the radius of the grid calculated as the major distance between extreme points of the signature, α_{GC} is the angle of the gravity centre and α_{max} is the angle of each sector as depicted in Fig. 3.1(c). Note that due to the particular choice of the grid centre and the grid radius, the features (3.1), (3.2) and (3.3) are translation and scaling invariant.

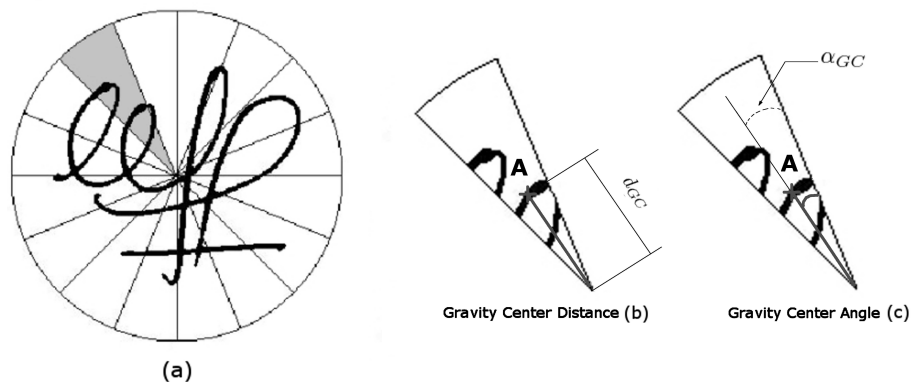


Figure 3.1: Features extracted from segmented sectors with the circular grid approach: (a) Segmented sector being analysed; (b) Gravity Centre Distance; (c) Gravity Centre Angle.

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.*

$$x_{sign} = [x_{PD}^T, x_{DGC}^T, x_{AGC}^T]^T, \quad (3.4)$$

where

$$x_{PD} = [x_{PD_1}, x_{PD_2}, \dots, x_{PD_N}]^T, \quad (3.5)$$

$$x_{DGC} = [x_{DGC_1}, x_{DGC_2}, \dots, x_{DGC_N}]^T, \quad (3.6)$$

$$x_{AGC} = [x_{AGC_1}, x_{AGC_2}, \dots, x_{AGC_N}]^T. \quad (3.7)$$

3.3 Signature Database

The GPDS300Signature CORPUS [VFTA07], which is a public and freely available (for research purposes) database, is used to perform the verification experiments. A description of this database can be found in Appendix B. There are 160 writers enrolled in the database. For each writer, there are 24 genuine signatures and 30 forged signatures. That is, a total of $160 \times 24 = 3840$ genuine and $160 \times 30 = 4800$ forged signatures. According to the authors in [VFTA07], forgeries in the database are simple and skilled ones. As already introduced in Section 2.1, the first ones are non-practised imitations of the genuine signature model, while the latter are practised imitations of the genuine signature model, where the forgers are allowed to practise the imitation for as long as they deem it necessary. Figure 3.2 shows a sample of a typical genuine (left) and forged (right)

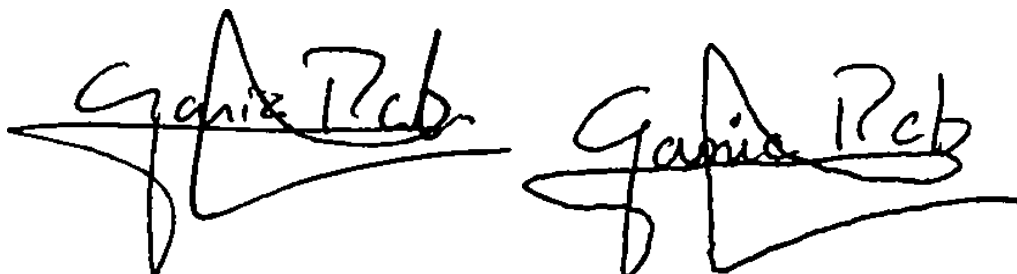


Figure 3.2: Typical genuine (left) and forged (right) signatures in the database.

signature in the database. Finally, in addition to test the proposed verification system against simple and skilled forgeries, fictitious signatures, which are signatures that the forger simply makes up, are also taken into account for testing purposes. Usually, this type of signatures are represented by a signature taken by chance from the database, that is a signature sample which belongs to a different writer from the writer under consideration. For a writer in the database, genuine signatures of all the other enrolled writers are then used as fictitious signatures.

It is important to point out that during AFHA 2011 an interesting discussion in order to unify the nomenclature and classification of forgeries between PR and FHE communities took place. Since AFHA 2011 is posterior to the work in [VFTA07] where the GPDS300Signature CORPUS is described, the definitions of simple, skilled and fictitious forgeries given in that work may not coincide with the definitions agreed in AFHA 2011 (which are the ones introduced in Section 2.1, and the ones used in this Chapter). For example, fictitious signatures were called *random signatures* or *random forgeries* before the discussion held in AFHA 2011. Nevertheless, the terminology agreed in AFHA 2011 is used throughout this Chapter (and the whole Thesis) in order to keep a correspondence with the current terminology used in the literature.

3.4 SVM-based Classifier Applied to Signature Verification

As already introduced in Chapter 2, different types of classifiers have been applied to solve the offline signature verification classification problem, being those based on ANN ([OSK05] and [FSV06]), HMM ([JYBS00], [JS01], [FAT05], [JBS05]), and SVM ([FAT05], [JBS05], [OSK05], [FSV06]), among the most frequently used. HMM-based classifiers have shown to be well suited for signature modeling since they are able to capture personal variability ([JS01], [FAT05]), while SVM is a quite recent technique of statistical learning theory developed by Vapnik ([Vap95], [Vap98]). Recent works like [JBS05], [OSK05] and [FSV06] have presented com-

parisons between SVM-based classifiers and other classification methods like ANN ([OSK05], [FSV06]) and HMM [JBS05], [BGS10], showing that SVM-based classifiers present several advantages with respect to the other techniques. In fact, SVM-based classifiers has been successfully used in signature verification applications ([JBS05], [OSK05], [FSV06]) since they have the ability to work with high-dimensional data, they provide high generalization performance without the need to add *a priori* knowledge and, generally, they provide better generalization performance than other classification methods when the amount of data is small.

In this Chapter, an SVM-based classifier is used for the classification stage of the offline signature verification system proposed. Fundamentals regarding the SVM technique are introduced in Appendix C. Basically, from a set of samples from two classes, a SVM classifier tries to find the hyperplane that maximizes the distance to either class, minimizing the misclassification error. Although in their basic form SVM were developed for the purpose of learning linear threshold functions, they have been extended to the nonlinear case by means of the use of kernels. Several kernels have been proposed in the literature for SVM-based classifiers ([Bur98], [SS01]). In this work, the widespread-used linear, polynomial and Radial Basis Functions (RBF)¹ kernels are considered.

To verify a signature, that is to verify the identity claimed by a writer, the feature vector (which is calculated as described in Section 3.2) is used as the input of an SVM classifier trained for the writer under consideration. The SVM classification process will determine whether the signature belongs to the genuine class or to the forged class. Then, the signature will be assumed as genuine and the writer's claimed identity will be true if it belongs to the first class, otherwise the signature will be considered as a forgery assuming the claimed identity to be false. The signature verification experiments were performed resorting to the SVM toolbox for Matlab described in [SGGR05].

3.5 Evaluation Protocol

For the proposed verification system, an SVM model is trained for each writer. Two classes are trained: the genuine class and the forged class. Only genuine signatures are used for training purposes. In real applications neither simple nor skilled forgeries are available during the training phase, then it would not be appropriate to use them for training purposes. It is important to highlight that although the database does contain simple and skilled forgeries, avoiding their use for training purposes results in a more realistic model. To train the genuine class

¹The RBF kernel is defined as

$$K(\mathbf{x}(n), \mathbf{x}(k)) = e^{-\|\mathbf{x}(n) - \mathbf{x}(k)\|^2 / \sigma^2}.$$

a subset of 14 out of the 24 genuine signatures available per writer is used. The corresponding forged class, meanwhile, is trained using a subset of the genuine signatures (the ones separated for training purposes) of the remainder writers in the database. This set of signatures can be interpreted as fictitious signatures for the writer under consideration. For each one of the remaining writers, 5 genuine signatures are chosen randomly from the 14 genuine signatures available for training. That makes, taking into account the 160 writers enrolled in the database (described in 3.3), a total of $159 \times 5 = 795$ fictitious signatures for training.

For testing purposes, the 30 simple and skilled forgeries exclusively reserved for this purposes together with the remaining 10 genuine signatures are used. The 10 genuine signatures of each writer are used to calculate the FRR, that is, the false rejection of genuine signatures. In addition, the 10 genuine signatures of all the other writers in the database are used as fictitious signatures for computing the FAR for fictitious signatures. The 30 forged signatures (simple and skilled signatures) per writer are used to calculate the FAR for simple and skilled forgeries. Unfortunately, simple and skilled forgeries are not discriminated in the database, then a unique FAR will be computed taking into account both types of forgeries.

3.6 Results and Discussion

The principal aim of the testing phase is to evaluate the circular grid performance. For this purpose, experiments with different number of grid divisions $N = 8$, $N = 16$, $N = 32$, $N = 64$ and $N = 128$, were carried out. Fig. 3.3 shows the FRR for the circular grid approach (top), and for the rectangular grid approach (bottom), for different number of divisions and three different kernels, namely, linear, RBF, and polynomial kernels. The FAR for simple and skilled forgeries for the same number of divisions, and the same kernels is shown in Fig. 3.4, for the circular (top) and the rectangular (bottom) gridding. Similarly the FAR for fictitious signatures is shown in Fig. 3.5. The FRR and the FAR presented in Figs. 3.3, 3.4 and 3.5 have been computed as the average of the FRR and the FAR corresponding to each one of the writers used for testing, respectively.

The proposed approach shows the best results when the number of divisions of the grid goes down. For $N = 8$ and $N = 16$ the results obtained with the Polynomial Kernel are promising, specially in the case of the FAR for simple and skilled forgeries, showing the system's capability to highlight the interpersonal variability. For the FRR, results are not that good, but they still are acceptable. Particularly, the best results are obtained with 16 divisions of the grid ($N = 16$) and a polynomial kernel, being the FRR equal to 18.75%, the FAR for simple and skilled forgeries equal to 2.125% and the FAR for fictitious signatures equal to 0.0727%.

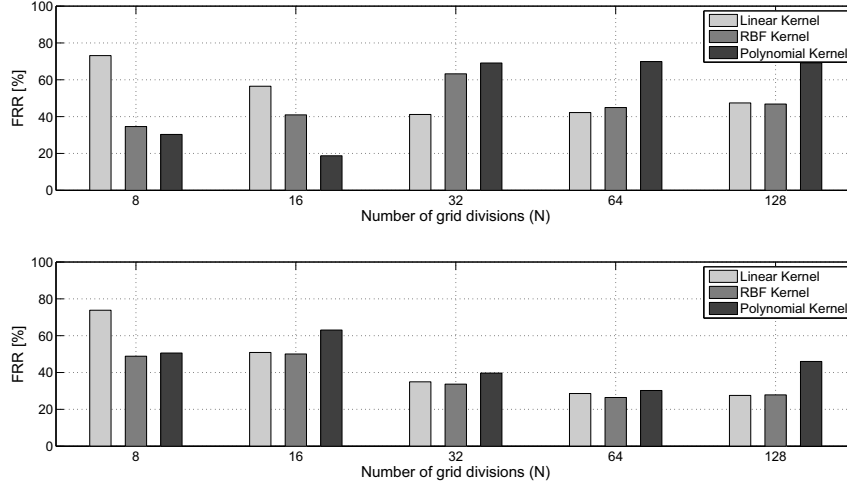


Figure 3.3: FRR for different number of divisions and kernels, for the circular (top) and the rectangular (bottom) grid approaches

For the conventional approach, instead, the best results are reached conforming the number of divisions of the grid is increased. Hence, the proposed approach has the advantage of getting good results while dealing with feature vectors in a lower dimensional space. This particularity can be related with the geometric structure of the proposed grid. As the number of divisions is increased, the size of the sectors is decreased. While not modifying the radius of the grid, only the angular amplitude of each sector is reduced. Then, the rate between the radius and the angular amplitude of a sector is not preserved. That makes the area near the centre of the grid suffer a higher relative reduction than the area near the contour, resulting in a geometric structure more sensitive to changes in the number of divisions. When the number of divisions increases the area of the sectors is so small that it does not make sense to compute the features inside them. Computing the features inside such small sectors would introduce important errors.

It can be noticed that the best results obtained with the proposed method (corresponding to $N = 16$ and polynomial kernel) show improvements with respect to the best results achieved with the rectangular grid approach (corresponding to $N = 128$ and RBF kernel). This is summarized in Table 3.1.

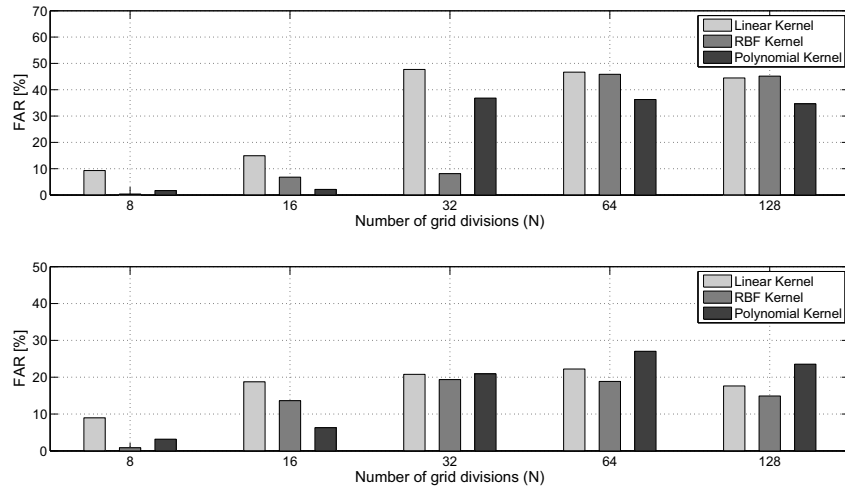


Figure 3.4: FAR (simple and skilled forgeries) for different number of divisions and kernels, for the circular (top) and the rectangular (bottom) grid approaches

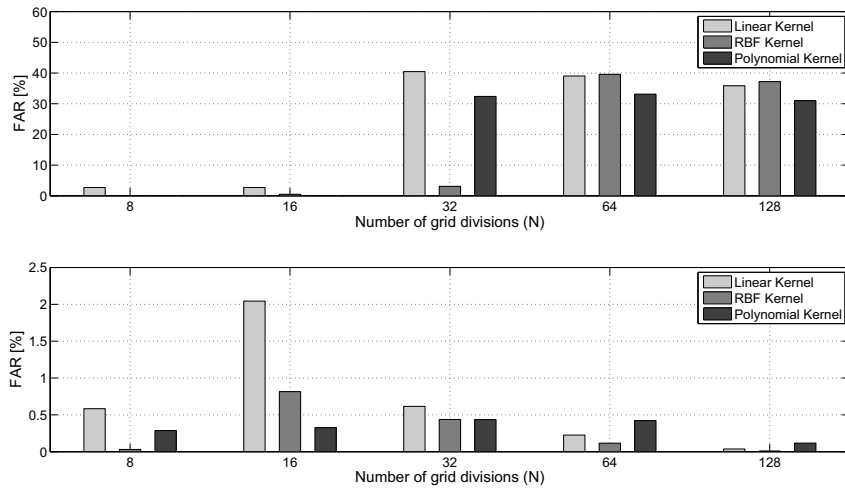


Figure 3.5: FAR (fictitious signatures) for different number of divisions and kernels, for the circular (top) and the rectangular (bottom) grid approaches

Table 3.1: Best results for circular and rectangular grid approaches

| | Circular Grid N=16, Polynomial kernel | Rectangular Grid N=128, RBF kernel |
|--------------------------------|------------------------------------------|---------------------------------------|
| FRR | 18.75% | 27.875% |
| FAR (simple and skilled forg.) | 2.125% | 14.9167% |
| FAR (fictitious signatures) | 0.0727% | 0.0106% |

3.7 Searching for Robustness against Signature Rotation

Although the verification results presented in Section 3.6 are good and outperform the ones obtained when using the traditional rectangular grids, the proposed feature extraction technique in Section 3.2 fails in solving an important issue that has to be addressed: the feature extraction should be robust against rotation of the signatures. Nor the proposed circular grid technique, neither the widely used rectangular one are robust against the signature rotation. Robustness against rotation is one of the major challenges when dealing with off-line signature verification systems. This problem has not been extensively dealt with in the literature. The idea here is to use the representation of the signature described in Section 3.2 incorporating a new step consisting in mapping the features to the Fourier Transform domain in order to achieve robustness against rotation. Similar techniques have been proposed in [WFTZ07b] and [CNPM04b].

3.7.1 Feature Extraction

The circular chart enclosing the signature showed in Fig. 3.1 is used, and the three graphometric features, namely, pixel density distribution x_{PD} , gravity centre distance x_{DGC} and gravity centre angle x_{AGC} , are computed for each sector as described in Subsection 3.2.2. Let the graphometric features for the i th sector of a signature S_0 be defined as in (3.1), (3.2) and (3.3). The i th sector of the grid is delimited by the angles $2i\pi/N$ and $2(i+1)\pi/N$. A generic feature calculated inside the i th sector of S_0 can be expressed as

$$x_{0_i} = f(S_0(\theta)), \quad \frac{2i\pi}{N} < \theta < \frac{2(i+1)\pi}{N},$$

with $i = 0, \dots, N-1$.

Suppose now that the same features are calculated inside the same i th sector of a rotated version of S_0 , namely, $S_\rho(\theta) = S_0(\theta - \rho)$ where the rotated angle is $\rho = k2\pi/N$ radians with $k = 1, 2, \dots, N$ in counterclockwise direction. The corresponding feature is given by

$$\begin{aligned}
x_{\rho_i} &= f(S_\rho(\theta)), & \frac{2i\pi}{N} < \theta < \frac{2(i+1)\pi}{N}, \\
&= f(S_0(\theta - \rho)), & \frac{2i\pi}{N} < \theta < \frac{2(i+1)\pi}{N}, \\
&= f(S_0(\theta)), & \frac{2(i-k)\pi}{N} < \theta < \frac{2(i+1-k)\pi}{N}, \\
&= x_{0_{i-k}},
\end{aligned} \tag{3.8}$$

with $i = 0, \dots, N-1$, $k = 1, \dots, N$.

From (3.8) it can be concluded that the features of the rotated signature are circularly shifted with respect to the corresponding features for the original signature.

Finally, the features of the original and rotated signatures are obtained by taking the N -point Discrete Fourier Transform (DFT) of x_{0_i} and x_{ρ_i} respectively. That is

$$\begin{aligned}
X_0(u) &= \frac{1}{N} \sum_{i=0}^{N-1} x_{0_i} e^{-j2\pi iu/N}, \\
X_\rho(u) &= \frac{1}{N} \sum_{i=0}^{N-1} x_{\rho_i} e^{-j2\pi iu/N}.
\end{aligned} \tag{3.9}$$

Replacing x_{ρ_i} by $x_{0_{i-k}}$ in (3.9) yields

$$\begin{aligned}
X_\rho(u) &= \frac{1}{N} \sum_{i=0}^{N-1} x_{0_{i-k}} e^{-j2\pi iu/N}, \\
&= \frac{1}{N} \sum_{i=-k}^{N-1-k} x_{0_i} e^{-j2\pi(i+k)u/N}, \\
&= e^{-j2\pi k/N} X_0(u),
\end{aligned} \tag{3.10}$$

with $u = 0, \dots, N-1$. It is clear then that $X_\rho(u)$ and $X_0(u)$ have the same absolute value, that is $|X_\rho(u)| = |X_0(u)|$, and then $[|X_0(0)|, \dots, |X_0(N-1)|]^T$ is a feature vector which is invariant against rotation. In addition, since the spectrum is symmetric at the central point, it is enough to keep the first $N/2 + 1$ values of the DFT for the representation of the signature.

Finally, the scaling, translation and rotation invariant feature vector X_{sign} is defined as

$$X_{sign} = [X_{PD}^T, X_{DGC}^T, X_{AGC}^T]^T, \tag{3.11}$$

where

$$\begin{aligned} X_{PD} &= [|X_{PD_0}|, |X_{PD_1}|, \dots, |X_{PD_{N/2}}|]^T, \\ X_{DGC} &= [|X_{DGC_0}|, |X_{DGC_1}|, \dots, |X_{DGC_{N/2}}|]^T, \\ X_{AGC} &= [|X_{AGC_0}|, |X_{AGC_1}|, \dots, |X_{AGC_{N/2}}|]^T. \end{aligned}$$

3.7.2 Evaluation protocol

The signatures in the GPDS300Signature CORPUS database are organized as follows: A randomly selected subset of 30 out of the 160 writers is separated and used for parameter optimisation purposes. This set of signatures is not used in the subsequent training and testing phases. The remainder 130 writers are organized as follows: The 30 forged signatures available per writer are used exclusively for testing, while the 24 genuine signatures available per writer are randomly divided into two groups. The first one, containing 13 signatures, is used for training purposes. The second one, consisting of 11 signatures, is used for testing. For each writer, the set of training samples was composed of 13 genuine signatures and 129 fictitious signature (1 genuine signature randomly chosen from the 13 available for each of the 129 remainder writers). The proportion of genuine samples to forged samples used for training was optimised over the optimisation subset of signatures.

In order to obtain reliable results, Monte Carlo techniques are used. The experiments are carried out randomly resampling the dataset into training and testing sets for each one of the writers. The resampling process is repeated 100 times.

Experiments are specially focused on testing the rotation invariance property of the proposed feature extraction technique. For that purpose, the signature model trained in each Monte Carlo instance, is tested over a dataset composed of the 11 original genuine signatures randomly chosen for testing, the 30 original forged signatures (simple and skilled forgeries) and rotated versions of both testing groups. Signatures used for testing are rotated 10, 20, 30, 40, 50 and 60 degrees in a counterclockwise direction. Experiments with different number of grid divisions $N = 8$, $N = 16$, $N = 32$, $N = 64$ and $N = 128$, and different types of kernels, namely, linear, polynomial and RBF are carried out. The internal parameters of the SVM-based classifiers, are optimised over the subset of signatures used exclusively for optimisation purposes. The best results, which are presented in Section 3.7.3, were obtained with the polynomial kernel and $N = 16$ grid divisions.

3.7.3 Results and Discussion

In order to show the improvements achieved by mapping the features extracted from the circular grid to the Fourier Transform domain, the same experiments were also carried out with the features prior to the mapping. Figure 3.6 shows the mean value of the FRR (top) and the FAR for simple and skilled forgeries (bottom) calculated with the proposed rotation invariant features (red) and the features prior to the mapping to the Fourier Transform domain (blue), averaged for the 130 writers tested. The rotation invariant property of the proposed features in the DFT domain can be clearly observed in Fig. 3.6. On the other hand, the verification errors of the features prior to the mapping to the DFT domain are strongly influenced by the rotation angle of the signature.

It is important to highlight that, in this case, the FAR for fictitious signatures was not computed. Researchers in the FHE community state that this type of attack seldom appears in real applications, and so they believe that it is not necessary to test the verification systems against them. During AFHA 2011 (posterior to the prior results presented in Section 3.6), FHE and PR researchers agreed in the current tendency of not taking into account fictitious signatures as attacks to the verification system. Then, from this Section on, this will be the case.

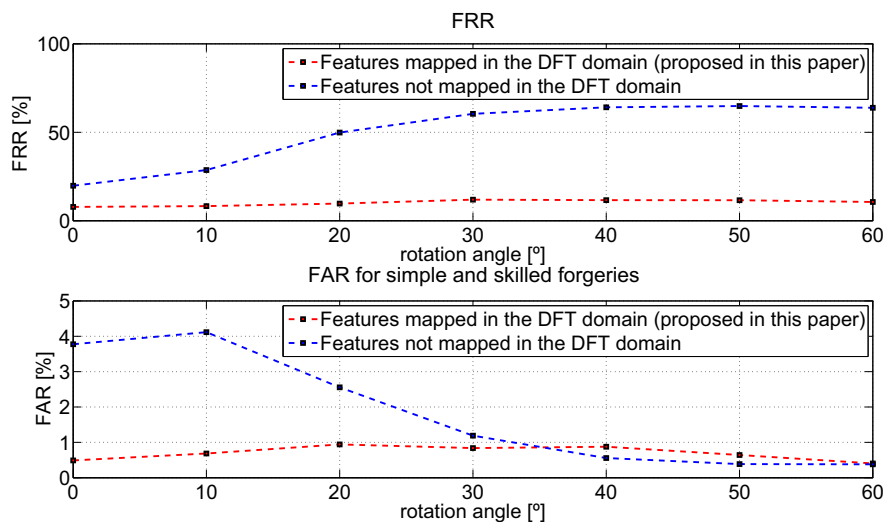


Figure 3.6: FRR (top) and FAR for simple and skilled forgeries (bottom) calculated with the proposed rotation invariant features (red) and with the features prior to the mapping to the DFT domain (blue).

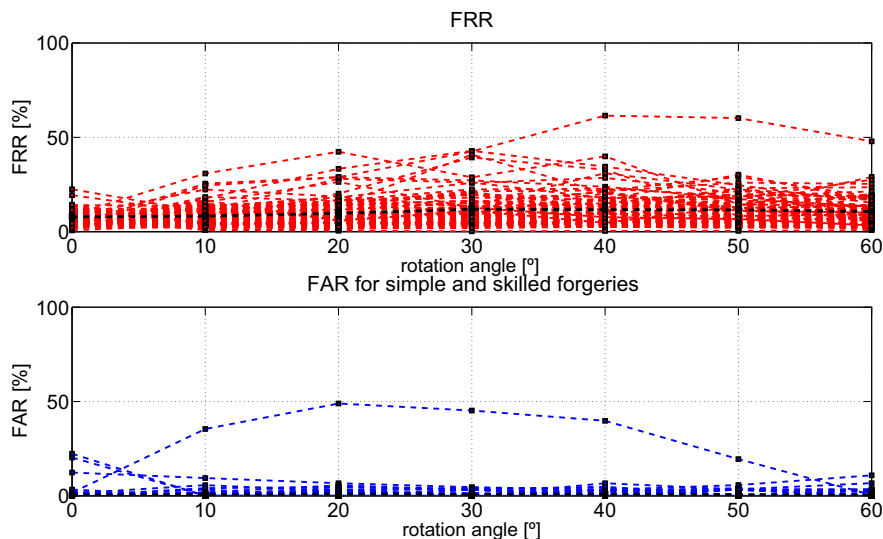


Figure 3.7: FRR (top) and FAR for simple and skilled forgeries (bottom) calculated for each of the 130 writers tested (rotation invariant features).

In Fig. 3.7 the FRR (top) and the FAR for simple and skilled forgeries (bottom) calculated with the proposed rotation invariant features, for each of the 130 writers tested are shown. From Fig. 3.7 it can be observed that the proposed method has a good performance in terms of the FAR except for only a few signatures while in terms of the FRR the performance is not that good for some of the signatures. An example of the type of signature for which the proposed verification system fails is shown in Fig. 3.8. From Fig. 3.8, it can be seen that the signatures for which the developed model is not well suited are those which present long thin lines underlying the signature which are too long with respect to the rest of the signature's body which, in these cases, is in an extreme of the underlying line. In this type of signatures, the centre of mass of the signature is not a good choice for the circular grid centre resulting in a poor verification performance.

In Table 3.2, results published in some related works together with the corresponding results obtained with the rotation invariant approach proposed in Section 3.7.1, are shown. In particular, Table 3.2 includes results from works that use a similar database (160 writers, 24 genuine signatures and 24 forged signatures) in which a polar representation of the image of the signature is used to compute the characteristic features ([FAT05], [VFTA08]) and from [PG10b] which uses the same database. Since no evaluation about the invariance property of the features are available for [FAT05] and [VFTA08], the results included in Ta-

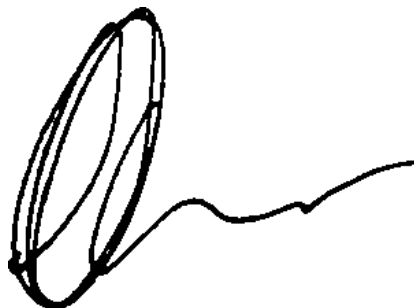


Figure 3.8: A signature that is not accurately modeled by the proposed feature extraction approach.

ble 3.2 are the ones calculated for the original signatures (without any rotation).

Table 3.2: Comparison between the results obtained with the proposed approach and other approaches proposed in the literature

| | FRR | FAR |
|------------------------|--------|--------|
| Proposed approach | 7.82% | 0.49% |
| Results in Table 3.1 | 18.75% | 2.125% |
| Ferrer et. al.[FAT05] | 14.1% | 12.6% |
| Vargas et. al.[VFTA08] | 10.01% | 14.66% |

It can be observed from Table 3.2 that the proposed method outperforms the methods of the state of the art in [FAT05] and [VFTA08], for both the FRR and the FAR.

3.8 Some Concluding Remarks

A new feature extraction approach based on a circular grid has been proposed 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 on a public database, quantified by the FRR and the FAR for simple and skilled, and fictitious signatures, using the proposed features have shown improvements with respect to the ones based on features extracted from rectangular grids.

In addition, a further advantage emerges from the use of circular grids to extract the features, that is, the possibility of developing a feature extraction

technique which is robust against the rotation of the signatures. The classification results show that the proposed signature verification system achieves the desired robustness against rotation of the signatures. In addition, the signature verification system proposed has a performance comparable to similar ones of the state of the art. In particular, the low FAR for simple and skilled forgeries obtained indicates an improvement in the capability of the system to highlight the interpersonal variability.

3.9 Some Ideas for Further Work

The presented results show that the FRR needs to be reduced. This means incrementing the verification process capability to absorb the intrapersonal variability. Efforts have to be done in this direction, while keeping the capability to remark the interpersonal variability. To address the problem, a possible strategy is to introduce new graphometric features (specially dynamic ones) and choose a suitable combination to achieve a better model of the signatures. Also modifying the number of samples used for training can be tested. Using much more signatures to train the forged class than the genuine class, may result in a system more adapted to interpersonal variability than to intrapersonal variability, and that is the case of the model used throughout this Chapter. It is likely that an improvement in the FRR value could be achieved by a better balance between the genuine and forged classes used for training.

Online Signature Verification: Features

In this Chapter, two different approximations of the time functions associated with the signing process are proposed for online signature verification. One of them is based on Legendre polynomials and the other one on wavelet decomposition. Several feature combinations are studied to evaluate their relevance for online signature verification. In addition, a consistency measure is proposed to quantify the discriminative power of the features.

4.1 Introduction

In online verification systems the signature is parameterized by different discrete time functions, *e.g.*, pen coordinates, pen pressure and pen inclination angles. Researchers have long argued about the effectiveness of the different time functions available for online signature verification. During SVC2004, the results using only pen coordinates outperformed those adding pen pressure and pen inclination angles [YCX⁺04]. Since then, several works have been presented concerning the most suitable set of parameters for modeling the signatures. In [KY05a], the authors assert that using pen coordinates leads to better performance than using pen pressure in addition to pen coordinates. In [LG05], several parameters are compared and the authors conclude that pen coordinates and some derived parameters are the most reliable ones. Even the time variability between training and testing data acquisition sessions was considered in [HGD09], where the authors conclude that pen pressure is the most unreliable time function, and pen inclination angles are too unstable (as shown by most of the researchers), being pen coordinates the most robust time functions in the presence of a long term time variability. On the other hand, some works showed an improved performance when combining the pen coordinates information with the pen pressure and pen inclination angles [MM07]. The conflicting results observed in the litera-

ture make the discussion still open. Moreover, most of the works do not consider the cultural origin of the signatures. This is a highly important issue that should be addressed in every signature verification system [PBP11b].

A desirable property for any feature is that it must have high consistency in the sense that the feature values of the genuine signatures should be close to each other while the ones of genuine and forged signatures should not. A well defined consistency model would allow to quantify the discriminative power of the features and to predict their effectiveness for verification purposes. A consistency model is introduced in [LBA96] and [LG05]. In [LBA96], the consistency model is used to select an optimal subset of global features from a larger global feature set. In [LG05], several local and global features are compared on the basis of their consistency, resulting pen coordinates and some derived features the most reliable ones. The lack of a widely used consistency model in the literature, makes its study an interesting issue.

In this Chapter, two different approximations of the time functions are proposed, one based on Legendre polynomials and another based on wavelet decomposition. The coefficients in these series expansions, computed resorting to least squares estimation techniques, are used as features to model them. Since not only the feature extraction but the feature selection is a crucial step for the verification success, several feature combinations are studied to analyse their relevance for online signature verification. Different time functions associated with the signing process are analysed in order to provide some insight on their actual discriminative power. Finally, a consistency measure is proposed to quantify the discriminative power of the features, in an attempt to make a contribution in the direction of feature selection.

The Chapter is organized as follows. The proposed feature extraction approaches are described in Section 4.2. In particular, the one corresponding to the time function approximations using Legendre polynomials is introduced in Subsection 4.2.2.1, while the one corresponding to the approximations using wavelet decomposition is presented in Subsection 4.2.2.2. The signature database used for the experiments, containing Western (Dutch) and Chinese signatures, is described in Section 4.3. Experiments carried out with different feature combinations are described in Section 4.4. The different feature combinations analysed are introduced in Subsections 4.4.1 and 4.4.2. Finally, a novel consistency measure that can be used for feature selection purposes is introduced in Section 4.5. Some concluding remarks are given in Section 4.6 and in Section 4.7 some future directions are discussed.

4.2 Feature Extraction

Several methods have been proposed in the literature for online signature verification. They differ basically in the way they perform the feature extraction

and in the classification approach they employ. The different features can be classified into local features, calculated for each point in the time sequence, and global features, calculated from the whole signature. Many researchers accept that approaches based on local features achieve better performance than the ones based on global features, but still there are others who favour the use of global features [RKD05], [GFRO05].

When using global features, feature vectors have a fixed amount of components regardless the signature length. This represents an advantage since it makes the comparison between two signatures easier with respect to the case of having different feature vector lengths. Several works in the literature have proposed a fixed-length representation of the signatures, among them [YK09], where the authors employ the FFT, can be mentioned. Further, a fixed-length model of the signatures can be required for certain biometric applications [TAK⁺00], [XVK⁺08].

In this Section, two different fixed-length representations of the signatures are proposed. One is based on the approximation of the different time functions by Legendre orthogonal polynomials, and the other on their representation using wavelets.

4.2.1 Time Functions and Preprocessing

4.2.1.1 Basic functions

Typically, the measured data consists of three discrete time functions: pen coordinates x and y , and pen pressure p . Depending on the acquisition device, the pen altitude and azimuth angles could also be available. In addition to the raw data, some other dynamic functions, such as, x and y velocities and accelerations, and log curvature radius can also be computed from them.

4.2.1.2 Normalisation

Depending on the given space to sign, signatures can be written in different sizes, writers can place them anywhere they want in the sheet of paper and many times they would sign in a rotated angle with respect to the one they usually sign. This makes size, translation and rotation normalisation fundamental preprocessing tasks.

- Size Normalisation:

A width normalisation is performed on the x and y pen coordinates of the signature. The width of the signature is previously fixed while the height is left to take the corresponding value in order to keep the original height-

to-width ratio. Then, the x and y pen coordinates are modified as:

$$x_{sn}(n) = \frac{x^o(n) - x_{min}}{x_{max} - x_{min}} \cdot W_{new}, \quad (4.1)$$

$$y_{sn}(n) = \frac{y^o(n) - y_{min}}{y_{max} - y_{min}} \cdot H_{new}, \quad (4.2)$$

where $(x^o(n), y^o(n))$ are the original point coordinates and $(x_{sn}(n), y_{sn}(n))$ are the corresponding ones after size normalisation, W_{new} is the new fixed width and H_{new} is the resulting new height computed as:

$$H_{new} = H^o \cdot \frac{W_{new}}{W^o}, \quad (4.3)$$

being W^o and H^o the original width and height, respectively.

- Translation Normalisation:

The signatures are centred by subtracting the corresponding mean values from the original x and y pen coordinates, that is:

$$x(n) = x_{sn}(n) - x_{mean}, \quad (4.4)$$

$$y(n) = y_{sn}(n) - y_{mean}, \quad (4.5)$$

where $(x_{sn}(n), y_{sn}(n))$ are the size normalised point coordinates and $(x(n), y(n))$ are the corresponding ones after translation normalisation.

- Rotation Normalisation:

Variations in the angle of signing can produced spurious variations between genuine signatures. It is common in the literature to perform a correction of the main direction of the signature by measuring its angle and rotating the signature until it has the direction of a predetermined baseline (see [FORG07] for an example). Nevertheless, it has been argued that the main direction of the signature is a distinctive feature and so, eliminating it could result in a loss of useful discriminative information. In line with this idea, in this work, no rotation normalisation has been applied.

Another widely used preprocessing technique is resampling. Due to the acquisition process, the measured data may contain noise or gaps introduced during the signing process. Resampling is used to correct these acquisition artifacts and, in addition, to get a fixed-length resampled time function. Several works in the literature use resampling to remove redundant points from the measured signals [MMO07]. In [MMO07], the effect of different resampling techniques on the verification performance is studied. The authors state that resampling has several advantages such as reducing the storage requirements and increasing the simplicity without compromising and even improving the system performance. On the

other hand, many other works in the literature do not use resampling as a pre-processing step [KY05a], [LG05], [FORG07], [YK09]. Moreover, in [KY05a] and [YK09] the authors observed that using resampling leads to worse verification performances, since it implies a significant loss of information. They conclude that it is convenient not to use resampling and that the disadvantage of not having a fixed-length signal, is not that important. Here, the proposed feature extraction techniques are based on approximations of the time functions associated with the signing process by Legendre polynomials and wavelet decomposition. The former, delivers a fixed-length feature vector, so that no resampling of the original time functions is required, while for the latter the obtained feature vector length would depend on the length of the original time function. Then, in the case of the approximation using wavelets, resampling of the original time functions is required.

4.2.1.3 Extended functions

Several extended functions that can be computed from the acquired functions have been used in the literature. In [KY05a] the incremental variations of the x and y pen coordinates are proposed. In [RKD05] several time functions, such as, the x and y velocities and accelerations and the log curvature radius, among others, are used as well as their first and second order time derivatives. In this work, the path velocity magnitude v_T , the path-tangent angle θ , the total acceleration a_T and the log curvature radius ρ [FORG07] are computed from the basic function set composed of x and y pen coordinates and pen pressure p . Let $n = 1, \dots, L_{sign}$ be the discrete time index of the measured functions and L_{sign} the time duration of the signature in sampling units, then the above mentioned extended functions are computed as:

- Path velocity magnitude: $v_T(n) = \sqrt{\dot{x}^2(n) + \dot{y}^2(n)}$
- Path-tangent angle: $\theta(n) = \arctan(\dot{y}(n)/\dot{x}(n))$
- Total acceleration: $a_T(n) = \sqrt{\dot{v}_T^2(n) + c^2(n)}$, where: $c(n) = v_T(n) \cdot \dot{\theta}(n)$
- Log curvature radius: $\rho(n) = \log(v_T(n)/\dot{\theta}(n))$

In all cases, the first order time derivatives are computed as [FORG07]:

$$\dot{f}(n) \approx \Delta f(n) = \frac{\sum_{\tau=1}^2 \tau(f(n+\tau) - f(n-\tau))}{2 \cdot \sum_{\tau=1}^2 \tau^2} . \quad (4.6)$$

Throughout this Chapter the whole set of features will be composed by the x and y pen coordinates, the pressure p , the above mentioned extended functions, *viz.* v_T , θ , a_T and ρ , their first order time derivatives (these are the features proposed in [FORG07]), and their second order time derivatives computed as in

(4.6) from the corresponding first order time derivatives. That is, $x, y, p, v_T, \theta, a_T$ and ρ , their first order time derivatives $dx, dy, dp, dv_T, d\theta, da_T$ and $d\rho$, and their second order time derivatives $d^2x, d^2y, d^2p, d^2v_T, d^2\theta, d^2a_T$ and $d^2\rho$, will be the extended time functions considered in this Chapter.

4.2.2 Orthogonal Time Function Representation

4.2.2.1 Time Function Approximation via Legendre Polynomials

In this Subsection, models of the time functions associated with the signing process, based on Legendre series approximations, are presented. The coefficients of the Legendre series are computed resorting to least squares techniques. A similar approach for the representation of handwritten digits has been proposed in [GW10], where the coefficients of the series approximation are computed using function moments.

- Orthogonal polynomials series expansions

A family of functions $\{g_i\}$ in (in general) an infinite dimensional functional space $H([a, b])$, defined in the domain $[a, b]$, is said to be orthonormal with respect to an inner product $\langle \cdot, \cdot \rangle$ in $H([a, b])$ if $\langle g_i, g_j \rangle = \delta_{ij}$ for all $\{i, j\}$, where δ_{ij} is the Kronecker delta.

Provided the inner product space $H([a, b])$ is complete with respect to the metric induced by the inner product, a set of orthonormal basis functions $\{h_i\}_{i=1}^{\infty}$ can be defined. In this case, any function $f \in H([a, b])$ can be uniquely represented by a series expansion in the orthonormal basis, that is

$$f = \sum_{i=1}^{\infty} \alpha_i h_i, \quad (4.7)$$

where

$$\alpha_i = \langle f, h_i \rangle. \quad (4.8)$$

It is not difficult to prove that the best (in the sense of the metric induced by the inner product) approximation of $f \in H([a, b])$ in an N -dimensional subspace is given by

$$f \approx \sum_{i=1}^N \alpha_i h_i. \quad (4.9)$$

- Coefficient estimation

The idea here is to approximate the time functions measured during the signature acquisition stage by a finite series expansion in orthonormal polynomials in the interval $[0, 1]$, and to use the series expansion coefficients

as features. Particularly, Legendre polynomials are considered in this case, and the approximation equation (4.9) becomes

$$f(t) \approx \sum_{i=1}^N \alpha_i L_i(t), \quad (4.10)$$

where $L_i(t)$ are the orthonormal Legendre polynomials¹ normalised to the interval $[0, 1]^2$.

Since the time functions $f(t)$ are unknown, the coefficients in the truncated series expansions (4.10) cannot be computed as in (4.8) but rather they have to be estimated from a set of M (usually larger than N) samples of the function at the time instants $\{t_1, t_2, \dots, t_M\}$.

In matrix form, equation (4.10) at the time instants $\{t_1, t_2, \dots, t_M\}$ can be written as

$$\underbrace{\begin{bmatrix} f(t_1) \\ f(t_2) \\ \vdots \\ f(t_M) \end{bmatrix}}_{\mathbf{f}} = \underbrace{\begin{bmatrix} L_1(t_1) & L_2(t_1) & \cdots & L_N(t_1) \\ L_1(t_2) & L_2(t_2) & \cdots & L_N(t_2) \\ \vdots & \vdots & \ddots & \vdots \\ L_1(t_M) & L_2(t_M) & \cdots & L_N(t_M) \end{bmatrix}}_{\mathbf{L}} \underbrace{\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_N \end{bmatrix}}_{\boldsymbol{\alpha}} \quad (4.11)$$

It is well known that the solution $\hat{\boldsymbol{\alpha}}$, in the least squares sense, of the over determined system of equations (4.11) is given by $\hat{\boldsymbol{\alpha}} = \mathbf{L}^\dagger \mathbf{f}$, where $\mathbf{L}^\dagger = (\mathbf{L}^T \mathbf{L})^{-1} \mathbf{L}^T$, stands for the left pseudo-inverse of \mathbf{L} .

To illustrate the above estimation procedure, the x and y pen coordinates associated with a signature, and the corresponding approximations using Legendre polynomials with orders $N = 21$, $N = 15$ and $N = 10$, are shown in Fig. 4.1.

The Best FIT³ between the measured and the approximated time functions, for the above mentioned Legendre polynomial orders, are given in Table 4.1.

¹The polynomials are orthonormal with respect to the standard inner product

$$\langle h_i(t), h_j(t) \rangle = \int_0^1 h_i(\tau) h_j(\tau) d\tau.$$

²Typically, Legendre polynomials are defined in the interval $[-1, 1]$.

³The Best FIT is defined as:

$$\text{Best FIT} = 100 \left(1 - \frac{\|x - x_{approx}\|}{\|x - x_{mean}\|} \right). \quad (4.12)$$

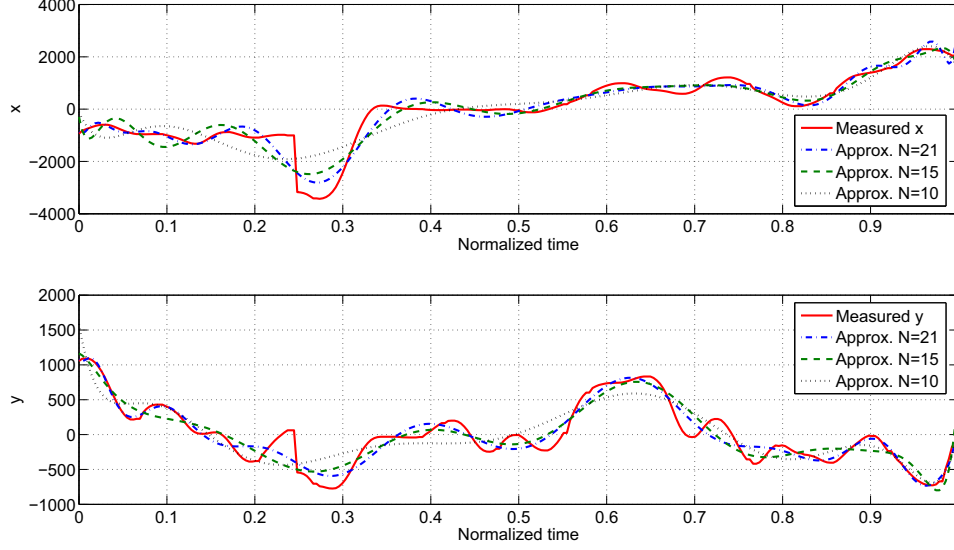


Figure 4.1: Time functions: x and y pen coordinates (red solid line) and their corresponding approximations by Legendre polynomials with orders $N = 21$ (blue dashed line), $N = 15$ (green dash-dotted line) and $N = 10$ (black dotted line).

It can be observed that a reasonable FIT is obtained for $N = 21$. Experimental results show that further increasing the polynomial orders does not substantially improve the approximation accuracy. This is an expected result, taking into account the bias-variance tradeoff inherent to least squares estimation from noisy data.

Table 4.1: Best FIT between the measured and the approximated time functions using Legendre polynomials of order N .

| N | 21 | 15 | 10 |
|-------------|---------|---------|---------|
| FIT_x [%] | 77.7955 | 68.9708 | 57.6664 |
| FIT_y [%] | 70.7341 | 62.9579 | 53.3995 |

4.2.2.2 Time function approximation via wavelets

An orthonormal wavelet bases in an inner product space can be generated by dilation and translation of a mother wavelet [Dau92]. An approximation of a function f similar to the one in (4.13) can also be performed using these wavelet bases. This linear approximation can be improved if one chooses *a posteriori*

the N bases h_i , depending on the function f , in such a way to minimise the approximation error. This is done by choosing the set of N bases that have the largest inner product amplitudes $|\langle f, h_i \rangle|$. The approximation would then be as follows:

$$f \approx \sum_{i \in I_N} \alpha_i h_i, \quad (4.13)$$

where I_N is an index set containing the indices corresponding to the largest inner products amplitudes.

This, results in a nonlinear approximation scheme since the approximation vectors change with the function f . Since the amplitude of the inner products in a wavelet bases is related to the regularity of the signal, the approximation scheme is equivalent to constructing and adaptive approximation grid, whose resolution is locally adapted to the signal regularity. For signals with isolated singularities, the wavelet-based approximation is more precise than a linear scheme, which maintains the same resolution over the whole signal support.

For the case of discrete time functions, the inner products are computed resorting to the Discrete Wavelet Transform (DWT) [Dau92], which decomposes the signal at different levels of resolutions, splitting it in low frequency (approximation) and high frequency (details) components.

The idea is to perform a multilevel decomposition of the time functions using the DWT and to use the approximation coefficients to represent them. As mentioned in Section 4.2.1, resampling of the time functions, previous to the DWT decomposition, is needed in order to have a fixed-length feature vector. An approach where all the coefficients, instead of only the approximation ones, are used to represent the time functions within the framework of online signature verification was presented in [CDW⁺12].

Figure 4.2 schematically depicts a filter bank representation of the multilevel (level of resolution $\ell = 2$) decomposition (left) and the approximate reconstruction⁴ (right) of the discrete time function $f(n)$ using the DWT. In the figure, LP represents a low pass filter having the scaling function as its impulse response, while HP represents a high pass filter having the mother wavelet as its impulse response. As mentioned above, the DWT ℓ -level approximation coefficients \mathbf{a}_ℓ will be used to model the corresponding time function.

To illustrate the above approximation procedure, the x and y pen coordinates associated with the same signature analysed in the case of the Legendre approximations in Fig. 4.1, and the corresponding approximations using the DWT (with the **db4** wavelet [Dau92]) with levels of resolutions $\ell = 1$, $\ell = 2$ and $\ell = 3$, are shown in Fig. 4.3. The time functions were resampled so that the resulting length is 256.

The Best FIT between the measured and the approximated time functions, for

⁴Note that only the approximation coefficients are used for the reconstruction of the signal in the right side of Fig. 4.2.

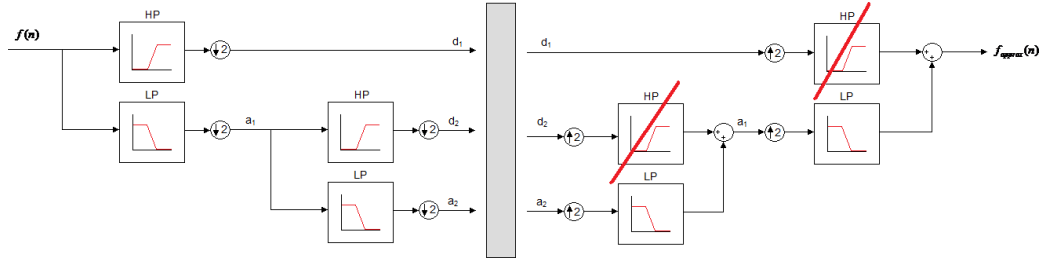


Figure 4.2: Filter bank scheme of the multilevel ($\ell = 2$) decomposition (left) and the approximate reconstruction (right).

the above mentioned levels of resolution, are given in Table 4.2. Also shown are the lengths of the resulting feature vectors for the different levels of resolution. It can be observed that reasonable FITs are obtained for the three levels of resolution considered in the table. The design parameter will then be the length of the resulting feature vector which will determine the level of resolution to be used.

Table 4.2: Best FIT between the measured and the approximated time functions using the DWT (db4). The lengths of the resulting feature vectors are shown in the last row.

| ℓ | 1 | 2 | 3 |
|--------------------------|---------|---------|---------|
| $FIT_x [\%]$ | 94.1014 | 93.2690 | 86.9960 |
| $FIT_y [\%]$ | 94.9623 | 93.8578 | 82.9524 |
| Length \mathbf{a}_ℓ | 131 | 69 | 38 |

4.3 Signature Database

The influence of the cultural origin of the signatures in the performance of the verification systems is important to be taken into account. There are not many works in the literature that consider non-Western signatures such as Chinese, Japanese, Arabic, etc.. In [PBP11b], an updated survey of non-English and non-Latin signature verification systems can be found. In that work, the authors point out that there are still many challenges in this research area. In line with this idea, the publicly available SigComp2011 Dataset [LMd⁺11] presented within ICDAR 2011 is used. This database is described in Appendix B. It has two separate datasets, one containing genuine and forged Western signatures (Dutch ones) and the other containing genuine and forged Chinese signatures. The available forgeries are skilled forgeries, which are imitations of the genuine signatures where the forger is allowed to practise the forgery for as long as she deem it necessary.

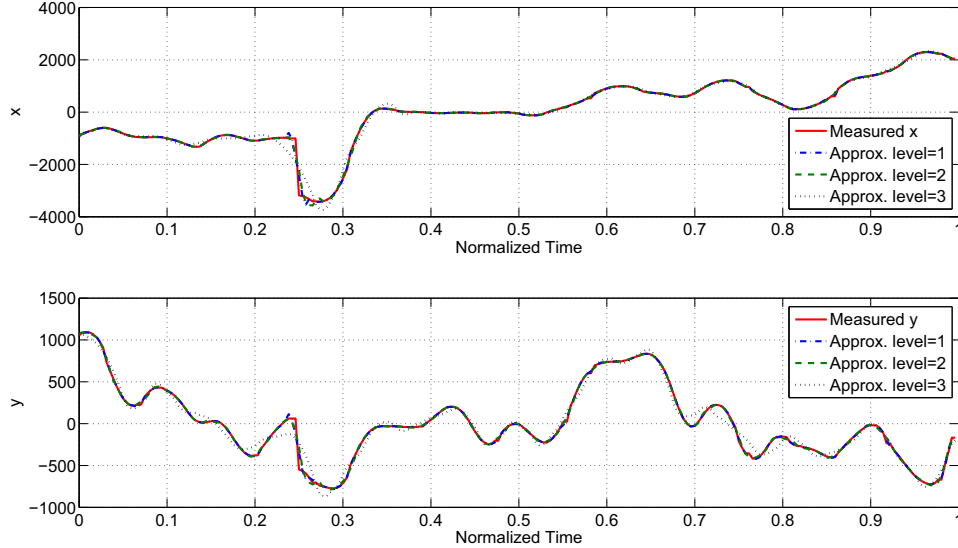


Figure 4.3: Time functions: x and y pen coordinates (red solid line) and their corresponding approximations by the DWT (db4) with levels of resolutions $\ell = 1$ (blue dashed line), $\ell = 2$ (green dash-dotted line) and $\ell = 3$ (black dotted line).

In Fig. 4.4, offline (left) and online (right) versions of a typical genuine (top) and forged (bottom) Dutch signature are shown. In Fig. 4.5, offline (left) and online (right) versions of a typical genuine (top) and forged (bottom) Chinese signature are shown. Each of the datasets in the SigComp2011 Dataset is divided into two sets, namely, the Training and Testing Sets. The measured data consists of three discrete time functions: pen coordinates x and y , and pen pressure p .

4.4 Feature Combination Experiments

In recent years, several works have been presented in the literature concerning the robustness of the different time sequences available to model the signatures. Nevertheless, the conflicting results presented so far makes the discussion still open. The idea of the experiments presented in this Chapter is to analyse different feature combinations based on different feature selection criteria in an attempt to give some answers to this widespread discussion about which are the best features for online signature verification.

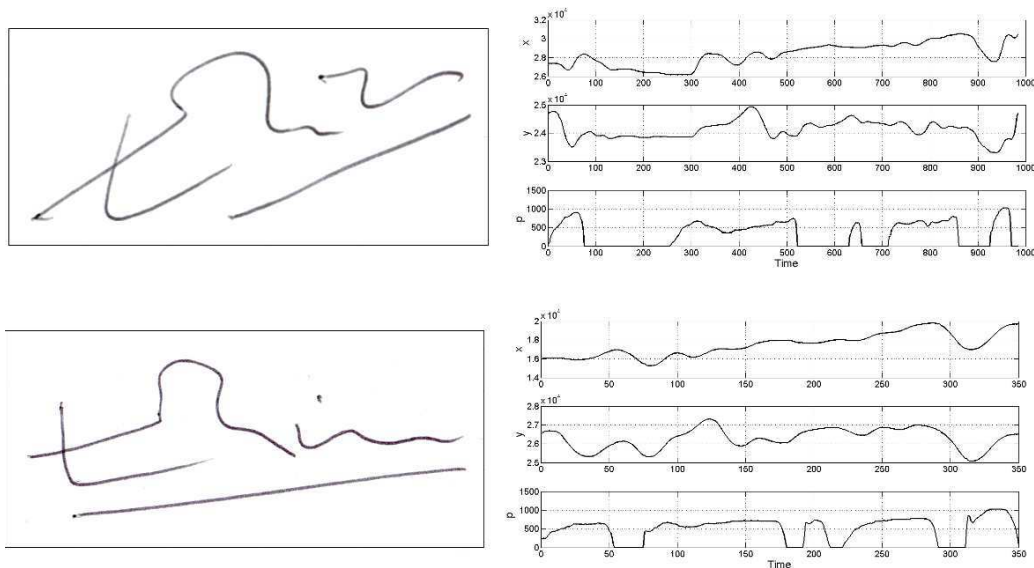


Figure 4.4: Offline (left) and online (right) versions of a typical genuine (top) and forged (bottom) Dutch signature.

4.4.1 Combination Experiments Using a Basic Features Set

In a first attempt to make some contributions towards the analysis of the discriminative power of the different time functions, a simple set of time functions is considered. This basic set contains only the measured data x , y and p , and the extended time functions v_T , θ , a_T and ρ , out from the whole set of time functions described in Subsection 4.2.1.3. All the possible combinations of these time functions are taken into account. For the sake of compactness of the notation, the different combinations, are coded as indicated in Table 4.3. Note the reader that whenever the x coordinate is considered, also the y coordinate is included in the combination, since no preferential direction would be expected to exist a priori. The same holds for the total velocity (magnitude v_T and direction θ).

4.4.1.1 Evaluation Protocol

The considered time functions for these experiments, *viz.*, x , y , p , v_x , v_y , a_T and ρ , are represented by their coefficients in the corresponding Legendre series expansion as described in Subsection 4.2.2.1. In addition, in order to analyse the robustness of the time functions against the use of different classification schemes, two well known state-of-the-art classifiers are used to perform the verification experiments, namely, SVM [Vap95] and RF[Bre01]. In recent years, SVM based classifiers have been successfully applied to the signature verification problem,

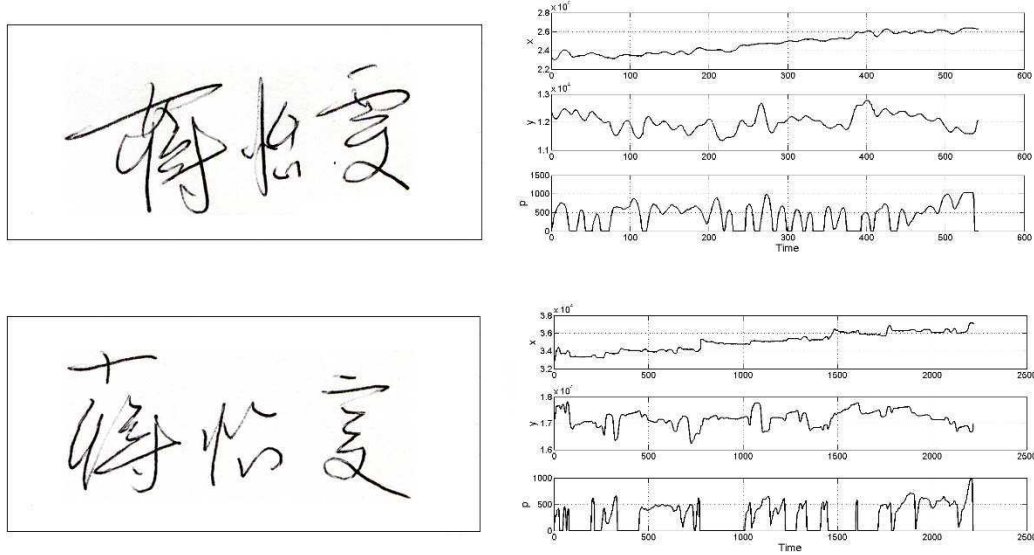


Figure 4.5: Offline (left) and online (right) versions of a typical genuine (top) and forged (bottom) Chinese signature.

as it was already shown in Chapter 3 for the offline case. In addition, several works have also been presented using SVM based classifiers for the online case [GGKS10], [KD12]. On the other hand, the advantages of RF based classifiers have not been widely exploited yet for the signature verification problem. Fundamentals of SVM based classifiers are given in Appendix C (as already mentioned in Chapter 3), while fundamentals of classifiers based on RF are given in Appendix D.

For each dataset, namely, Chinese and Dutch, the optimisation of the meta-parameters of the system is performed over the corresponding Training Set while the corresponding Testing Set is used for independent testing purposes.

The tuning parameters to adjust are the order of the Legendre polynomials and the internal parameters of the classifiers. To select the most suitable order for the Legendre polynomials, tests varying this parameter from 1 to 25 were carried out. For the SVM classifier, the tuning parameters⁵ are the scale σ^2 in the RBF kernel, and the regularization parameter $C > 0$ providing a tradeoff between model complexity and the training error, in the SVM cost function. The linear and polynomial kernels were also tested, but the RBF gave the best results. For the RF classifier, there are basically two tuning parameters to adjust, namely, the number of trees to grow and the number of randomly selected splitting variables to be considered at each node. In general, the sensitivity to those parameters is not meaningful [LW02], and the default values are a good choice.

⁵These parameters were optimised using the `tune.svm` routine of the `e1071` Package [CVGR05], with values within the range 10^{-10} to 10^{10} .

Table 4.3: Tested feature combinations.

| | | | | | |
|---------------|------------------------|--------------------|------------------|-----------------------|----------------------|
| Code | 1 | 2 | 3 | 4 | 5 |
| Feature Comb. | xy | p | $v_T\theta$ | a_T | ρ |
| Code | 6 | 7 | 8 | 9 | 10 |
| Feature Comb. | xyp | $xyv_T\theta$ | xya_T | $xy\rho$ | $pv_T\theta$ |
| Code | 11 | 12 | 13 | 14 | 15 |
| Feature Comb. | pa_T | $p\rho$ | $v_T\theta a_T$ | $v_T\theta\rho$ | $a_T\rho$ |
| Code | 16 | 17 | 18 | 19 | 20 |
| Feature Comb. | $xypv_T\theta$ | $xypa_T$ | $xyp\rho$ | $xyv_T\theta a_T$ | $xv_T\theta\rho$ |
| Code | 21 | 22 | 23 | 24 | 25 |
| Feature Comb. | $xya_T\rho$ | $pv_T\theta a_T$ | $pv_T\theta\rho$ | $pa_T\rho$ | $v_T\theta a_T\rho$ |
| Code | 26 | 27 | 28 | 29 | 30 |
| Feature Comb. | $xypv_T\theta a_T$ | $xypv_T\theta\rho$ | $xypa_T\rho$ | $xyv_T\theta a_T\rho$ | $pv_T\theta a_T\rho$ |
| Code | 31 | | | | |
| Feature Comb. | $xypv_T\theta a_T\rho$ | | | | |

To obtain statistically significant results, a 5-fold cross-validation (5-fold CV) is performed over the Testing Set to estimate the testing errors. For each instance of the 5-fold CV, a signature model is trained for each writer in the dataset. As already explained in Section 3.5, skilled forgeries are not available during the training phase, then only genuine signatures are used for training purposes. When training a signature model for a particular writer, two classes are involved, namely, genuine and forged. For training the genuine class, the subset of genuine signatures of that writer available in the corresponding training set of the 5-fold CV is used, while the subset of genuine signatures of all the remaining writers in the dataset available in the corresponding training set of the 5-fold CV is used as fictitious signatures for training the forged class. For testing purposes, the subset of genuine and forged signatures of the writer under consideration available in the corresponding testing set of the 5-fold CV are used. Only skilled forgeries are considered to calculate the testing errors, since fictitious signatures, as already discussed in Section 3.7.3, seldom appear in real situations.

To evaluate the performance, the EER is calculated, using the Bosaris toolkit [Bd06], from the DET curve as the point in the curve where the FRR equals the FAR. The cost of the log-likelihood ratios \hat{C}_{llr} and its minimal possible value \hat{C}_{llr}^{min} [Bd06] are computed using the toolkit as well. A smaller value of \hat{C}_{llr}^{min} indicates a better performance of the system. Using these measurements to evaluate the performance of a signature verification system has been proposed in [LMd⁺11], where the importance of computing the likelihood ratios was highlighted since they allow FHEs to give an opinion on the strength of the evidence [GFRO05],

although they are not in the position to make a leap of faith and judge about guilt or no guilt.

4.4.1.2 Results and Discussion

The verification performance for each combination in Table 4.3 is quantified by the EER , \hat{C}_{llr} and \hat{C}_{llr}^{min} , over the Dutch and Chinese Testing Sets. The experiments were performed using the state-of-the-art classification techniques RF and SVM. For the RF classifier, the number of trees was set to 500 and the number of randomly selected splitting variables was equal to \sqrt{P} , where P is the dimension of the feature vector, for both datasets. For the SVM classifier the internal parameters were set to the optimal values $\sigma^2 = 10^7$ and $C = 1$, for the Dutch dataset and $\sigma^2 = 10^7$ and $C = 10$ for the Chinese dataset. Finally, the order of the Legendre polynomials was set to $N = 21$, for both datasets, since as already discussed in Section 4.2, further increasing the order does not improve the approximation accuracy. Figures 4.6 and 4.7 show the verification errors corresponding to the Dutch and the Chinese Testing Sets, respectively, when using RF (left) and SVM (right) as the classifiers, for all the feature combinations in Table 4.3. In Figs. 4.6 and 4.7, the feature combinations have been included in a nondecreasing order of the errors, from left (best performance) to right (worst performance).

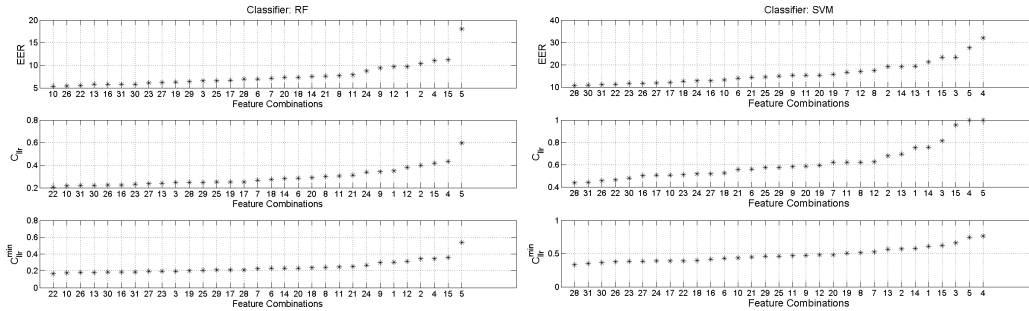


Figure 4.6: EER (top), \hat{C}_{llr} (middle) and \hat{C}_{llr}^{min} (bottom), for the Dutch Testing Set when using RF (left) and SVM (right) as classifiers.

Figure 4.6 shows that, for the Dutch data, whenever the pen pressure or the acceleration are included in a feature combination, the error rate is improved, independently of the classifier being used. Further, it can be noted that the inclusion of the velocity ($v_T\theta$) is also helpful since it is contained in the feature combinations that perform better, for both classifiers. For the Chinese data (Fig. 4.7) it is also the case that whenever the pen pressure or the acceleration is included in a feature combination, the error rate is improved, independently of the classifier being used. Like for the Dutch data, the velocity ($v_T\theta$) is also useful, for both classifiers. Finally, it is important to remark that, for this data,

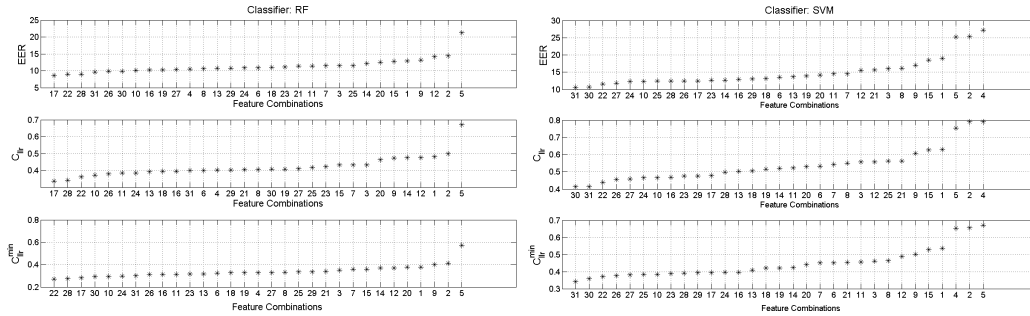


Figure 4.7: EER (top), \hat{C}_{ur} (middle) and \hat{C}_{ur}^{min} (bottom), for the Chinese Testing Set when using RF (left) and SVM (right) as classifiers.

the x and y pen coordinates are included in most of the best combinations (in the sense of the verification errors), for both classifiers. This suggests that the position information is highly discriminative for this type of data. Moreover, comparing the improvements in the verification performances when using the position information in the case of the Duch data and the Chinese data, it would be reasonable to say that the position information is likely to be more useful for the Chinese data than for the Dutch data. Chinese signature style is, in most of the cases, close to the Chinese handwriting style, consisting of one or more multi-trace characters, while Western signatures can adopt several different styles. Due to the nature of Chinese characters, it is likely that the position information (x and y pen coordinates) has more discriminative power than in the case of Dutch data.

As already mentioned above, from Figs. 4.6 and 4.7 it can be seen that incorporating the pen pressure improves the verification results, independently of the classifier being used, revealing a high discriminative power of the feature by itself. In addition, and maybe more interestingly, incorporating the pen pressure has been shown to improve the verification results independently of the dataset being considered. This could be suggesting that the reliability of the pen pressure is not influenced by the considered cultural origin of the signatures, depending mainly on the writer. Then, it would be reasonable to state that the pen pressure is a useful feature regardless of the cultural origin of the signatures. Although in the present experiments only Dutch (as an example of Western signatures) and Chinese signatures are considered, the independence of the data is very important since it means that the pen pressure could be a useful feature not highly influenced by the signature style, depending mostly of the writer. Of course, it is mandatory to analyse more data from different cultures in order to make further conclusions, but this observation can be a promising starting point.

The discriminative power of the pen pressure has long been questioned in the literature of online signature verification. This discussion, as it was mentioned in Section 4.1, has particularly been intensified since the SVC2004 was

held [YCX⁺04]. The results discussed here, suggest that the pen pressure is a useful feature when combined with other features. This observation agrees with many other reported results in the literature, such as [MM07] and [RKD05], where it is stated that the pen pressure is a useful feature to distinguish between writers when used in combination with other time functions. In addition, the results suggest that the usefulness of the pen pressure is independent of the classifier being used and not significantly influenced by the cultural origin of the signatures being considered, which is clearly an important advantage of the feature.

Table 4.4 shows the best results for the Dutch (left) and Chinese (right) data obtained using RF and SVM as classifiers. Regarding the Dutch data, the best results are obtained by the feature combinations $pv_T\theta a_T$ (for RF) and $xypa_T\rho$ (for SVM), while for the Chinese data, the best results are obtained by the feature combinations $pv_T\theta a_T$ (for RF) and $xypv_T\theta a_T\rho$ (for SVM). This makes sense since the pen pressure and the acceleration are reliable features, for both datasets. In the case of using the RF classifier, including the velocity ($v_T\theta$) helps to get better results, while in the case of using the SVM classifier it is necessary to incorporate more features to get better results. In addition, the results obtained when using RF are better than the ones obtained with SVM. Unfortunately, there are still no conclusive results in the literature regarding which one, between RF and SVM, is the best classifier, independently of the chosen features, in applications of handwriting recognition. For instance, the results in [BP10] show that SVM outperforms RF as a classifier, for the particular features (different from the ones chosen here) considered in that paper, while in [PG12] and [PGL12] the results using RF outperform the ones using SVM.

Table 4.4: Best verification results for the Dutch (left) and Chinese (right) Datasets, for both classifiers.

| Dutch Dataset | | | | | Chinese Dataset | | | | |
|------------------|--------|-------|-----------------|-----------------------|------------------------|--------|-------|-----------------|-----------------------|
| Features | Class. | EER | \hat{C}_{llr} | \hat{C}_{llr}^{min} | Features | Class. | EER | \hat{C}_{llr} | \hat{C}_{llr}^{min} |
| $pv_T\theta a_T$ | RF | 5.5 | 0.2039 | 0.1652 | $pv_T\theta a_T$ | RF | 8.93 | 0.3620 | 0.2722 |
| $xypa_T\rho$ | SVM | 10.68 | 0.4368 | 0.3323 | $xypv_T\theta a_T\rho$ | SVM | 10.54 | 0.4139 | 0.3419 |

From Table 4.4 it can be noticed that for the Chinese data more features are needed than for the Dutch data in order to get better results. In addition, the results for the Dutch signatures are better than those for the Chinese ones. Generally speaking, Chinese signatures appear to be more complex than Dutch signatures, in the sense that they have multiple separated characters composed by multiple traces leading to discontinuities in the time functions associated with the signing process. Then it is not surprising that more features are needed to model them in order to reach better verification results and that these results are not as good as in the case of the Dutch data. This is in line with the observations

in [LMd⁺11], indicating that Chinese signatures are more challenging and that a lot of research has to be done on this type of data.

For the purposes of comparison, the best results obtained in these experiments are shown in Table 4.5 together with the best commercial and non-commercial systems in the SigComp2011 competition [LMd⁺11] for the Dutch (left) and Chinese (right) data. Since the *EER* was not reported in the competitions results in [LMd⁺11], the accuracy (Acc) has been included instead. The performance comparison is performed taking into account the \hat{C}_{llr}^{min} since, as already mentioned in Subsection 4.4.1.1, the lower the \hat{C}_{llr}^{min} is the better the system performance will be. Even though the results are not as good as the corresponding to the best commercial system (*xyzmo*⁶, see [LMd⁺11]), they would have ranked first among the non-commercial systems and second among all the participants.

Table 4.5: Best verification results for the Dutch (left) and Chinese (right) Datasets.

| | | Dutch Dataset | | | Chinese Dataset | | |
|---------------------|--------|-----------------|-----------------------|-----------------------|-----------------|-----------------------|-----------------------|
| Features | Class. | EER | \hat{C}_{llr} | \hat{C}_{llr}^{min} | EER | \hat{C}_{llr} | \hat{C}_{llr}^{min} |
| $pv_T\theta a_T$ | RF | 5.5 | 0.2039 | 0.1652 | 8.93 | 0.362 | 0.2722 |
| System | Acc. | \hat{C}_{llr} | \hat{C}_{llr}^{min} | Acc. | \hat{C}_{llr} | \hat{C}_{llr}^{min} | |
| commercial | 96.27 | 0.2589 | 0.1226 | 93.17 | 0.4134 | 0.2179 | |
| 1st. non-commercial | 93.49 | 0.4928 | 0.2375 | 84.81 | 0.5651 | 0.3511 | |

4.4.2 Combination Experiments Using a More Complex Feature Set

It would be interesting to evaluate the discriminative power of more features than in the previous experiments presented in Subsection 4.4.1. The whole set of features presented in Subsection 4.2.1.3, that is, x , y , p , v_T , θ , a_T , ρ , their first order time derivatives, dx , dy , dp , dv_T , $d\theta$, da_T , $d\rho$, and their second order time derivatives d^2x , d^2y , d^2p , d^2v_T , $d^2\theta$, d^2a_T , $d^2\rho$, is then considered as the initial set of features in this Subsection. In general, it is believed that the more features are used for representing the signatures the best the results will be. This is not strictly true, since it is usually more important the significance of the features than the number of them. The idea here is to analyse different feature combinations selected out of the whole set of features conforming different selection criteria. In particular, the following feature combinations will be considered:

- **Automatically selected features (ASF):** Automatic feature selection is a useful technique not only to reduce the size of an original set of features, but also to provide some insight on the actual discriminative power

⁶<http://www.xyzmo.com>

of the features. An automatic feature selection based on the variable importance provided by the Random Forest algorithm (further described in Appendix D) is then proposed. As already mentioned in Section 4.3, the datasets in the SigComp2011 Database are divided into the Training and Testing Sets for the Dutch and Chinese data. The feature selection is performed over the Training Sets for both datasets.

- **Features in [FORG07]:** The initial set of features proposed in this Subsection is based on a set of widely used online features in the literature. In [FORG07] the measured and extended time functions (x , y , p , v_T , θ , a_T and ρ) and their first order time derivatives (dx , dy , dp , dv_T , $d\theta$, da_T and $d\rho$) are used. It is always interesting to compare the different selection criteria against the criteria used by other authors. For this reason, the feature combination used in [FORG07] is also considered here.
- **Original set of features:** The measured (x , y and p) and the extended functions (v_t , θ , a_T and ρ). Note that this is the same feature set analysed in the previous experiments (see Subsection 4.4.1).
- **All features:** The whole set of features presented in Subsection 4.2.1.3, that is, x , y , p , v_T , θ , a_T , ρ , their first order time derivatives, dx , dy , dp , dv_T , $d\theta$, da_T , $d\rho$, and their second order time derivatives d^2x , d^2y , d^2p , d^2v_T , $d^2\theta$, d^2a_T , $d^2\rho$.

4.4.2.1 Evaluation Protocol

In this case, the two representations of the time functions proposed in Section 4.2 are used, that is, the one based on Legendre polynomials described in Subsection 4.2.2.1 and the one based on the DWT described in Subsection 4.2.2.2. A comparison between these two different approximations is then carried out on the basis of the verification results obtained when using each one of them. In addition, for the representation based on the DWT, two different wavelets, namely, **db4** and **bio6.8**, were employed.

The verification results of the previous experiments presented in Subsection 4.4.1.2, showed that the RF based classifier outperformed the SVM based one regarding the verification error rates. Then, a RF classifier will be used for the experiments carried out with the time function combinations defined in Subsection 4.4.2.

For the case of Legendre polynomials representations, the tuning parameter to adjust is the order of the Legendre polynomials. As already done for the experiments in Subsection 4.4.1, to select the optimal order, this parameter was varied from 1 to 25. For the representation based on DWT approximations, the user has to choose the mother wavelet, the length of the resampled functions and the level of resolution for the approximation. In this case, two different

mother wavelets, namely, **db4** and **bio6.8**, were used. The length of the resulting feature vector is determined by the length of the resampled functions and the level of resolution. Regarding the RF classifier, the parameters to adjust are the number of trees to grow and the number of randomly selected splitting variables to be considered at each node. As already mentioned in Subsection 4.4.1.1, the default values are usually a good choice for these parameters. Then, in addition of outperforming SVM based classifiers (for the present application), the RF based classifier also has the advantage of not being hard to adjust.

To obtain statistically significant results, a 5-fold cross-validation (5-fold CV) like the one already described in Subsection 4.4.1.1 is performed over the Testing Set to estimate the testing errors. Also here, the performance is evaluated on the basis of the EER, the \hat{C}_{ur} and its minimal possible value \hat{C}_{ur}^{min} .

4.4.2.2 Results and Discussion

The verification performance for each combination described in Section 4.4.2 is quantified by the EER , \hat{C}_{ur} and \hat{C}_{ur}^{min} over the Dutch and Chinese Testing Sets. The meta-parameters for the RF classifier and the Legendre polynomials are set to the optimised values already used in the experiments presented in Subsection 4.4.1. For the case of the approximation based on wavelets, the time functions were resampled resulting in a normalised length of 256. The resolution level was set to 3, in order to obtain a feature vector of a reasonable length.

The verification error rates for the different feature combinations listed in Subsection 4.4.2 and the different time function approximations, are shown in Tables 4.6 and 4.7, for the Dutch and Chinese data, respectively. The best results obtained when using each of the proposed time function approximations are highlighted in **boldfaced** style. For the purposes of comparison, the verification results for the best commercial and non-commercial systems in the SigComp2011 competition [LMd⁺11] are also included in the last two rows of Tables 4.6 and 4.7.

From Tables 4.6 and 4.7 it can be seen that the combination using all the features and the combination using the automatically selected features obtain the best results, for both datasets. This shows that the feature selection done by the RF algorithm is a meaningful one. In addition, the length of the resulting feature vector is smaller than the corresponding to the set including all the features. Table 4.8 summarizes the sets of automatically selected features for the three different approximation techniques (the one using Legendre polynomials and the other two using **db4** and **bior6.8** wavelets) and for the Dutch (top section of the table) and Chinese (bottom section) datasets. Note that a different number of features are selected for each dataset, namely, 13 and 17 features for the Dutch and Chinese datasets, respectively. Again, more features are needed to model Chinese signatures suggesting their complexity.

Regarding the pen pressure, the results obtained in these experiments show

Table 4.6: Verification results for the Dutch Dataset

| Dutch Dataset | | | | | | | | | |
|------------------------|----------------------|-----------------|-----------------------|-----------------|-----------------|-----------------------|-----------------------|-----------------|-----------------------|
| Features | Legendre Polynomials | | | db4 wavelets | | | Bio6.8 wavelets | | |
| | EER | \hat{C}_{U_r} | $\hat{C}_{U_r}^{min}$ | EER | \hat{C}_{U_r} | $\hat{C}_{U_r}^{min}$ | EER | \hat{C}_{U_r} | $\hat{C}_{U_r}^{min}$ |
| ASF | 5.01 | 0.1903 | 0.1594 | 6.58 | 0.2426 | 0.2049 | 5.68 | 0.2295 | 0.1971 |
| [FORG07] | 5.18 | 0.2023 | 0.1746 | 6.87 | 0.2466 | 0.2144 | 6.06 | 0.2553 | 0.2067 |
| $xypv_T\theta a_T\rho$ | 6.4 | 0.2402 | 0.2077 | 8.04 | 0.2984 | 0.2485 | 7.24 | 0.2714 | 0.2309 |
| All feat. | 5 | 0.1879 | 0.1537 | 6.78 | 0.2491 | 0.2055 | 6.16 | 0.2320 | 0.1947 |
| System | Acc. | | | \hat{C}_{U_r} | | | $\hat{C}_{U_r}^{min}$ | | |
| commercial | 96.27 | | | 0.2589 | | | 0.1226 | | |
| 1st. non-commercial | 93.49 | | | 0.4928 | | | 0.2375 | | |

Table 4.7: Verification results for the Chinese Dataset

| Chinese Dataset | | | | | | | | | |
|-----------------------------|----------------------|-----------------|-----------------------|-----------------|-----------------|-----------------------|-----------------------|-----------------|-----------------------|
| Features | Legendre Polynomials | | | db4 wavelets | | | Bio6.8 wavelets | | |
| | EER | \hat{C}_{U_r} | $\hat{C}_{U_r}^{min}$ | EER | \hat{C}_{U_r} | $\hat{C}_{U_r}^{min}$ | EER | \hat{C}_{U_r} | $\hat{C}_{U_r}^{min}$ |
| ASF | 7.95 | 0.3313 | 0.2651 | 7.455 | 0.2962 | 0.2483 | 6.9 | 0.2874 | 0.2443 |
| [FORG07] | 9.38 | 0.3825 | 0.2764 | 8.41 | 0.3338 | 0.2621 | 8.12 | 0.3271 | 0.2619 |
| $x,y,p,v_T,\theta,a_T,\rho$ | 9.23 | 0.3539 | 0.3030 | 9.08 | 0.3548 | 0.2945 | 9.91 | 0.3727 | 0.3086 |
| All feat. | 8.66 | 0.3175 | 0.2647 | 7.9 | 0.3126 | 0.2476 | 7.26 | 0.2974 | 0.2421 |
| System | Acc. | | | \hat{C}_{U_r} | | | $\hat{C}_{U_r}^{min}$ | | |
| commercial | 93.17 | | | 0.4134 | | | 0.2179 | | |
| 1st. non-commercial | 84.81 | | | 0.5651 | | | 0.3511 | | |

that not only the pen pressure is always among the best features but also its first order time derivative is, since the automatic feature selection always include them into the optimal sets of features. This, for the Dutch and Chinese datasets. Moreover, for the latter, the second order time derivative of the pen pressure is also included in the selected feature sets. The best set of features for any of the feature extraction approaches being used and any of the datasets being considered contains the pen pressure and its first order time derivative. These results show that the pen pressure is useful independently of the feature extraction being used and, again, not highly influenced by the cultural origin of the signatures being considered. This is good news, since it provides support to the suggestions made in Subsection 4.4.1.2.

In [FORG07], the first order time derivatives are considered into the feature set since it has been shown that they are highly effective as discriminative parameters

Table 4.8: Optimal feature set selected by the automatic feature selection for each of the feature extraction approaches for the Dutch (top) and Chinese (bottom) datasets.

| Dutch | |
|-----------------------------|--------------------------------------------------------------------------------|
| Feature extraction approach | Optimal feature set |
| Legendre polynomials | $p, dx, dp, x, d^2\rho, a_T, y, d\rho, d^2y, d^2x, v_T, d^2\theta, dy, \rho$ |
| Wavelets db4 | $x, a_T, y, v_T, p, dp, \rho, dx, \theta, dy, d^2x, d^2y, dv_T$ |
| Wavelets Bio6.8 | $a_T, x, v_T, y, p, dy, dx, dp, \theta, d^2x, \rho, d^2y$ |
| Chinese | |
| Feature extraction approach | Optimal feature set |
| Legendre polynomials | $dx, d^2x, dp, p, y, d^2p, dy, dv_T, d\theta$ |
| Wavelets db4 | $yxpv_Ta_Tdydx^2y\theta\rho dpd^2xd\theta^2pdv_Tdpd^2\theta$ |
| Wavelets Bio6.8 | $x, y, v_T, p, a_T, dy, \rho, dx, dp, \theta, d^2x, d^2y, d\theta, d^2p, dv_T$ |

regarding verification with other behavioural traits, such as speech. In [RKD05], not only the first but the second order time derivatives of an original set of features are analysed. The results obtained when only the original set of features is used are clearly outperformed by the ones obtained when the features proposed in [FORG07] are used, that is, incorporating the first order time derivatives to the original set. This observation shows that the set of first order time derivatives is useful to reach better verification results. Nevertheless, to incorporate the second order time derivative seems not to be that significantly helpful, since incorporating them to the features used in [FORG07] does not improve the verification results as much as incorporating the first order time derivatives to the original feature set. Then, to use the first order time derivatives contributes to reach a better model of the signatures but to use the second order time derivatives does not contribute so much. This contribution of the first order time derivatives could be related to the power of the time derivatives to model the details. Nevertheless, it is important to note that, in addition to highlighting the details, the derivatives can also introduced spurious information like noise and produce other signal artifacts. Then, taking into account the first order time derivative of a signal may contribute to improve performance, but the second order time derivative may introduce non useful information. This is particularly noticeable in the cases of the time derivatives of the extended time functions, *viz*, v_t , θ , a_T and ρ that are more complex signals than the measured ones (x , y and p).

From Tables 4.6 and 4.7 it can be seen that the results obtained when using the automatically selected feature sets outperform the ones obtained when using the set of features used in [FORG07]. In line with the observations done in the previous paragraph, it can be observed that the main difference between these feature combinations is the way in which they incorporate the time derivatives

to the original set of features. The first order time derivatives of the measured functions, that is, dx , dy and dp are included in the optimal feature sets. On the other hand, dv_T and $d\theta$ are not always included, dp is only included once and da_T is not included in any optimal feature set. Instead of including these first order time derivative features, the second order time derivatives of the most simple time functions, that is, d^2x , d^2y and d^2p are included. This shows that it is likely more useful to use the second order time derivative of simple features than the first order time derivatives of more complex ones. As mentioned above, the time derivatives are capable to highlight details but also useless information. If the features are not so representative or they are unstable, the first order time derivatives of these features will be nothing but noisy data. On the other hand, the second order time derivatives of the simple features will still keep useful information and will be capable to highlight the details without increasing drastically the artifacts of the original signal.

From Table 4.6, it can be observed that the best results for the Dutch data are reached when using the Legendre polynomials approximation of the time functions. Dutch signatures are likely to be written in a continuous way, so that, the time functions associated with these signatures should be smooth and so polynomials, Legendre ones in this case, can accurately approximate them. In the case of using the approximation based on DWT, the results are not as good as in the case of using the Legendre approximation.

From Table 4.7, it can be observed that the best results for the Chinese data are obtained when using the DWT approximations of the time functions. As already mentioned above, Chinese signature style is, in most of the cases, close to the Chinese handwriting style, consisting of one or more multi-trace ideograms, and Chinese characters usually convey their meaning through pictorial resemblance to a physical object. This causes the time functions associated with the Chinese signatures to have several discontinuities. Then, a polynomial approximation, such as the one based on Legendre polynomials used here, is not good enough to model this type of signals.

Finally, the verification results obtained for the Dutch signatures are better than those for the Chinese ones, confirming the observations in Subsection 4.4.1 regarding the complexity of Chinese signatures.

4.5 Consistency Measure

In Subsections 4.4.1 and 4.4.2 different feature combinations have been analysed and their discriminative power has been evaluated on the basis of the obtained verification performance. It would be desirable to evaluate the discriminative capability of the features in a previous stage, that is, without the need of computing the verification errors. If such evaluation would be possible, it would be useful to predict the effectiveness of the features for verification purposes. Some works,

such as [LBA96] and [LG05], have been presented in the literature addressing this problem. In [LBA96], a consistency model of the features is proposed to quantify their discriminative power. Based on this model, an optimal subset of global features out of a larger global feature set is selected. In [LG05], several local and global features are compared on the basis of their consistency, resulting pen coordinates and some features derived from the pen coordinates the most reliable ones.

A well defined consistency model would allow to quantify the discriminative power of the features and to predict their effectiveness for verification purposes. Unfortunately, no standard and widely accepted consistency model exists in the literature. In this Section, a new consistency model is proposed.

4.5.1 Definition

An important property of a feature is its discriminative capability. Features associated with genuine signatures should be close to each other while distances between features associated with genuine and forged signatures should be relatively large. This property is usually called *consistency* of the feature.

A measure of consistency based on the features would be difficult to compute since they may have different lengths. It is then more reasonable to define a consistency measure based on the distances among features and not on the features themselves. The consistency of a given feature will then be computed based on the statistics of the intraclass (for the genuine signature class) and interclass (between the genuine and forged signature classes) distances. A consistency factor d , for each signer, could then be defined as follows:

$$d = \frac{\mu_D(C_g, C_f) - \mu_D(C_g, C_g)}{\sqrt{\sigma_D^2(C_g, C_g)} + \sqrt{\sigma_D^2(C_g, C_f)}}, \quad (4.14)$$

where C_g and C_f stand for the genuine and the forged classes, respectively, and where $\mu_D(C_g, C_g)$ and $\sigma_D^2(C_g, C_g)$, and $\mu_D(C_g, C_f)$ and $\sigma_D^2(C_g, C_f)$ are the sample means and sample variances of the genuine intraclass distances and the genuine-forged interclass distances, respectively.

The consistency factor in (4.14) is normalised in such a way that, under the assumption of Gaussian distributions for the involved distances, it equals 1 when the means $\mu_D(C_g, C_f)$ and $\mu_D(C_g, C_g)$ are separated by the sum of the respective standard deviations. This is illustrated in Fig. 4.8. The larger the consistency factor, the more consistent the features are.

An alternative definition of consistency, based on the Fisher ratio [Bis06]

$$J = \frac{(\mu_D(C_g, C_f) - \mu_D(C_g, C_g))^2}{\sigma_D^2(C_g, C_g) + \sigma_D^2(C_g, C_f)}, \quad (4.15)$$

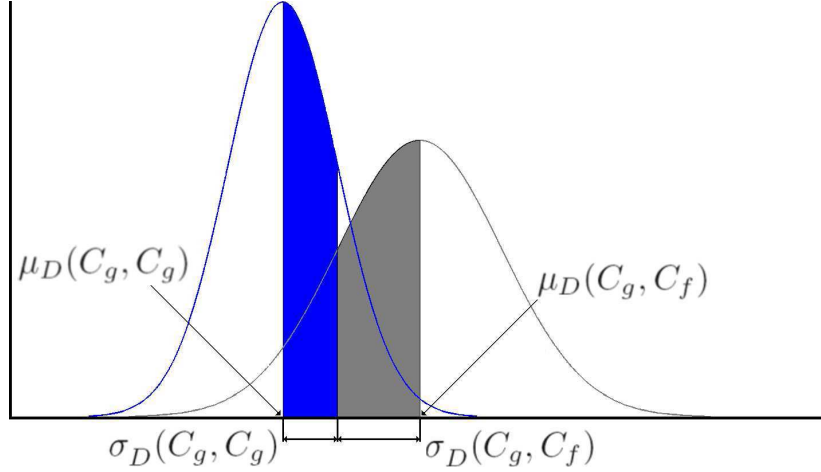


Figure 4.8: Genuine intraclass distance distribution (blue) and genuine-forged interclass distance distribution (grey), for the case $d = 1$.

would be the following:

$$\tilde{d} = \sqrt{J} = \frac{\mu_D(C_g, C_f) - \mu_D(C_g, C_g)}{\sqrt{\sigma_D^2(C_g, C_g) + \sigma_D^2(C_g, C_f)}}. \quad (4.16)$$

Here, the interpretation of the normalisation in Fig. 4.8 is not possible any more. The definition in (4.16) has been employed in [LG05] and [RKD05].

4.5.2 Evaluation Protocol

The proposed consistency factor defined in (4.14) quantifies the discriminative power of a particular combination of time functions. Then, the idea here is to compute the consistency factor value of different time function combinations and evaluate their discriminative power based on this value. It is important to highlight that, in addition to compute the consistency factor of the features and quantify their discriminative power, it is also necessary to evaluate the reliability of the proposed consistency measure. It is reasonable to expect that a simple set of features will be better suited than a more complex one to test the consistency factor reliability. Then, to evaluate the consistency factor capability to measure the discriminative power of the features, the set of features used in the experiments presented in Subsection 4.4.1, referred to as the original set of features in Subsection 4.4.2 will be used here. The consistency factor introduced in (4.14) is then used to quantify the discriminative capability of the different feature combinations listed in Table 4.3.

It would be desirable that, based on the consistency factor value, it would be possible to select the most suitable combination of time functions to be used for

a verification system. This selection has to be done in the training stage. The consistency factor should then be computed with the signatures available during this stage. It is the common case that, when training a verification system, skilled forgeries are not available. For this reason, the consistency factor for a particular writer will be computed using the genuine signatures corresponding to all the remaining writers as fictitious signatures to train the forged class. This will result in larger consistency factors compared with the ones that would be obtained using skilled forgeries for each writer. In any case, since the database does contain skilled forgeries, the consistency factor will also be computed using them. In this way, a comparison between the consistency measure computed resorting only to genuine signatures (interpreted as fictitious signatures for training the forged class) and resorting to skilled forgeries can be performed. A good correlation between these two quantities would be indicating the reliability of the proposed consistency measure.

Finally, the correlation between the consistency factor and the corresponding verification error for each feature combination is also analysed. A high correlation would be indicating that the consistency factor could be used to single out the best feature combination, in terms of the verification error. This, without the need of explicitly computing the error. In this way, the consistency factor value could be useful to perform feature selection. In addition, this correlation would be another indicator of the reliability of the proposed consistency measure.

As already said in Section 4.3, the datasets in the SigComp2011 Database are divided into two sets, namely, the Training Set and the Testing Set. The consistency factor will then be calculated as mentioned above over the Dutch and Chinese Training Sets. The correlation between the consistency factor and the verification errors, and the correlation between the consistency factor computed resorting only to genuine signatures and resorting to skilled forgeries, will be computed over the Dutch and Chinese Training Sets as well.

4.5.3 Results and Discussion

The proposed consistency factor d (in (4.14)) was computed for each of the combinations listed in Table 4.3, using only the genuine signatures, over the 10 authors in the Dutch Training Set and over the 10 authors in the Chinese Training Set. Figures 4.9 and 4.10 show the boxplots associated with the consistency factors for each feature combination over the 10 writers in the Training Set (left), and a detail of the boxplots associated with the two best (larger consistency factors) and two worst (smaller consistency factors) combinations (right), for the Dutch and Chinese datasets, respectively. The different feature combinations have been included in a nonincreasing order of the consistency factor, from left (most consistent) to right (least consistent).

Figure 4.9 shows that the most consistent combinations for the Dutch data are the one containing the pressure, the total acceleration and the log curvature radius

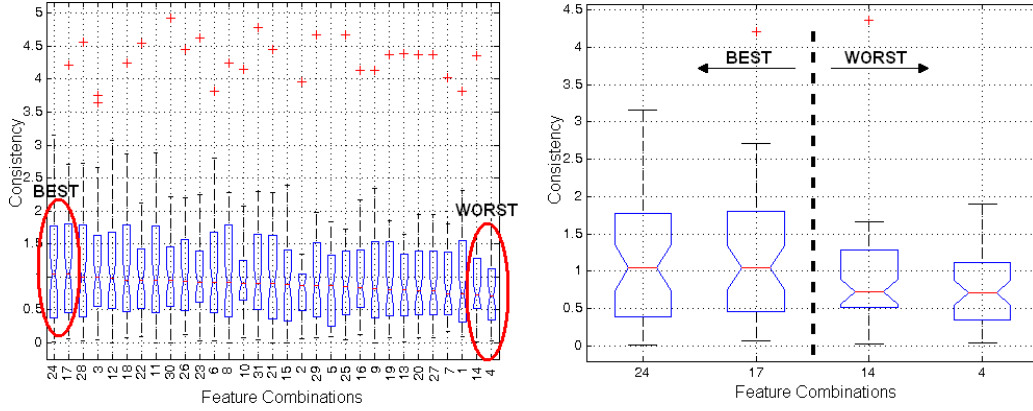


Figure 4.9: Boxplots for the consistency factor over the 10 authors in the Training Set for the Dutch data (left), and detail of the two most and the two least consistent feature combinations (right).

($pa_T\rho$), and the one containing the pen coordinates, the pressure and the total acceleration ($xypa_T$). The overlapping notches in the boxplots would indicate that the difference between the medians is not statistically significant. This is not the case if one compares the best (left most boxplot) and the worst (right most boxplot) combinations, where no overlapping is present, indicating that the difference between the medians is statistically significant. Similar comments, *mutatis mutandi*, hold for the boxplots in Fig. 4.10 corresponding to the Chinese data. For this dataset, the most consistent combinations are the one containing the pen coordinates and the pressure (xyp), and the one containing the pen coordinates, the pressure and the velocity ($xypv_T\theta$).

The results obtained in this Section, show that whenever the pen pressure is incorporated to a feature combination, the consistency factor is improved in almost all the cases, for both datasets. This indicates that the pen pressure, when combined with the other time functions, is a reliable feature. These observations agree with the ones made in the previous Subsections 4.4.1.2 and 4.4.2.2 and many other reported results in the literature (see for example [MM07] and [RKD05]), where it is stated that the pen pressure is a useful feature to distinguish between writers when used in combination with other time functions. Of course, the pen pressure is not powerful when it is employed as the unique feature. This result agrees with the one presented in [LG05] where the authors used only one feature at a time to compute the consistency. However, using only one feature at a time is not a very realistic case. In addition, the most consistent combinations (shown in Fig. 4.9 and 4.10), for both datasets, contain the pen pressure, suggesting, in line with the previous discussions in Subsections 4.4.1.2 and 4.4.2.2, that the reliability of the pen pressure is not highly influenced by the considered cultural origin of the signatures.

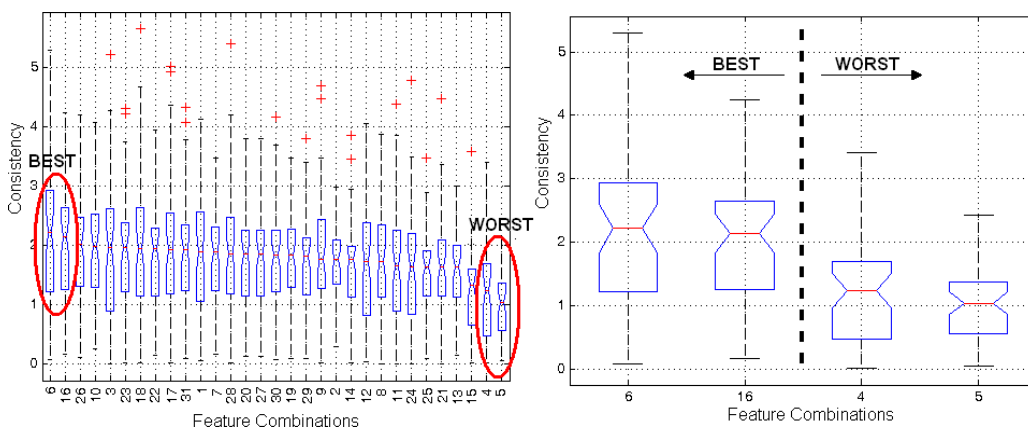


Figure 4.10: Boxplots for the consistency factor over the 10 authors in the Training Set for the Chinese data (left), and detail of the two most and the two least consistent feature combinations (right).

For the Dutch data, in addition to the pen pressure, the acceleration is also present in the most consistent combinations. Moreover, as it is the case of the pen pressure, when the acceleration is incorporated to a particular feature combination it improves the consistency factor in most of the cases. In addition, these two features are also the ones that improve the obtained verification error rate of a feature combination whenever they are incorporated (see Subsection 4.4.1.2). This fact is important because it means that whenever an improvement in the consistency factor is made by incorporating one of these features to a feature combination it can be expected an improvement in the verification performance of that combination. Then, in these cases, there would be no need to explicitly compute the error rate in order to evaluate its improvement, because it would be predicted by the improvement in the consistency factor. Again, the acceleration by itself is not a highly consistent feature, in agreement with [LG05]. For the Chinese data, the x and y pen coordinates are present in the most consistent combinations, and they improve the consistency in most of the cases when incorporated to a particular feature combination. This is in line with the analysis in Subsection 4.4.1.2, where it was argued that the position information (x and y pen coordinates) is likely to have more discriminative power for this type of data, improving the verification error rates whenever is included in a feature combination.

The position information is likely to be more consistent for the Chinese data than for the Dutch data, as already discussed in Subsection 4.4.1.2. On the other hand, for the Dutch data, the information about the changes in the velocity of writing (the acceleration) is likely to be more consistent. Dutch signatures are, in general, irregularly shaped, then it is likely that the position information is less important than the information regarding the rate of change of the position. Even

further, the changes in the velocity seem to be more important than the velocity itself. The acceleration points out these changes, revealing typical hesitations of the forgers. This may not be the case for the Chinese signatures because as they are signatures containing separated characters, the velocity present lots of changes and zero velocity moments in the genuine as well as in the forged signatures. Then, a hesitation of the forger is difficult to distinguish from a stop in the natural writing process of a genuine writer.

In order to assess the ability of the proposed consistency factor d to predict the verification performance of the system (in the sense of the classifiers error rates), the correlation between the consistency factor d and the verification error \hat{C}_{ur}^{min} is calculated for the feature combinations listed in Table 4.3. Spearman's correlation coefficient [GC03] is used to quantify this correlation. The verification performance presented in Subsection 4.4.1.2 corresponding to the experiments described in Subsection 4.4.1, is the verification performance for each of the combinations considered for the consistency computation and so, the one taken into account in the present correlation analysis. The results for the correlation between the proposed consistency factor d and the verification error \hat{C}_{ur}^{min} (showed in Figs. 4.6 and 4.7) are shown in Figs. 4.11 and 4.12, for the Dutch and Chinese data, respectively, where the feature combinations have been included in a nondecreasing order of the correlation coefficient, from left (highest absolute correlation) to right (lowest absolute correlation). An alternative measure of correlation would be the Pearson correlation coefficient [GC03] which is design to measure the strength of the linear dependence between two variables. While the Pearson correlation coefficient is limited to linear dependence, Spearman's correlation coefficient is more general, in the sense that it measures the monotonic dependence between two variables (not restricted to a linear function), being the best option for the current analysis.

Figure 4.11 shows that, for the Dutch data, the correlation between the proposed consistency factor d and the \hat{C}_{ur}^{min} , and the correlation between the consistency factor \tilde{d} (in (4.16)) and the \hat{C}_{ur}^{min} , are very similar. On the other hand, Fig. 4.12 shows that, for the Chinese data, the correlation between the consistency factor and the \hat{C}_{ur}^{min} is slightly better in the case of using the consistency factor \tilde{d} (in (4.16)), specially when using the SVM classifier. Nevertheless, the best correlations obtained by d (-0.6003 and -0.6457 for the RF and SVM classifiers, respectively) in the case of the Dutch data and the ones in the case of the Chinese data (-0.6911 and -0.6347 for the RF and SVM classifiers, respectively) are indicating an acceptable correlation between the consistency factor and the verification error. That is, d could be used as an indicative of the verification performance to be expected for a particular feature combination (and a particular classifier), for both datasets. From Figs. 4.11 and 4.12 it can be seen that the correlation between the consistency factor and the verification error is highly dependent on the classifier being used. This correlation is an effective and maybe

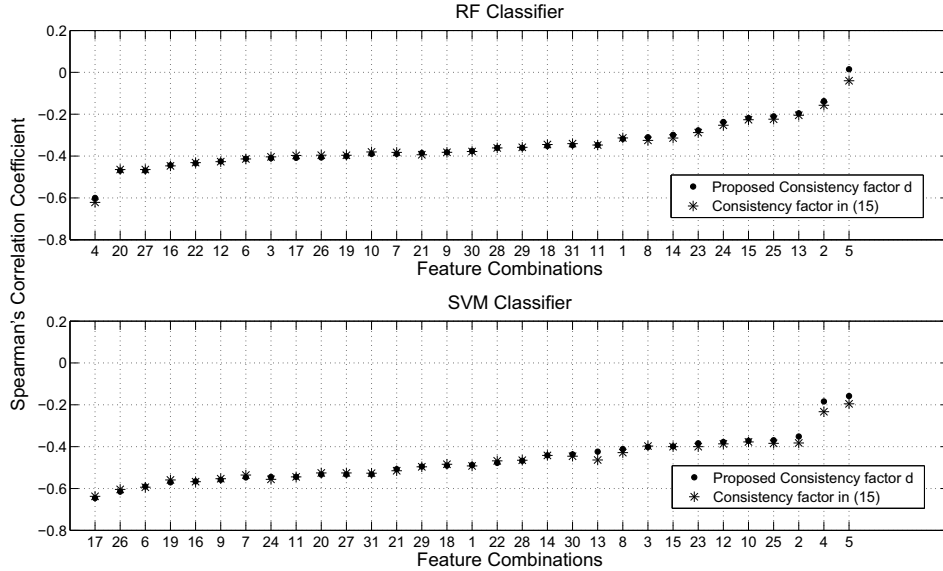


Figure 4.11: Spearman's correlation coefficient between the Consistency factors (d and \tilde{d}) and the \hat{C}_{llr}^{min} calculated for the signatures in the Dutch Training Set using the RF classifier (top) and the SVM classifier (bottom).

more accurate indicative of the verification performance than the consistency factor itself. Nevertheless, it is important to highlight that if the best feature combination is to be chosen based either on the consistency or the correlation, the consistency factor has the advantage of not being dependent on the classifier being used.

While the decision about which feature combination is to be used must be necessarily made based on consistency factors computed with genuine signatures, in real situations the verification system is likely to be subjected to skilled forgeries. For the consistency factor, computed during the training stage, to be reliable, it must have a high correlation with the consistency factor computed resorting to skilled forgeries. Since the Dutch and Chinese Training Sets do contain skilled forgeries, this correlation can be computed. As in the case of the correlation between the consistency and the verification error rates, also here the Spearman's correlation coefficient is employed to quantify this. Figures 4.13 and 4.14 show the boxplots for Spearman's correlation coefficient over all the feature combinations listed in Table 4.3, for d and \tilde{d} , for the Dutch and Chinese Training Sets, respectively. From Figs. 4.13 and 4.14 it can be observed that the correlation values using the consistency factor d are better than the ones using \tilde{d} . This is an important advantage of the proposed consistency factor d . Despite the fact that the correlation for the Dutch data (Fig. 4.13) is higher than for the Chinese data

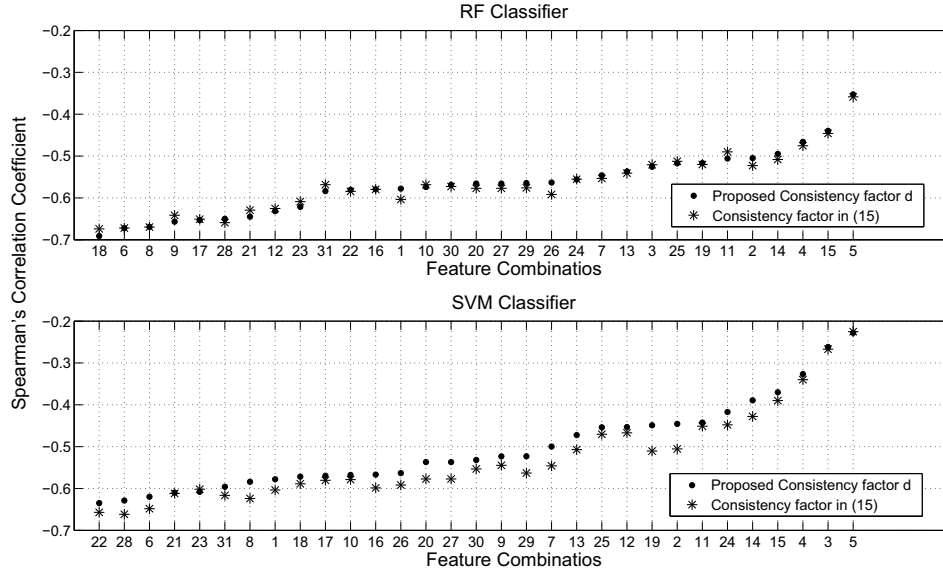


Figure 4.12: Spearman's correlation coefficient between the Consistency factors (d and \tilde{d}) and the \hat{C}_{llr}^{min} calculated for the signatures in the Chinese Training Set using the RF classifier (top) and the SVM classifier (bottom).

(Fig. 4.14), both correlation coefficients are high enough to allow the use of the consistency factor computed using only genuine signatures in a real situation in which skilled forgeries are not available to train the system but they are present to test the system. Of course, this is strongly dependent on the forgeries quality.

For the purposes of summarizing, Tables 4.9 and 4.10 show the verification errors, quantified by the EER , \hat{C}_{llr} and \hat{C}_{llr}^{min} (previously shown in Figs. 4.6 and 4.7), for the three most consistent feature combinations (top left section of the tables), the three ones that obtained the highest correlation between the consistency factor and the verification error \hat{C}_{llr}^{min} using RF (bottom left section) and using SVM (bottom right section), and the three ones that obtained the highest correlation between the consistency factor computed using only genuine signatures and using skilled forgeries (top right section), over the Dutch and Chinese Testing Sets, respectively. The best results in Tables 4.9 and 4.10 using RF and SVM classifiers, are highlighted in **boldfaced** style.

From Tables 4.4 and 4.9, it can be observed that using consistency as the criterion for feature selection leads to the feature combination with the best error rate which corresponds, in this case, to the SVM classifier (shaded cells in Table 4.9). Similarly, from Tables 4.4 and 4.10, it can be observed that using the correlation between the consistency and the \hat{C}_{llr}^{min} for SVM as the feature selection criterion leads to the feature combination with the best verification performance

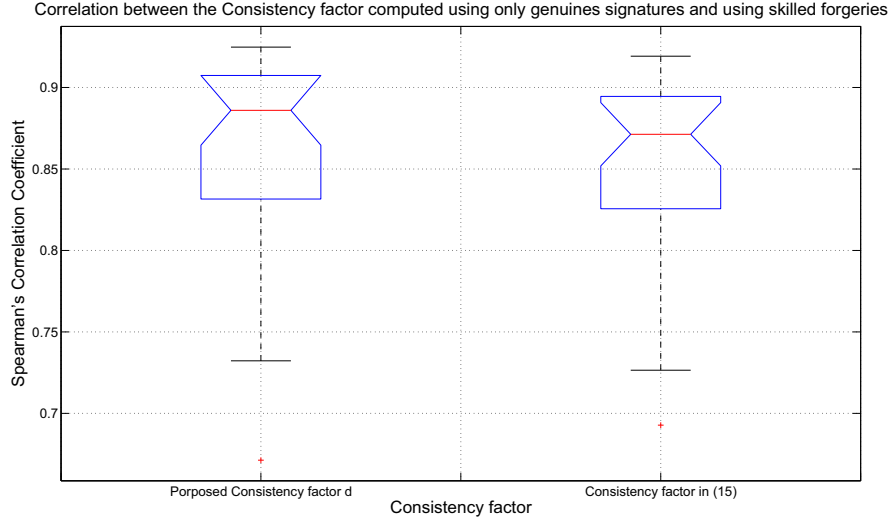


Figure 4.13: Boxplots for Spearman's correlation coefficient between d (left) and \tilde{d} (right) computed using only genuine signatures and using skilled forgeries from the Dutch Training Set, over all the combinations tested.

which corresponds, in this case, to the RF classifier (shaded cells in Table 4.10). Although the consistency and correlation measures in Tables 4.9 and 4.10 not always lead to the best combination, that is the one with the lowest error rate, they lead to combinations with low errors, and so they can be used as feature selection criteria.

4.6 Some Concluding Remarks

This Chapter was devoted to the analysis of different features for online signature verification. An in depth analysis of different combinations of the time functions was performed in an attempt to give some insight to the actual discriminative power of the features.

Two different approximation techniques to represent the time functions associated with the signing process were proposed, one based on Legendre polynomials and the other based on wavelet decomposition. Both of them proved to be well suited for modeling the time functions since the obtained verification results are among the best ones reported in the state-of-the-art over the same datasets. In addition, the proposed signature models would allow for a dimensionality reduction with respect to the case of using all the points in the time functions. The optimal order of the Legendre polynomials is 21, while the time functions length was normalised to 256 before wavelet decomposition, resulting in representations

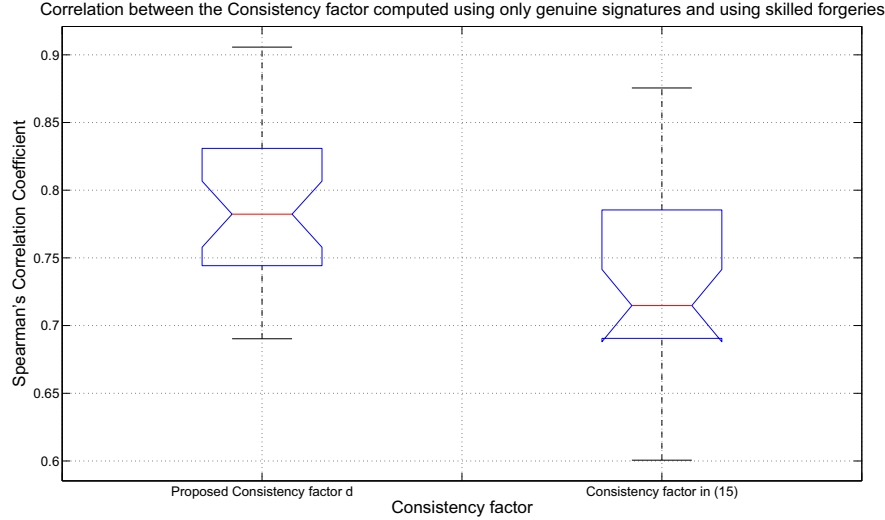


Figure 4.14: Boxplots for Spearman’s correlation coefficient between d (left) and \tilde{d} (right) computed using only genuine signatures and using skilled forgeries from the Chinese Training Set, over all the combinations tested.

with 34 and 48 parameters for the **db4** and **bior6.8** wavelets, respectively. Considering, for instance, a signature with about 1500 points, the dimensionality reduction would be in the order of $1500/22 \simeq 68$ in the case of using Legendre polynomials, $1500/34 \simeq 44$ and $1500/48 \simeq 31$ in the case of using **db4** and **bior6.8** wavelets, respectively.

An exhaustive study was carried out over a different sets of features widely used in the online signature verification literature. The discriminative power of the pen pressure has particularly been analysed, and it was shown to be a useful feature. Moreover, when analysing the first and second order time derivatives, the derivatives of the pen pressure proved to be useful too.

Although in this work only Dutch and Chinese signatures are considered, the fact that the pen pressure discriminative power is not highly influenced by the cultural origin of the data is very important since it means that the pen pressure could be a useful feature for any signature style. Of course, it is mandatory to analyse more data from different cultures in order to make further conclusions, but the results presented here can be a promising starting point.

A new consistency factor was defined to quantify the discriminative power of the features. The capability of the proposed consistency factor to quantify the discriminative power of the features and to predict their behaviour concerning to the verification performance, was evaluated on the basis of a correlation analysis. A good correlation between the consistency factor computed using only genuine signatures and using skilled forgeries was shown. This robustness property is

Table 4.9: Verification results over the Dutch Testing Set for the best, regarding consistency and correlation, feature combinations.

| Most Consistent | | | | | Highest corr.: cons. gen. and skilled for. | | | | |
|-----------------------------------------------------|------------|--------------|-----------------|-----------------------|------------------------------------------------------|------------|--------------|-----------------|-----------------------|
| Comb. | Class. | EER | \hat{C}_{llr} | \hat{C}_{llr}^{min} | Comb. | Class. | EER | \hat{C}_{llr} | \hat{C}_{llr}^{min} |
| $pa_T\rho$ | RF | 8.75 | 0.3379 | 0.2695 | $xya_T\rho$ | RF | 7.65 | 0.3109 | 0.2525 |
| | SVM | 12.78 | 0.5154 | 0.3889 | | SVM | 14.25 | 0.5563 | 0.4460 |
| $xypa_T$ | RF | 6.65 | 0.2531 | 0.2122 | $xypa_T$ | RF | 6.65 | 0.2531 | 0.2122 |
| | SVM | 12.04 | 0.5056 | 0.3901 | | SVM | 12.04 | 0.5056 | 0.3901 |
| $xypa_T\rho$ | RF | 6.91 | 0.2478 | 0.2131 | $pvt\theta a_T\rho$ | RF | 5.84 | 0.2208 | 0.1862 |
| | SVM | 10.68 | 0.4368 | 0.3323 | | SVM | 10.94 | 0.4782 | 0.3644 |
| Highest corr.: cons. and \hat{C}_{llr}^{min} (RF) | | | | | Highest corr.: cons. and \hat{C}_{llr}^{min} (SVM) | | | | |
| Comb. | Class. | EER | \hat{C}_{llr} | \hat{C}_{llr}^{min} | Comb. | Class. | EER | \hat{C}_{llr} | \hat{C}_{llr}^{min} |
| a_T | RF | 11.06 | 0.4336 | 0.3450 | $xypa_T$ | RF | 6.65 | 0.2531 | 0.2122 |
| | SVM | 31.95 | 1.000 | 0.7665 | | SVM | 12.04 | 0.5056 | 0.3901 |
| $xypv_T\rho$ | RF | 7.29 | 0.2885 | 0.2307 | $xypv_T\theta a_T$ | RF | 5.44 | 0.2247 | 0.1791 |
| | SVM | 15.31 | 0.5857 | 0.4823 | | SVM | 11.63 | 0.4567 | 0.3758 |
| $xypv_T\theta\rho$ | RF | 6.20 | 0.2347 | 0.1957 | xyp | RF | 6.96 | 0.2869 | 0.2303 |
| | SVM | 11.88 | 0.5176 | 0.3854 | | SVM | 13.92 | 0.5577 | 0.4294 |

important since, in real applications, skilled forgeries are not available in the training phase but they do appear when testing the system. In addition, a good correlation between the consistency values and the verification errors was also shown, suggesting that the former could be used to select the optimal (i.e., leading to the smallest verification error) feature combination.

Finally, in order to give an insight of the influence of the cultural origin of the signatures in the verification performance of the systems, Dutch and Chinese signatures have been analysed. The presented results, showed that Chinese signatures are more complex and lot of work has still to be done on this type of signatures.

4.7 Some Ideas for Further Work

To actually use the proposed consistency factor to perform feature selection would be a good opportunity to confirm the results shown in this Chapter.

The analysis of signatures from different cultural origins gives an interesting starting point to further studies. It would be also interesting to continue in this direction incorporating to the Dutch and Chinese signatures more signatures from other countries.

It is important to note that the discriminative power of the pen pressure can easily be influenced by the signature acquisition process. If this process is done

Table 4.10: Verification results over the Chinese Testing Set for the best, regarding consistency and correlation, feature combinations.

| Most Consistent | | | | | Highest corr.: cons. gen. and skilled for. | | | | |
|-----------------------------------------------------|------------|--------------|-----------------|-----------------------|------------------------------------------------------|------------|--------------|-----------------|-----------------------|
| Comb. | Class. | EER | \hat{C}_{llr} | \hat{C}_{llr}^{min} | Comb. | Class. | EER | \hat{C}_{llr} | \hat{C}_{llr}^{min} |
| xyp | RF | 10.93 | 0.3998 | 0.3239 | $v_T\theta$ | RF | 11.50 | 0.4336 | 0.3505 |
| | SVM | 13.54 | 0.5320 | 0.4523 | | SVM | 16.06 | 0.5583 | 0.4612 |
| $xypv_T\theta$ | RF | 10.27 | 0.3948 | 0.3116 | $pv_T\theta\rho$ | RF | 11.17 | 0.4238 | 0.3161 |
| | SVM | 12.91 | 0.4673 | 0.3965 | | SVM | 12.63 | 0.4765 | 0.3884 |
| $xypv_T\theta a_T$ | RF | 9.85 | 0.3786 | 0.3112 | pv_xv_y | RF | 10.07 | 0.3727 | 0.2955 |
| | SVM | 12.41 | 0.4547 | 0.3752 | | SVM | 12.36 | 0.4654 | 0.3842 |
| Highest corr.: cons. and \hat{C}_{llr}^{min} (RF) | | | | | Highest corr.: cons. and \hat{C}_{llr}^{min} (SVM) | | | | |
| Comb. | Class. | EER | \hat{C}_{llr} | \hat{C}_{llr}^{min} | Comb. | Class. | EER | \hat{C}_{llr} | \hat{C}_{llr}^{min} |
| $xyp\rho$ | RF | 11.01 | 0.3940 | 0.3262 | $pv_T\theta a_T$ | RF | 8.93 | 0.3620 | 0.2722 |
| | SVM | 13.11 | 0.5052 | 0.4202 | | SVM | 11.57 | 0.4377 | 0.3703 |
| xyp | RF | 10.93 | 0.3998 | 0.3239 | $xyp a_T\rho$ | RF | 9.02 | 0.3408 | 0.2739 |
| | SVM | 13.54 | 0.5320 | 0.4523 | | SVM | 12.39 | 0.4969 | 0.3906 |
| $xyp a_T$ | RF | 10.59 | 0.4058 | 0.3308 | xyp | RF | 10.93 | 0.3998 | 0.3239 |
| | SVM | 16.16 | 0.5505 | 0.4626 | | SVM | 13.54 | 0.5320 | 0.4523 |

in carefully controlled conditions, the pressure is likely to be a very distinctive feature. As a future work, it would be interesting to study its discriminative power in a FHEs casework context.

Online Signature Verification: Exploiting characteristics

In this Chapter, a pre-classification stage based on global features is incorporated to an online signature verification system based on the wavelet representation of the time functions presented in Section 4.2 and a RF classifier, for the purposes of improving its performance.

5.1 Introduction

There exist two types of features for online signature verification, namely, local (calculated for each point in the time sequence), and global (calculated from the whole signature). Many researchers accept that approaches based on local features achieve better performance than those based on global features, but still there are others who favour the use of global features [RKD05], [FNL⁺05]. Indeed, it is interesting to use global features since they have the advantage of being simple features, usually more intuitive than local ones, and can be easily computed and compared. Which features, local or global, are better suited to represent the signatures is not an easy question to answer. Further, it would be reasonable to expect that local and global features could provide complementary information [FNL⁺05]. Then, a combination of global and local features would be a good alternative. The way in which they should be combined is not trivial and constitutes an interesting and still open challenge.

The different representation levels (global and local) should be combined, according to the application needs, in such a way that can make one or another to be more important. In [Pla94], a multilevel online signature verification system which uses three different signature representations, one based on global features and the other two based on local ones, is presented. In addition, several fusion strategies have been proposed in the literature to combine local and global features. For instance, in [FNL⁺05], two widely known decision level fusion strategies, namely, **sum** and **max** rules are used for this purpose and compared.

Finally, some approaches using a pre-classification stage based on global features, such as signature total time duration and pen down duration, for the purposes of early detecting bad forgeries, have been proposed in the mid 1990's in [Pla94] and [LBA96].

In this Chapter, global based features are used for pre-classification purposes. The idea is to pre-classify signatures, declaring as forgeries those that are far away from their mean, in terms of the global based features. This could help to quickly recognize and classify gross forgeries, speeding up and simplifying the verification process. The remaining signatures continue with the subsequent classification stage which consists in extracting features providing a more detailed representation (wavelet approximations of the associated time functions as described in Section 4.2), and classifying them on the basis of a RF classifier like the one used in the experiments presented in Chapter 4.

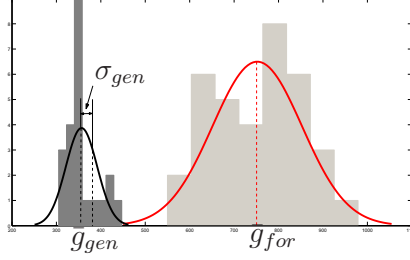
Two different approaches for pre-classification are proposed. On one hand, each individual global based feature is used separately and its discriminative capacity is analysed. On the other hand, the global based features are used in a combined form (unique feature vector) and their combined discriminative power is studied.

The Chapter is organized as follows. Section 5.2 introduces the proposed pre-classification approach. The feature extraction is described in Section 5.3. In particular, Subsection 5.3.1 focuses on the global based features, while Subsection 5.3.2 focuses on the time function based features. The evaluation protocol used to perform the experiments is described in Section 5.4. In Section 5.5 the results are presented and discussed. In particular, Subsection 5.5.1 addresses the univariate case and Subsection 5.5.2 addresses the multivariate case. Finally, some concluding remarks are given in Section 5.6 and in Section 5.7 some future directions are discussed.

5.2 Pre-classification

The idea of the pre-classification is to exploit the intrinsic characteristics of the features based on global parameters and the ones based on time functions for different tasks. When using global based features, it would be expectable to get a rough and quick signature representation that could be useful to detect some anomalies of the signature. On the other hand, if a more precise representation is needed, using the time function based features could provide it, at the cost of a more time consuming feature extraction.

Global based features are then used in this paper for pre-classification purposes. It is reasonable to expect that some global based features, such as signature total time duration, pen down duration and average pressure, for the genuine samples would be far away from the corresponding ones for the forged samples. This is illustrated in Fig. 5.1 (left), for the case of a single global based feature,



Algorithm 1

If $|g_{test} - \bar{g}_{train}|^2 > \alpha^2 \sigma_{train}^2$
 then signature=FORGERY
 else continue classification

Figure 5.1: Left: Distribution of the global based feature signature total time duration for the genuine and forged signatures of a particular author. Right: Decision rule.

where the distributions of the global based feature “signature total time duration” for the genuine (left) and forged (right) signatures of a particular author, are depicted.

The idea is then to classify as a forgery those signatures for which the global based features differ significantly from the corresponding genuine feature mean. In particular, the decision rule in Fig. 5.1 (right) is considered. There, g_{test} denotes the global based feature corresponding to the test signature, \bar{g}_{train} and σ_{train}^2 are the global based feature sample mean and sample variance over the genuine training set, respectively, and α is a coefficient defining the threshold.

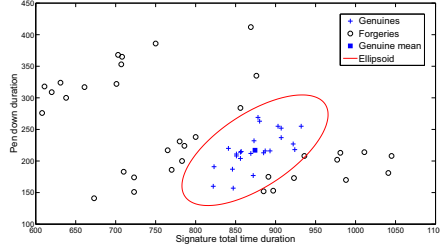
An alternative would be to consider a feature vector containing all the global based features. The natural extension to the multivariable case of the decision rule in Algorithm 1 would be Algorithm 2 in Fig. 5.2 (right). There, \mathbf{g}_{test} denotes the global based feature vector corresponding to the test signature, $\bar{\mathbf{g}}_{train}$ and Σ_{train} are the global based feature vector sample mean and sample covariance over the genuine training set, respectively, and α is a coefficient defining the threshold. The decision rule means that signatures whose feature vectors lie outside the hyperellipsoid define as

$$(\mathbf{g}_{test} - \bar{\mathbf{g}}_{train})^T \Sigma_{train}^{-1} (\mathbf{g}_{test} - \bar{\mathbf{g}}_{train}) = \alpha^2,$$

are considered as forgeries. Figure 5.2 (left) illustrates this for the case of a two-dimensional feature vector.

Coefficient α^2 is computed, for the global based feature vector, in three different ways for comparison purposes, namely, as: i. the maximum, or ii. the mean, or iii. the minimum, over all the authors in the Training Set of the database, of the maximum, over all the signatures of each author, of the positive definite quantity (note the reader that this is the so-called Mahalanobis distance):

$$d(\mathbf{g}_{test}, \bar{\mathbf{g}}_{train}) \triangleq (\mathbf{g}_{test} - \bar{\mathbf{g}}_{train})^T \Sigma_{train}^{-1} (\mathbf{g}_{test} - \bar{\mathbf{g}}_{train}). \quad (5.1)$$

**Algorithm 2**

If $(\mathbf{g}_{test} - \bar{\mathbf{g}}_{train})^T \Sigma_{train}^{-1} (\mathbf{g}_{test} - \bar{\mathbf{g}}_{train}) > \alpha^2$
 then signature=FORGERY
 else continue classification

Figure 5.2: Left: Distribution of the global feature vectors for the genuine and forged signatures of a particular author. In this case, the feature vector is composed by the signature total time duration and the pen down duration. Right: Decision rule.

That is (for case i.),

$$\alpha^2 = \max_A \max_{A_i} \{ (\mathbf{g}_{test} - \bar{\mathbf{g}}_{train})^T \Sigma_{train}^{-1} (\mathbf{g}_{test} - \bar{\mathbf{g}}_{train}) \}, \quad (5.2)$$

where A is the set of all the authors in the Training Set and A_i denotes the i -th author in the same set.

For the univariate case in Algorithm 1, equation (5.2) becomes:

$$\alpha^2 = \max_A \max_{A_i} \left\{ \frac{|g_{test} - \bar{g}_{train}|^2}{\sigma_{train}^2} \right\}.$$

A different approach is considered in [LBA96], where only the signature total time duration is used for pre-classification purposes. There, the threshold is computed as a fraction of the \bar{g}_{train} , and it is heuristically set to 0.2.

5.3 Feature Extraction

The set of time functions used in this Chapter is the same as the one used in Chapter 4 (presented in Subsection 4.4.2). That is, the measured time functions x , y , and p , the extended functions v_T , θ , a_T and ρ , and their first (dx , dy , dp , dv_T , $d\theta$, da_T and $d\rho$), and second (d^2x , d^2y , d^2p , d^2v_T , $d^2\theta$, d^2a_T and $d^2\rho$) order time derivatives. Previous to the feature extraction, the original pen coordinates are normalised regarding scale and translation as in Subsection 4.2.1.2.

5.3.1 Global based features

Several global based features can be extracted from the measured and extended time functions. These features should be selected to be discriminative enough in order for the proposed pre-classification to be successful. The following global based features, corresponding to the better ranked ones by the feature selection performed in [RKD05] and [FNL⁺05], are used in this paper:

- signature total time duration T ,
- pen down duration T_{pd} ,
- positive x velocity duration T_{vx} ,
- average pressure \bar{P} ,
- maximum pressure P_M ,
- time at which the pressure is maximum T_{P_M} .

5.3.2 Time function based features

The set of time functions used in this Chapter, *viz.*, x , y , p , v_T , θ , a_T , ρ , and their first (dx , dy , dp , dv_T , $d\theta$, da_T and $d\rho$), and second (d^2x , d^2y , d^2p , d^2v_T , $d^2\theta$, d^2a_T and $d^2\rho$) order time derivatives, are modeled using the DWT coefficients as described in Subsection 4.2.2.2. In particular, the **db4** wavelet is used in this Chapter.

5.4 Evaluation protocol

The SigComp2011 Dataset is used for the verification experiments. For each dataset, the optimisation of the meta-parameters of the system is performed over the corresponding Training Set while the Testing Set is used for independent testing purposes. The meta-parameters are: α for the pre-classification stage, and the ones corresponding to the wavelet decomposition approach and the RF classifier listed in Subsection 4.4.2.

To obtain statistically significant results, a 5-fold cross-validation (5-fold CV) is performed over the Testing Set to estimate the verification errors. For each instance of the 5-fold CV, a signature of a particular writer from one of the testing sets in the 5-fold CV is fed to the system. After its preprocessing, the global based features (\mathbf{g}_{test}) are extracted from it. Then, the pre-classification is performed as follows: the distance in (5.1) between \mathbf{g}_{test} and $\bar{\mathbf{g}}_{train}$ (sample mean computed over the current writer's genuine signatures available in the training set of the 5-fold CV) is computed. If this distance is larger than the threshold (α^2), the signature is declared to be a forgery. If this is not the case, the signature is subjected to the subsequent classification stage, as follows: the DWT approximation coefficients are computed for the different time functions being considered. Then, a RF classifier is trained by the current writer's genuine class in the training set of the 5-fold CV, and a forged class consisting of the genuine signatures of all the remaining writers in the same set. The result of the verification process is then either the result of the pre-classification (the input signature is considered a

forgery), or the result of the RF classifier. If the result is given by the pre-classification, the verification process is speeded up.

5.5 Results and Discussion

To evaluate the individual discriminative power of the global based features, experiments using each one of them separately for pre-classification purposes were carried out. These results are presented in Subsection 5.5.1. On the other hand, experiments using all the global based features in a combined feature vector for pre-classification were also carried out. The correlation between the individual global based features is taken into account through the global based feature vector covariance. This is important, since the features are not likely to be independent. The corresponding results are presented in Subsection 5.5.2. In both cases, the meta-parameters of the wavelet decomposition and the RF classifier, are set to their corresponding optimised values presented in Subsection 4.4.2.

5.5.1 Univariate case

The verification results, with pre-classification, for the six global based features, and the three different values of α , are shown in Table 5.1, for the Dutch (left) and Chinese (right) data. The best results are indicated in **boldfaced** style. For comparison purposes, also the results without pre-classification are included in the last row section of Table 5.1. The results in [PGL12], based on Legendre polynomials representations (without pre-classification), are also included in that section of the table.

From Table 5.1, it can be seen that the proposed pre-classification does improve the error rates with respect to the case of not using it, and also the ones in [PGL12]. This is not the case for every value of α . The actual values of α belong to the intervals: $[3, 4]$, $[2, 3]$ and $[1, 2]$, for the **max**, the **mean** and the **min** criteria, respectively. The maximum α (α_{max}) defines a conservative threshold, while α_{min} allows for more signatures to be pre-classified at the cost of larger errors (in the sense of classifying genuine signatures as forgeries). Then, α_{mean} is a trade-off value of α . The results confirm this, since using α_{max} leads, in most of the cases, to better results than using α_{mean} , while using α_{min} leads to the worst results. For the Dutch data, the error rates improve only when using α_{max} . This shows that many genuine signatures are beyond $2\sigma_{gen}$. For the Chinese data, the error rates are still improved when using α_{mean} , achieving the best error rate in this case. Then, for this data, the threshold can be chosen to be a more robust one.

The pre-classifications based on P_M , T_{P_M} and T_{pd} did not get any error rate improvements. In the cases of P_M and T_{P_M} , this seems to be reasonable. The time at which the pressure is maximum T_{P_M} is probably an unstable feature, since

Table 5.1: Verification results for the Dutch (left) and Chinese (right) Datasets

| | Dutch Dataset | | | | | | Chinese Dataset | | | | | |
|----------------------|-------------------|----------|--------------------|-----------|-------|-----------|-------------------|----------|--------------------|-----------|-------|-----------|
| | T | T_{pd} | T_{v_x} | \bar{P} | P_M | T_{P_M} | T | T_{pd} | T_{v_x} | \bar{P} | P_M | T_{P_M} |
| | α_{max} | | | | | | | | | | | |
| EER | 5.35 | 6.22 | 4.95 | 4.93 | 6.58 | 7.08 | 5.65 | 7.64 | 7.01 | 7.05 | 7.84 | 7.91 |
| \hat{C}_{Ur} | 0.237 | 0.238 | 0.228 | 0.212 | 0.254 | 0.270 | 0.224 | 0.301 | 0.264 | 0.263 | 0.291 | 0.284 |
| \hat{C}_{Ur}^{min} | 0.201 | 0.205 | 0.16 | 0.168 | 0.214 | 0.240 | 0.193 | 0.253 | 0.218 | 0.215 | 0.243 | 0.251 |
| | α_{mean} | | | | | | | | | | | |
| EER | 5.98 | 6.61 | 6.64 | 6.01 | 8.17 | 8.91 | 4.93 | 7.63 | 6.88 | 6.92 | 8.42 | 8.85 |
| \hat{C}_{Ur} | 0.256 | 0.272 | 0.26 | 0.255 | 0.301 | 0.303 | 0.2 | 0.293 | 0.274 | 0.2719 | 0.321 | 0.333 |
| \hat{C}_{Ur}^{min} | 0.226 | 0.241 | 0.241 | 0.227 | 0.276 | 0.284 | 0.178 | 0.250 | 0.228 | 0.231 | 0.283 | 0.286 |
| | α_{min} | | | | | | | | | | | |
| EER | 14.94 | 15.91 | 13.24 | 10.63 | 11.83 | 10.2 | 11.4 | 14.8 | 10.47 | 10.22 | 12.67 | 12.16 |
| \hat{C}_{Ur} | 0.388 | 0.44 | 0.383 | 0.348 | 0.408 | 0.338 | 0.333 | 0.453 | 0.338 | 0.333 | 0.427 | 0.384 |
| \hat{C}_{Ur}^{min} | 0.382 | 0.424 | 0.365 | 0.335 | 0.399 | 0.323 | 0.318 | 0.414 | 0.311 | 0.314 | 0.387 | 0.358 |
| | Without pre-class | | Results in [PGL12] | | | | Without pre-class | | Results in [PGL12] | | | |
| EER | 6.78 | | 5.91 | | | | 7.9 | | 10.03 | | | |
| \hat{C}_{Ur} | 0.249 | | 0.237 | | | | 0.312 | | 0.36 | | | |
| \hat{C}_{Ur}^{min} | 0.205 | | 0.195 | | | | 0.247 | | 0.2969 | | | |

people is not likely to be consistent in the time where the pen pressure reaches a peak. The value of P_M , is likely to be dependent on the writing surface, the pen, etc., making it hard to make pre-classification decisions based only on this feature. On the other hand, in the case of T_{pd} , the result was unexpected, since it is believed that forgers are not able to accurately reproduce the pen down time of the genuine writers. The global based features used for pre-classification, leading to error rate improvements, were: T , T_{v_x} and \bar{P} . In the case of T , the results do nothing but confirm the well known fact that this feature is a good discriminator [LBA96]. The value of \bar{P} , is likely to be more consistent than the other considered pressure based features, since people may not make considerable changes in the average pressure when signing. Finally, T_{v_x} proved to have a high discriminative power, since it is probable that forgers may go back (*i.e.*, negative x velocity) several times during the writing process.

The best error rates were achieved using α_{max} and T_{v_x} , and using α_{mean} and T , for the Dutch and Chinese data, respectively. For the Chinese data, the result is not surprising since T is a highly discriminative feature. For the Dutch data, the result could be explained based on the fact that, in most of the cases, horizontal traces are more significant than vertical ones. Then, differences in the time in which the writer is writing forward would indicate that a forged signing process is taking place.

5.5.2 Multivariate case

Three different ways of computing α are described in Section 5.2 for the multivariate case. For the univariate case discussed above, the results using α_{max} are, in most of the cases, better than those using α_{mean} , while using α_{min} leads to the worst results. This is also the case for the multivariate case analysed here. Moreover, in this case, the results using α_{mean} and α_{min} are not good, and for this reason they are not shown here. The verification results, with the pre-classification performed based on the feature vector combining all the considered global based features for α_{max}^2 , are shown in Table 5.2, for the Dutch (left) and Chinese (right) data, respectively. The best results are indicated in **boldfaced** style.

Table 5.2: Verification results for the Dutch and Chinese Datasets

| | | Dutch Dataset | | Chinese Dataset | |
|-------------------------|-----------------------|--------------------|------------|--------------------|--|
| α_{max}^2 | EER | 11.23 | | 13.01 | |
| | \hat{C}_{llr} | 0.3046 | | 0.3533 | |
| | \hat{C}_{llr}^{min} | 0.2978 | | 0.3494 | |
| $\lambda\alpha_{max}^2$ | EER | 3.09 | | 4.99 | |
| | \hat{C}_{llr} | 0.1589 | | 0.2173 | |
| | \hat{C}_{llr}^{min} | 0.1121 | | 0.1696 | |
| System | Commercial | 1st non-commercial | Commercial | 1st non-commercial | |
| \hat{C}_{llr} | 0.2589 | 0.4928 | 0.4134 | 0.5651 | |
| \hat{C}_{llr}^{min} | 0.1226 | 0.2375 | 0.2179 | 0.3511 | |

The results shown in Table 5.2 corresponding to α_{max} , are not good. This is an unexpected result, since it would be reasonable to expect that increasing the feature space dimensionality would increase the signatures separability. Figure 5.3 shows an example of the distribution of feature vectors composed by the signature total time duration and the pen down duration for the genuine and forged signatures of an author in the Training Set. Of course, this is a two-dimensional simplification of the multivariate case proposed here where the six global based features are combined into a feature vector. Nevertheless, Fig. 5.3 can be used to illustrate the idea that combining the global based features increases the discriminative power. The ellipsoid in solid line in Fig. 5.3 is the one defined by α_{max}^2 computed resorting to (5.2). Note that, even though α_{max} is being used, there are genuine signatures that lie outside this ellipsoid which will be wrongly classified as forgeries. This is probably due to the fact that α is always computed over a separate set of genuine signatures used exclusively for training purposes, without taking into account the forgeries which are also available in the Training Set. However, it is clear from Fig. 5.3 that it is possible to enlarge the ellipsoid in such a way that less genuine signatures lie outside it so that the results can be

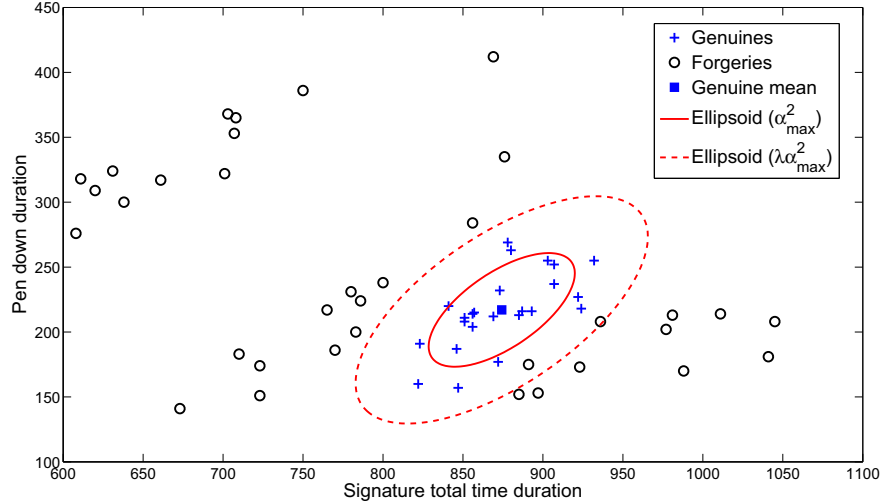


Figure 5.3: Distribution of the global feature vectors (composed by the signature total time duration and the pen down duration) for the genuine and forged signatures of an author in the Training Set.

improved. This is also illustrated in Fig. 5.3, where a larger ellipsoid containing the original one has been plotted in dashed line. Then, by redefining the threshold (enlarging it) a better result would be expected to be obtained. This is the idea here, where a factor $\lambda > 1$ is proposed in order to obtain a more suitable threshold by multiplying α_{max}^2 by $\lambda > 1$.

It is important to note that the genuine and forged signatures are, in general, more overlapped when the global based features are considered individually than when they are combined in a unique feature vector. Note that, if the dimensionality of the feature vectors in Fig. 5.3 is reduced (by projecting them to one axis), the corresponding representations of the genuine and forged signatures will not be easily separable. The results obtained when applying each of the three α criteria will differ in the amount of genuine signatures that will be wrongly classified as forgeries but, it will not be possible, with only one feature, to avoid misclassifications. Multiplying α by a factor (as proposed for the multivariate case) will not make any improvements, since the two classes are strongly overlapped. For this reason, the proposed threshold can not be easily modified to improve the results (as in the multivariate case). On the other hand, for the multivariate case, genuine and forged signatures are not that overlapped. Then, the fact that some genuine signatures are being missclassified is not due to the overlapping of the classes but to the threshold definition. Note that, the proposed redefinition of the threshold only makes sense in the multivariate case.

The verification results, for the multivariate case and the redefined threshold

$\lambda\alpha_{max}^2$, are also shown in Table 5.2. Parameter λ was optimised over the Training Set, being equal to 5 for the Dutch and equal to 4 for the Chinese signatures. These results show that the new threshold leads to improvements with respect to the case of using α_{max}^2 . The results also show improvements with respect to the cases where no pre-classification is performed (last row section in Table 5.1).

The results in the multivariate case outperform the best ones in the univariate case. This confirms that the global based features have a high discriminative power when they are combined. Note that, the increment in the signatures separability due to the increment of the feature space dimensionality, is only possible if the selected global based features are discriminative enough. This is the case here since the global based features used correspond to the better ranked ones among several commonly used global based features ([RKD05] and [FNL⁺05]). If this were not the case, increasing the feature space dimensionality would not necessarily improve the results.

For comparison purposes, the results for the best commercial and non-commercial systems in Sig-Comp2011 [LMd⁺11] are included in the last row section of Table 5.2. Note, from Tables 5.1 and 5.2 that, even for the univariate case, the best results are comparable to those in the state-of-the-art. Moreover, the results for the multivariate case outperform those in the state-of-the-art.

In addition, the pre-classification helps to simplify and speed up the verification process. In Fig. 5.4, the percentage of the pre-classified signatures out of the total amount, for each of the three α criteria (univariate case), and both thresholds α_{max}^2 and $\lambda\alpha_{max}^2$ (multivariate case), are shown for the Dutch (left) and Chinese (right) data, respectively. Note that, still in the most conservative cases (α_{max} for the univariate and $\lambda\alpha_{max}^2$ for the multivariate cases), an important part of the whole set of signatures is discarded making the system to further process less signatures. In particular, 25% and 40% of the Dutch and Chinese signatures, respectively, for the univariate case, and 30% and 55% of the Dutch and Chinese signatures, respectively, for the multivariate case, are discarded.

5.6 Some Concluding Remarks

The addition of a pre-classification stage based on global features to an online signature verification system has been proposed. This approach proved to be a good alternative to combine global and time function based features, since it has shown to be capable of exploiting the discriminative power of the global based features to improve the overall performance with respect to the case where only time function based features are used and no pre-classification is carried out. In addition, the obtained results are comparable to the ones of the state-of-the-art. The method has the advantage of being very simple, since it is based only on global based features, but proved to be powerful, allowing significant improvements in verification errors, process speed and simplicity of the whole

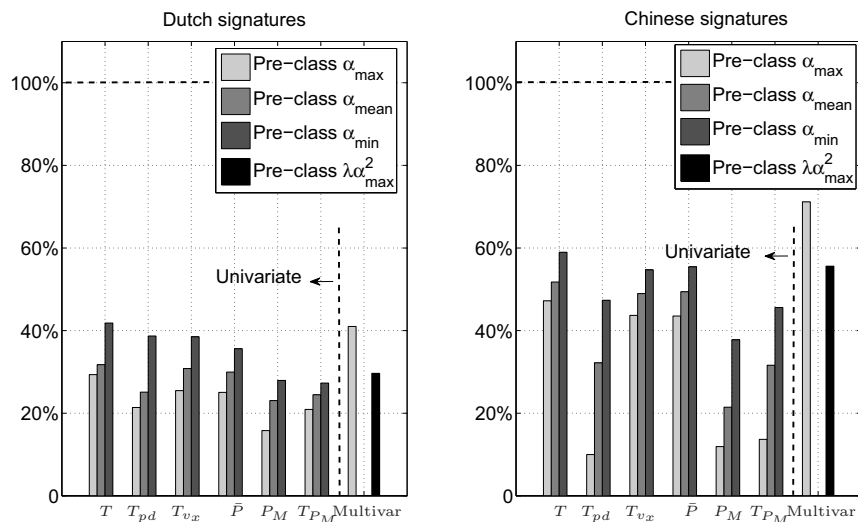


Figure 5.4: Percentage of Dutch (left) and Chinese (right) signatures pre-classified as forgeries, using the different definitions of the threshold.

signature verification system.

Experiments using each one of the global based features separately to perform the pre-classification were carried out to evaluate their individual discriminative power. The total signature time duration, the pen down duration and the average pressure showed to be discriminative enough so that when used for pre-classification the obtained verification results outperform the ones obtained when no pre-classification is performed.

Experiments combining the global based features in a unique feature vector, taking into account their correlation, were also carried out. The verification results obtained in this case outperform the ones obtained when using the features individually. This was expected since, being the features discriminative enough, an increment in the feature space dimensionality is likely to improve the separability between genuine and forged signatures.

5.7 Some Ideas for Future Work

Since the use of global features for pre-classification purposes proved to have several advantages, it would be interesting to select particular global and time function based features in order to adapt the system to different applications. For example, the use of FHE based features in such a system would be of a great benefit for both, FHE and PR communities.

Online Signature Verification: Bridging the gap between FHE and PR

In this Chapter, the discriminative power of some features which seems to be relevant to signature analysis by FHEs is studied for online signature verification. In addition, the discriminative power of this set of features is particularly compared to the discriminative power of an automatically selected feature set. To combine these two types of features is also proposed on the basis of two information fusion schemes. Finally, the feasibility of using only FHEs based features for automatic online signature verification is evaluated. For this purpose, global and time function based features which are relevant to FHEs are combined.

6.1 Introduction

To bridge the gap between the PR and the FHE communities is an important task in the field of signature verification, being crucial for PR researchers to develop useful tools for FHEs. In recent years, several workshops and tutorials like, AFHA 2011 and 2013, have been carried out in order for PR and FHE researchers to meet each other and discuss the current issues in the field. Many signature verification competitions addressing forensic significant scenarios have also been carried out within ICDAR and ICFHR (which are two of the main conferences in the field) like the ones presented in [BvdHCKL09], [LvdHFM10], [LMd⁺11], [LMA⁺12], [MLA⁺13a]. As a result, some works addressing the problem of forensic handwritten signature verification have recently been published in the literature [ML12], [MLD13a], [MLDF14], [MLD13b], [MALD13]. In [ML12] the performance evaluation used by FHE is analysed, while in [MLDF14] a comparative analysis between man and machine based techniques is presented. Local features

are proposed for forensic signature verification in [MLD13a] and [MALD13]. Even the problem of the disguised signatures has been taken into account in [MLD13b].

The above discussion shows some examples of how researchers are currently working towards a more fluent collaboration in order for both communities to be benefited. Although some improvements have been achieved, there is still much work to be done. In this direction, it would be interesting to investigate an online feature set containing features which are relevant to FHEs. This could help them to better understand the signatures and the writer behaviour. Some results regarding the use of features motivated by FHEs are presented in [SBOJ07] and [PL07]. In their daily casework, FHEs work with the offline specimens of the signature, making it not possible to look at online features. Nevertheless, they have plenty of experience in the interpretation of some dynamic information that can be inferred from the offline signature [BFBR11], [Wil12], [CM12]. Usually, FHEs try to understand the forgery process from the forgers point of view. For a successful forgery, the forger must imitate all habits of the authentic writer and the qualities of the authentic signature, and must discard all conflicting elements of his own writing. There will be then a trade-off between accuracy and velocity. Then, to produce an accurate copy of the specimen signature, the forger will likely write slowly, resulting in bad line fluency and hesitations that will be visible for the FHEs. On the other hand, if the forger focuses on the writing velocity to make the forgery more fluent he will aim at a variation that fits within the authentic writer's variability. A monotonous pressure over all the signature, the presence of more pressure in unusual places or slightly different curves at specific spots, can also be signs of forgeries.

In this Chapter, the idea is to show that an automatic online signature verification system based only on the use of a small set of FHE based features could provide verification results comparable with the ones in the state-of-the-art, and to show that FHEs could benefit by combining automatically selected features and the features they usually employ in their daily casework. For this purpose, the discriminative power of a set of features relevant to FHEs is analysed and compared to the discriminative power of automatically selected feature sets. For both types of features, the feature extraction is performed using the wavelet approach introduced in 4.2.2.2 and a RF classifier is used for classification purposes. After this analysis, the question arises whether these two types of features could be combined in order to enhance their discriminative power and improve the system's verification performance. Two different fusion strategies, namely, feature level fusion and decision level fusion, are proposed to this purpose. Finally, the feasibility of using only FHEs based features for automatic online signature verification is evaluated. Both, global features and features based on the wavelet representation of the time functions associated with the signing process, which are relevant to FHEs, are considered. Two combination approaches of global and time function FHE based features are proposed. One of them, consists in

a pre-classification of the signatures (the one introduced in Chapter 5) based on FHE global based features so that gross forgeries can be discarded, followed by a RF classifier using time function based FHE features. The other one, consists in a decision level fusion of two RF classifiers using global and time function FHE based features, respectively.

The Chapter is organized as follows. Section 6.2 describes some time function based features that are relevant to FHEs, and defines the set of time function based FHEs features to be considered in this work. The comparison between the discriminative power of a set of FHE based features and an automatically selected feature set is performed in Section 6.3. In Section 6.4, the combination of these two set of features is proposed on the basis of two well-known fusion information techniques. Finally, the feasibility of performing automatic online signature verification based only on FHE based features is studied in Section 6.5. Some concluding remarks are given in Section 6.6 and some future directions are discussed in Section 6.7.

6.2 FHE Time Function Based Features

In recent years, much effort has been devoted to try to incorporate the forensic handwriting expertise into the field of automatic signature verification. In particular, some works in the literature [SBOJ07], [PL07] address this problem by proposing the computation of features motivated by forensic handwriting examination.

As mentioned in Section 6.1, FHEs work with the static image of the signature, so it is not possible for them to look at online features. However, they can infer dynamic properties from the signature image, to some extent. FHEs consider velocity and curvature as distinctive features. On the other hand, the acceleration and the pen position (which can be established by striae and inkless starts) are less useful for them. Pen pressure is not a useful feature either, since it is strongly dependent on external factors such as the writing material and surface. Although pressure is not a useful feature for FHEs, pressure fluctuations are interesting for them, since they are highly individualistic to the writer. In this Chapter, the following time functions are considered as the FHE based features:

- velocity (v_T (magnitude) and θ (direction)),
- curvature (ρ),
- first order time derivative of the pressure (dp).

It is important to note that these features were selected based on FHEs criteria for Latin scripts. It is likely that for FHEs who examine Chinese scripts, the criteria would be different. Since information about FHEs criteria for Chinese

scripts was not available for the author, the same FHE based feature set is used in this Chapter for both signature styles.

Each of these selected time functions is approximated by the wavelet decomposition introduced in Subsection 4.2.2.2. In particular **db4** wavelets are used in this Chapter.

6.3 FHE vs. Automatically Selected Features

The SigComp2011 Dataset is used for the verification experiments. The meta-parameters of the verification system, that is, the ones for the wavelet decomposition and the RF classifier, are set to their corresponding optimised values presented in Subsection 4.4.2. To obtain statistically significant results, a 5-fold CV like the one already used for the experiments in Chapter 4 (described in Subsection 4.4.1.1) is performed over the Testing Set to estimate the testing errors. The performance is evaluated on the basis of the EER , the \hat{C}_{llr} and its minimal possible value \hat{C}_{llr}^{min} .

The verification results are shown in Table 6.1 for the Dutch (left) and Chinese (right) data, respectively. The first row corresponds to the results obtained using the proposed FHE based features, while the second row corresponds to the results obtained using the automatically selected features in the second row of Table 4.8. The verification results corresponding using the four best ranked automatically selected features are shown in the third row. For the purposes of comparison, the verification results for the best commercial and non-commercial systems in the SigComp2011 competition are also included in the last two rows of Table 6.1.

Table 6.1: Verification results for the Dutch (left) and Chinese (right) Datasets

| | Dutch Dataset | | | Chinese Dataset | | |
|----------------|---------------|-----------------|-----------------------|-----------------|-----------------|-----------------------|
| Features | db4 wavelets | | | db4 wavelets | | |
| | EER | \hat{C}_{llr} | \hat{C}_{llr}^{min} | EER | \hat{C}_{llr} | \hat{C}_{llr}^{min} |
| FHEs' feat. | 9.59 | 0.3408 | 0.2966 | 10.27 | 0.3454 | 0.2760 |
| ASF | 6.58 | 0.2426 | 0.2049 | 7.455 | 0.2962 | 0.2483 |
| ASF (4 feat.) | 9.67 | 0.3365 | 0.2948 | 9.91 | 0.3948 | 0.3265 |
| System | Acc. | \hat{C}_{llr} | \hat{C}_{llr}^{min} | Acc. | \hat{C}_{llr} | \hat{C}_{llr}^{min} |
| commercial | 96.27 | 0.2589 | 0.1226 | 93.17 | 0.4134 | 0.2179 |
| 1st. non-comm. | 93.49 | 0.4928 | 0.2375 | 84.81 | 0.5651 | 0.3511 |

From Table 6.1, it can be observed that the best verification results are obtained when using the automatically selected features (second row), for both datasets. In the case of using the features based on the FHEs criterion (first row), the verification performance is not as good as the one corresponding to the

automatically selected features, but it is still a very good performance for both datasets. This result is promising since these features have a meaningful interpretation by FHEs. This would suggest that, in case the verification system has to be limited to take into account only FHE based features, its performance would not be substantially deteriorated. In fact, if all the features that FHEs look at could be implemented (which is hard to do since some features used by FHEs, such as line quality and ink intensity variations, are not appropriately defined to be computed automatically), the performance might even be better. Moreover, taking into account other results in the state-of-the-art reported over the same datasets (last two rows of Table 6.1), it can be concluded that the performance of systems using only the FHE based features will still be comparable to the ones reported in the state-of-the-art. The fact that the results obtained when using the automatically selected features outperform the ones obtained when using the FHEs based features is probably due to the fact that the automatic feature selection incorporates more features to model the signatures. Anyway, the automatic selection is always keeping the FHE based features among the selected ones, that is, the FHE based features are considered important by the automatic feature selection algorithm. These features have been thoroughly investigated by FHEs and are generally accepted by the FHE community.

To perform a fairer comparison between FHE based features and automatically selected ones, the number of automatically selected features was set to equal the number of FHE based features (that is 4). The results using only the four most important features automatically selected are included in the third row of Table 6.1. The subset of selected features are x , a_T , y and v_T for the Dutch, and y , x , p and v_T for the Chinese datasets, respectively. It can then be observed that, when having a limitation in the number of features to be considered, the automatically selected features do not coincide with the FHE based features (only the v_T is kept) and the discriminative power of the set is deteriorated. This is indicating that for the automatically selected features to be significant it is likely that the size of the resulting set has not to be limited in such a way that more features can be taken into account and so can be their corresponding interaction when performing the selection to reach a more meaningful selected set of features. This is not the case when using the FHE based features, since they are highly discriminative by themselves. Then, when having a limitation in the size of the feature set, the FHE based features would be more reliable and a better design option. In fact, they outperform (or equal) the results obtained by the reduced set of automatically selected features.

There are certain signatures that are wrongly classified when using the FHE based features, while they are correctly classified when using the automatically selected ones. Figure 6.3 shows the FHE based features, and the corresponding wavelet approximation, for a sample (shown in the bottom row) of a Dutch (left) and a Chinese (right) wrongly classified signature, respectively. Note that the

wavelet approximations are not so good due to the fact that the time functions corresponding to the shown signature samples are not smooth. Experiments (not included here) showed that incorporating the detail coefficients improves the approximation accuracy, at the cost of increasing the feature vector length. This would not be a limitation in the case of using the set of FHE based features since this set contains only four features. In any case, to pick this type of signatures out of a database and to analyse their stability could be an interesting issue for further work. As will be shown later in Section 6.4, the combination of FHE based features and automatically selected ones will improve the overall error rates, even when dealing with this type of signatures.

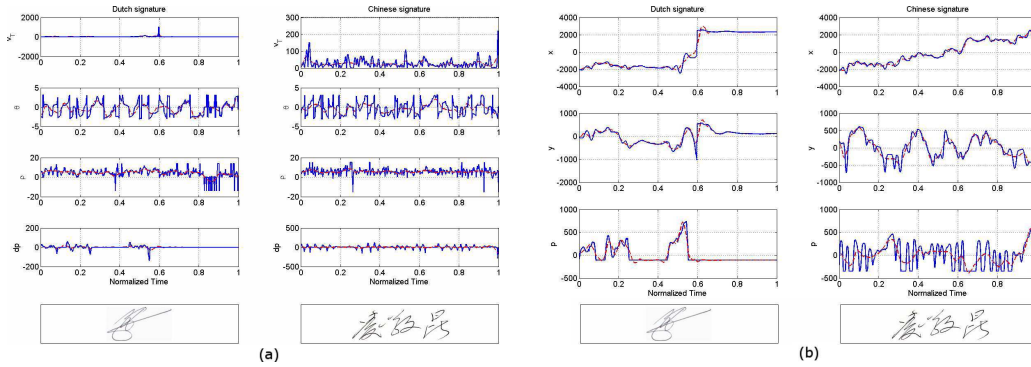


Figure 6.1: Wrongly classified signatures with FHEs based features. (a) Original v_T (first row), θ (second row), ρ (third row) and dp (fourth row) (blue solid line) and their corresponding approximations (red dashed line), for the Dutch (left) and Chinese (right) data. Bottom row: Associated signature images. (b) Original x (top), y (middle) and p (bottom) (blue solid line) and their corresponding approximations (red dashed line), for the Dutch (left) and Chinese (right) data.

6.4 On Combining FHE based Features and Automatically Selected Features

Based on the results in Table 6.1, the question arises whether the discriminative capability of both FHE based features and automatically selected ones could be combined to improve the performance of the system. Traditionally, three main approaches for information fusion can be distinguished, namely, early or feature level fusion, intermediate or classifier level fusion, and late or decision level fusion. In the feature level case, the feature vectors coming from different sources are concatenated to obtain a combined feature vector which is then used in the classification stage. In the classifier level approach, which is typically encountered

in applications where HMM and Dynamic Bayesian Networks (DBN) are used to model the different signals involved, a composite classifier is generated by combining the individual classifiers used to process the different signals. Finally, in the late or decision level fusion approach, a final decision is obtained by combining the probability/likelihood scores from the separate classifiers processing the different signals. A good overview on information fusion techniques for the case of audiovisual signals, in applications of Human-Computer Interfaces can be found in [SSR10]. References abound on the use of information fusion techniques in the field of biometrics, see for instance [ZCAM12] and [Lak13] on ear, and [KP10] on iris biometrics, respectively. Fusion information has also been used in the field of signature verification, as it is shown in [RJI12] and [GFRO05] for the offline and online approaches, respectively. Figure 6.4 schematically depicts the three main approaches for information fusion. In this Section, feature level and decision level approaches will be considered to combine the discriminative capabilities of FHE based features and automatically selected features. In this case, classifier level fusion is not possible due to the particular classifier being used.

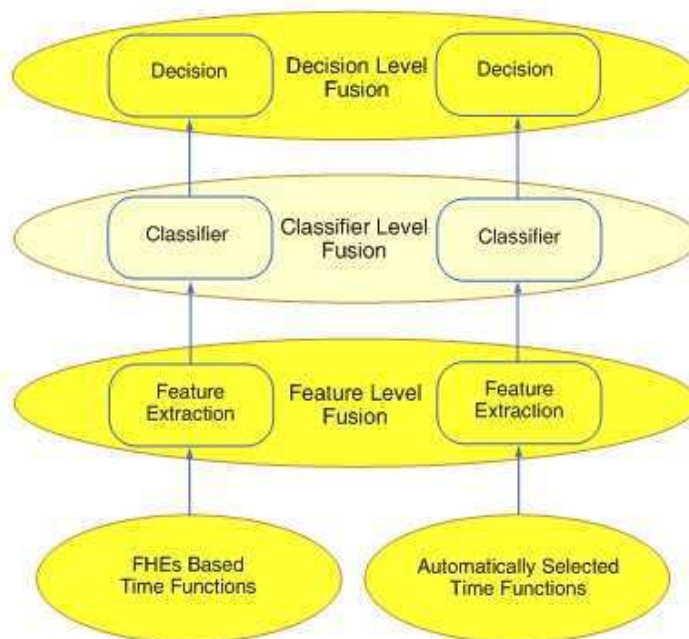


Figure 6.2: Fusion information schemes.

It is important to note that the different sets of automatically selected features considered here contain different features for the Dutch and Chinese datasets (see Table 4.8). Then, it would be reasonable to expect that a verification system based on the fusion of the FHE based features and the automatically selected

ones (at feature level or decision level), not only would combine the discriminative capabilities of both feature sets, but also could adapt better to the different signature styles (Dutch and Chinese). That is, since the automatically selected feature sets are particularly suited for each signature style, so will be the combination (although the same FHE based feature set is used for both signatures styles), giving to the verification system more adaptability and flexibility.

Regarding a combination/fusion at feature level, it is clear that since the automatically selected feature set includes the FHE based features and also the reduced set of automatically selected features, the only combination that would make sense is the one between the FHE based features and the reduced set of automatically selected features. A feature vector consisting of the concatenation of the four FHE based features and the first three ranked features in the automatically selected feature set is considered in this case. Note that only the first three automatically selected features are being considered since the fourth feature is v_T , which is one of the four FHE based features. For the Dutch dataset these three features are: x , a_T and y , while for the Chinese dataset, they are: y , x and p . The verification results for this fusion strategy are shown in the third row of Table 6.2. It can be observed that for the Dutch dataset an improvement in the verification errors with respect to the ones obtained with each set of features separately (rows 1 and 2 of Table 6.2) is achieved with the combined feature vector. On the other hand, for the Chinese dataset only the results obtained with the reduced set of automatically selected features are improved with the combined feature vector.

Table 6.2: Verification results for the decision level fusion between the FHE based features and the whole set of automatically selected features, for the Dutch (left) and Chinese (right) Datasets.

| | Dutch Dataset | | | Chinese Dataset | | |
|-----------------------|---------------|-----------------|-----------------------|-----------------|-----------------|-----------------------|
| | EER | \hat{C}_{llr} | \hat{C}_{llr}^{min} | EER | \hat{C}_{llr} | \hat{C}_{llr}^{min} |
| FHE based feat. | 9.59 | 0.3408 | 0.2966 | 10.27 | 0.3454 | 0.2760 |
| ASF (4 feat.) | 9.67 | 0.3365 | 0.2948 | 9.91 | 0.3948 | 0.3265 |
| Feature Level Fusion | 7.93 | 0.316 | 0.2606 | 8.19 | 0.3349 | 0.2905 |
| Decision Level Fusion | 7.3 | 0.2649 | 0.2333 | 7.69 | 0.3146 | 0.2668 |

Regarding a combination/fusion at decision level, independent classifiers are used for each of the feature sets to be combined and the final decision is computed as a combination of the likelihood scores associated with each classifier. In particular, the following widely used combination/fusion rule *weighted geometrical combination rule* [KHDM98], is considered in this Section:

$$P_{fused} = P_{FHE}^{1-\gamma} \cdot P_{ASF}^{\gamma}, \quad (6.1)$$

where P_{fused} is the likelihood score for the combined scheme, P_{FHE} and P_{ASF} are the likelihood scores for the FHE based features and the automatically selected feature set, respectively, and $0 \leq \gamma \leq 1$ is a user defined parameter weighting the individual likelihood scores. In order to compare the results of this fusion approach with the ones obtained with the fusion at feature level, here also the combination is performed between the FHE based features and the reduced set of automatically selected features. Note that, also for this fusion approach, only the first three automatically selected features are used. In this way, v_T is only taken into account once. The value of γ is optimised over the Training Set for each dataset. Figure 6.3 shows the C_{llr}^{min} error as a function of γ for the Dutch (left) and Chinese (right) datasets. From Fig. 6.3 the optimal values are $\gamma^{Dutch} = 0.7$ and $\gamma^{Chinese} = 0.44$ for the Dutch and the Chinese datasets, respectively. From Table 6.2 it can be observed that an improvement in the verification errors with respect to the ones obtained with each set of features separately (rows 1 and 2 of Table 6.2) is achieved with the fusion at decision level, for both datasets.

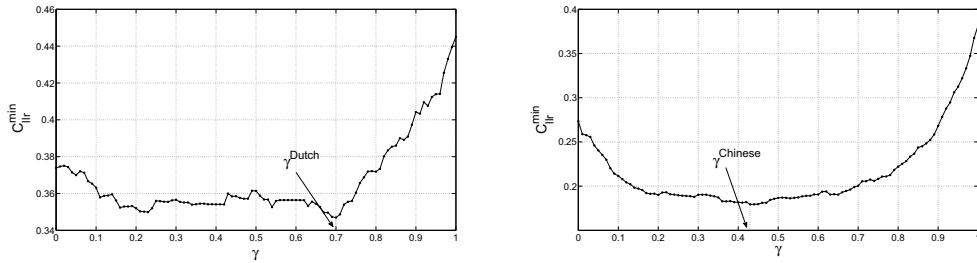


Figure 6.3: C_{llr}^{min} error as a function of γ over the Training Set for the Dutch (left) and Chinese (right) datasets.

From the above discussion, it can be observed that for Chinese data the verification performance of the individual systems is only improved when fusion at the decision level is performed. The verification results obtained with the FHE based features are much better than the ones obtained with the reduced set of automatically selected features (see Table 6.2). Due to this notorious difference in the individual discriminative power of each one of the combined set of features, it is likely that the simple combination of these features in a unique feature vector (feature level fusion) would not have enough discriminative power to outperform the results obtained when using only the FHE based features. On the other hand, the fusion at the decision level does improve the verification results obtained with each set of features separately. Then, the use of weights to fuse the individual likelihood scores contributes to combine the individual discriminative power in a more efficient way. Note that the optimal value of γ is $\gamma^{Chinese} = 0.44$, meaning that the likelihood scores obtained with the FHE based features (which are more discriminative than the reduced set of automatically selected features) have a higher weight in the combination. For the Dutch data, it has already been shown

that the fusion of both feature sets reaches better verification results than the ones obtained with the individual systems, for both fusion approaches. The fact that both fusion approaches yield verification result improvements, is probably due to the fact that the individual verification performances are quite similar. Note that the fusion at decision level (row 4 in Table 6.2) yields the best verification results (highlighted in **boldfaced** style) for both datasets, and therefore this would be the fusion approach of choice for the considered features and datasets. In addition, the verification results obtained in this case are comparable to the ones in the state-of-the-art (last two rows in Table 6.1). It is important to highlight that these very good verification results are obtained using only features that are meaningful to FHEs and some other features (the three most important from the automatically selected feature set) that are very simple and easily interpretable. This is promising since it shows that using only few features and, even more important, simple and easy to interpret ones (specially those relevant to the FHEs) good results can be achieved avoiding the use of too many features or more complex ones, which would mean more computational load and less interpretability of the whole system.

For the sake of completeness, the combination at decision level between the FHE based features and the whole set of automatically selected features has been also carried out. The results for this fusion are shown in the last row of Table 6.3. It can be observed that these results outperform the ones obtained with the FHE based features (row 1 in Table 6.2), but do not outperform the ones obtained with the whole set of automatically selected features (row 2 in Table 6.3). Note that, for the Chinese data, the fusion results are almost equal to the ones corresponding to the whole set of automatically selected features, while for the Dutch data, the fusion results are worse than the ones corresponding to the whole set of automatically selected features. It is likely that the fusion results do not outperform the ones obtained using the whole set of automatically selected features because the FHE based features are contained in the whole set of automatically selected features. Based on these comments it can be concluded that this combination is not worth taking into account, since the results of the combination do not improve the best results obtained individually, and the resulting number of features is large and many of them are difficult to interpret. The attentive reader has already probably noticed that none of the fusion strategies considered yield results outperforming the ones obtained using the whole set of automatically selected features. However, the results obtained with the decision level fusion of the FHE based features and the reduced set of automatically selected features are not that far from the best results, so that by trading-off accuracy with system complexity, this fusion approach would be the best strategy.

Table 6.3: Verification results for the decision level fusion between the FHE based features and the whole set of automatically selected features, for the Dutch (left) and Chinese (right) Datasets.

| | Dutch Dataset | | | Chinese Dataset | | |
|-----------------------|---------------|----------------|----------------------|-----------------|----------------|----------------------|
| | EER | \hat{C}_{lr} | \hat{C}_{lr}^{min} | EER | \hat{C}_{lr} | \hat{C}_{lr}^{min} |
| FHE based feat. | 9.59 | 0.3408 | 0.2966 | 10.27 | 0.3454 | 0.2760 |
| Auto. Selec. feat. | 6.58 | 0.2426 | 0.2049 | 7.455 | 0.2962 | 0.2483 |
| Decision Level Fusion | 6.8 | 0.2452 | 0.2125 | 7.6 | 0.2926 | 0.2444 |

6.5 Automatic Online Signature Verification based only on FHE Features: an Oxymoron?

Finally, after the analysis carried out in Sections 6.3 and 6.4, an interesting question arises: is it possible to develop an automatic online signature verification system based exclusively on FHE based features in such a way that reliable results could be obtained? In this Section, the idea is to make some contributions towards answering this question. The set of time function based features relevant to FHEs described in Section 6.2 (hereafter referred to as TFFHE features) together with a set of global based features also relevant to FHEs (GFHE features) are considered in this case. In an attempt to answer the above question, these two types of features are combined in different ways so that their main characteristics can be exploited.

How to combine global and time function based features have already been discussed in Chapter 5. Here, two different verification schemes are proposed. One of them consists in the pre-classification stage presented in Chapter 5 based on GFHE features, followed by a RF classifier based on the TFFHE features, and the other one in a decision level fusion like the one used in Section 6.5.2 between GFHE and TFFHE features.

6.5.1 FHE Based Global Features

In Chapter 5, a set of global based features corresponding to the better ranked ones by the feature selection performed in [RKD05] and [FNL⁺05] is used, namely, signature total time duration T , pen down duration T_{pd} , positive x velocity duration T_{vx} , average pressure \bar{P} , maximum pressure P_M and the time at which the pressure is maximum T_{P_M} . This set of features proved to be well suited for pre-classification purposes as the results in Section 5.5 demonstrate.

Global based features are useful for FHEs when doing their daily casework. Moreover, these features are necessary for them to perform the examination of the signatures. Looking at global features can reveal that a forgery process has

taken place since poor forgeries are dissimilar in global features. Pen trajectories are very writer specific, in particular, the path it follows towards the starting point. FHEs can make an estimation of the duration of the signature by looking at characteristics related to the writing speed (thin lines, feathering, connected strokes) and the total line length. On the other hand, global features computed based on the pressure time function, are not so reliable since pressure levels will fluctuate in the inktrace. For example, taking the average pressure, will result in a significant loss of information. Then, global features based on the pressure time functions would not be of the choice of an FHE.

The above analysis suggests that global features based on the pen trajectories (time and space) are relevant to FHEs. In addition, FHEs can estimate the time duration of the signature. Then, global based features associated with the pen trajectories such as signature total time duration and pen down duration would be good choices. On the other hand, as already mentioned above, global features based on the pressure time function would not be good choices since they have not practical usability. Then, from the original set of global based features used in Chapter 5, the following ones would be relevant to FHEs and they will be considered as the GFHE features in this Section:

- signature total time duration T
- pen down duration T_{pd} .

Of course, there are other global characteristics that FHEs look at when doing their daily casework. The ones considered here have the advantage of being a subset of some of the most widely used and better ranked global features in the PR literature and, also important, of being very simple to compute.

The GFHE as it is the case of the TFFHE features presented in Section 6.2 have been selected based on FHE criteria for Latin scripts. A similar comment to the one done in the case of the TFFHE features stands for this case.

6.5.2 Global and Time function FHE based features fusion

Two different fusion strategies schemes are considered to combine the discriminative capabilities of both GFHE and TFFHE features. In the first approach, the GFHE features are used for the multivariate pre-classification described in Chapter 5 in such a way that a signature whose GFHE feature vector is far away from its corresponding mean in the training set is discarded, declaring it as a forgery. If this is not the case, the signature is represented by the TFFHE features and then fed to a RF classifier in a subsequent stage. In the second approach, a combination of GFHE and TFFHE features based on decision level fusion is proposed. In this case, the decision of two RF classifiers fed by each type of features (GFHE and TFFHE) is fused based on the combination rule introduced in Section 6.4.

6.5.3 Evaluation Protocol

The SigComp2011 Dataset is used for the verification experiments. The optimisation of the meta-parameters of the proposed verification systems is performed, for each dataset in the SigComp2011 Dataset, over the corresponding Training Set, while the Testing Set is used for independent testing purposes. To obtain statistically significant results, a 5-fold CV is performed over the Testing Set to estimate the verification errors. For the pre-classification scheme, a 5-fold CV like the one described in Section 5.4 is performed. In this particular case, the global based features T and T_{pd} are computed in order to construct the GFHE feature vector (\mathbf{g}_{test}), while the DWT approximation coefficients are computed for the considered time functions (v_T , θ , ρ and dp) associated with the input signature, in order to construct the TFFHE feature vector for training the subsequent RF classifier if needed. For the decision level fusion scheme, the first step for each instance of the 5-fold CV is to extract the GFHE and TFFHE features from a signature of a particular writer from one of the testing sets in the 5-fold CV. Then, the classification is performed as follows: two RF classifiers are trained, one of them using GFHE features and the other one using TFFHE features. Each classifier is trained by a genuine class consisting of the current writer's genuine class in the training set of the 5-fold CV, and a forged class consisting of the genuine signatures of all the remaining writers in the same set. The result of the verification process is then the combination of the outputs of these two classifiers computed as in (6.1).

6.5.4 Results and Discussion

For both approaches, the meta-parameters of the RF classifier are set like in Subsection 4.4.1. On the other hand, for the wavelet approximation of the time functions, some changes are introduced in the selection of the meta-parameters in order to improve the results. In Subsection 4.4.2, the time functions were resampled to a normalised length of 256, and the different feature sets being considered were represented using the **db4** wavelet decomposition with a resolution level of 3. In this case, the time function normalised length remains the same, but it is considered the fact that a better approximation accuracy would be obtained with a lower resolution level, at the cost of increasing the feature vector length. To increase the length of the feature vector is not a limitation when using TFFHE since this set contains only four features. Then, for these experiments, the wavelet resolution level is set to 2.

Parameters α and β have to be chosen for the pre-classification and the decision level fusion approaches, respectively. Parameter α is computed resorting to (5.2) over the Training Sets corresponding to the Dutch and Chinese datasets. Figure 6.4 shows the \hat{C}_{llr}^{min} error as a function of β for the Dutch (left) and Chinese (right) datasets. From Fig. 6.4, the optimal values are $\beta^{Dutch} = 0.13$ and

$\beta^{Chinese} = 0.06$ for the Dutch and the Chinese datasets, respectively.

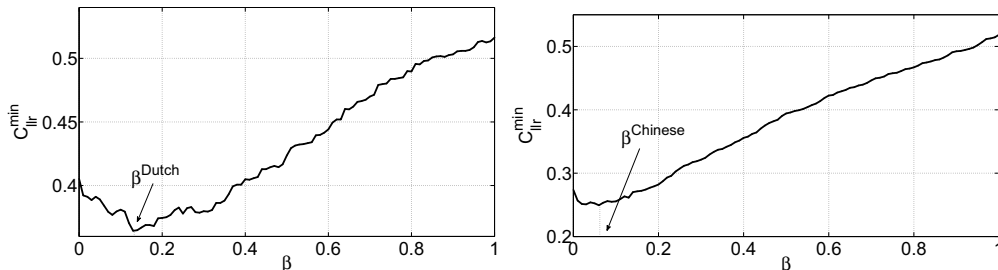


Figure 6.4: \hat{C}_{lr}^{min} error as a function of β over the Training Set for the Dutch (left) and Chinese (right) datasets.

The verification results for the case of the pre-classification approach (PC) are presented in the first row of Table 6.4, while the ones corresponding to the case of the decision level fusion approach (DLF) are presented in the second row of Table 6.4, for the Dutch (left) and Chinese (right) datasets, respectively. For comparison purposes, the verification results corresponding to the automatically selected features considered in Section 6.3 which are presented in Table 6.1, are included in the third row of Table 6.4. The best results in Table 6.4 are highlighted in **boldfaced** style.

Table 6.4: Verification results for the Dutch (left) and Chinese (right) Datasets

| | Dutch Dataset | | | Chinese Dataset | | |
|---------------------------|---------------|----------------|----------------------|-----------------|----------------|----------------------|
| | EER | \hat{C}_{lr} | \hat{C}_{lr}^{min} | EER | \hat{C}_{lr} | \hat{C}_{lr}^{min} |
| Pre-Classification | 4.42 | 0.222 | 0.178 | 5.98 | 0.265 | 0.211 |
| Decision Level Fusion | 8.12 | 0.335 | 0.298 | 7.63 | 0.299 | 0.251 |
| ASF | 6.58 | 0.2426 | 0.2049 | 7.455 | 0.2962 | 0.2483 |

It can be observed that the results obtained with the ASF features have been improved by the present pre-classification approach. This is an important result since it shows that using only features that are fully interpretable by FHEs it is possible not only to achieve a good performance but also a better performance than the one obtained when using more complex and difficult to interpret automatically selected features. In addition, the results obtained when using the pre-classification approach are comparable to the ones of the best non-commercial verification systems reported in [LMd⁺11] (see last two rows of Table 6.1).

On the other hand, no improvements over the verification results obtained using ASF features are achieved when using the decision level fusion approach. In fact, the results obtained when using the pre-classification approach clearly

outperform the ones corresponding to the decision level fusion approach. This shows that these global based features are not useful when using the decision level fusion for combining the features. Even more, the values of β^{Dutch} and $\beta^{Chinese}$ are low, pointing out a poor contribution of the global based features in the fused score. Further, it is likely that even incorporating more global based features, the pre-classification approach will still get better results than the decision level fusion approach since the former can exploit the intrinsic characteristics of the global based features in a better way. This observation is in line with the ones made in Chapter 5 where the benefits of using global based features for pre-classification purposes are discussed.

It can then be concluded that, based only on FHE based features, the pre-classification is the best combination approach. As already mentioned, it achieves results comparable with those in the state-of-the-art, being even better than the results obtained when more complex and difficult to interpret automatically selected features are used. It is also important to note that the proposed combination approaches have the advantage of using a small number of FHE based features.

The presented results provide an answer to the question in the title of the Section showing that the oxymoron is not true. Specifically, the exclusive use of FHE based features for automatic online signature verification is technically sound, when properly combined, since error rates comparable to those of the state-of-the-art are obtained.

6.6 Some Concluding Remarks

The feasibility of using only FHE based features for automatic online signature verification has been shown in this Chapter. Two different approaches combining global features and features based on the wavelet representation of the time functions associated with the signing process which are relevant to FHEs, have been studied. The results show that the approach based on the pre-classification outperforms the one based on the decision level fusion. The best results obtained with the proposed methods which use exclusively FHE based features are comparable to those of the state-of-the-art over the same datasets, being specially remarkable the fact that they outperform the results using a RF classifier based on more complex and difficult to interpret automatically selected features. The results are promising since they could make automatic signature verification techniques useful tools for the FHE community.

6.7 Some Ideas for Future Work

During the analysis of the verification performance obtained with the FHE based features it has been pointed out that there are some types of signatures that are not well classified with these features. To pick this type of signatures out of a database and analyse their stability would be an interesting issue to further research.

It is likely that, if more FHE based features, either global or time function based ones, could be computed automatically and could be incorporated to the verification system, the results could be improved.

Online Signature Verification: When the whole is greater than the sum of the parts

Is it possible to combine all the online features proposed throughout the previous Chapters in order to achieve better verification results? In this Chapter, this question is addressed. Global features, automatically selected time function based features and FHE time function based features are combined in different ways so that their main characteristics can be exploited.

7.1 Introduction

Throughout the previous Chapters, different features have been proposed and analysed for online signature verification. In Chapter 4, the discriminative power of several time function combinations has been evaluated. In that case, a set of automatically selected features achieved the best verification results. In Chapter 5, the discriminative power of a set of widely used global based features was analysed and used for pre-classification purposes. Finally, in Chapter 6, features relevant to FHEs were studied. The use of each of these type of features proved to be interesting, resulting not only in good verification performances, but also providing different advantages to the verification systems. The question arises whether it would be possible to combine these feature sets in such a way that the verification performance could be improved with respect to the case of using each one of them individually.

The automatically selected time function based features presented in the second row of Table 4.8, the global based features presented in Subsection 5.3.1, and the FHE time function based features presented in Subsection 6.2 are then combined. The idea is to combine them in such a way that their main characteristics can be exploited. In particular, the use of two different combination

strategies that have already proved to be useful techniques in Chapters 5 and 6 is proposed. One of them is based in the multivariate pre-classification introduced in Subsection 5.5.2 and the other one is based on the decision level fusion strategy presented in Section 6.4.

The Chapter is organized as follows. The proposed feature set combinations are presented in Section 7.2. In particular, the decision level fusion approach is introduced in Subsection 7.2.1, and the pre-classification approach is presented in Subsection 7.2.2. In Section 7.3, the evaluation protocol is described. Subsection 7.3.1 focuses on the decision level fusion experiments and Subsection 7.3.1 focuses on the pre-classification ones. The results are presented and discussed in Section 7.4. Finally, some concluding remarks are given in Section 7.5 and in Section 7.6 some future directions are discussed.

7.2 Feature Combination Approaches

As already mentioned in Section 7.1, three different feature sets will be combined, *viz.*, the automatically selected time function based features corresponding to the ones in the second row of Table 4.8, the global based features presented in Subsection 5.3.1 (T , $T_p d$, T_{vx} , \bar{P} , P_M and T_{P_M}) and the FHE based time functions presented in Subsection 6.2 (v_T , θ , ρ and dp). In this Chapter, these feature sets will be referred to as ASF, GF and TFFHE, respectively.

A decision level fusion like the one used in Subsection 6.4 and the multivariate pre-classification introduced in Section 5.5.2 are used to combining the ASF, GF and TFFHE feature sets. Subsection 7.2.1 is devoted to the description of the decision level fusion approach experiments, while Subsection 7.2.2 is devoted to the pre-classification approach experiments.

7.2.1 Decision Level Fusion

Feature level and decision level approaches could be used to combine the discriminative capabilities of the different feature sets considered here, *viz.*, GF, ASF and TFFHE.

Regarding a fusion at feature level, it is clear that since the ASF set contains the TFFHE set, feature level fusion between these two sets would not make sense. Two separate experiments fusing GF features with ASF features on one hand, and fusing (at a feature level) GF features with TFFHE features on the other, were carried out. The verification results obtained (not shown here) in both cases did not improve the ones corresponding to the case of using the ASF feature set and the TFFHE feature set individually.

Regarding a fusion at decision level, independent RF classifiers are used for each of the feature sets to be combined and the final decision is computed as a combination of the likelihood scores associated which each classifier. In this

case, the decision of three RF classifiers fed by each type of features (GF, ASF and TFFHE) is fused based on the *weighted geometrical combination rule* already used in Section 6.4, that is:

$$P_{fused} = P_{GF}^{\beta} P_{ASF}^{\gamma} P_{TFFHE}^{(1-\beta-\gamma)}, \quad (7.1)$$

where P_{fused} is the likelihood score for the combined scheme, P_{GF} , P_{ASF} and P_{TFFHE} are the likelihood scores for the classifiers based on GF, ASF and TFFHE features, respectively, and $0 \leq \beta \leq 1$ and $0 \leq \gamma \leq 1$ are user defined parameters weighting the individual likelihood scores.

7.2.2 Pre-classification

Two different experiments employing the pre-classification approach are considered. One of them, uses GF features for pre-classification while the subsequent classification stage is performed on the basis of a decision level fusion of two RF classifiers fed by ASF and TFFHE features, respectively. The other one, uses TFFHE features for pre-classification while for the subsequent classification stage employs a decision level fusion of two RF classifiers fed by ASF and GF features, respectively.

7.3 Evaluation protocol

The SigComp2011 Dataset is used for the verification experiments. The optimisation of the meta-parameters of the proposed verification systems is performed, for each dataset in the SigComp2011 Database, over the corresponding Training Set, while the Testing Set is used for independent testing purposes. To obtain statistically significant results, a 5-fold CV is performed over the Testing Set to estimate the verification errors. Particular details for the 5-fold CV procedure are given in Subsections 7.3.1 and 7.3.2 for the decision level fusion approach and the pre-classification approach, respectively.

7.3.1 Decision level fusion

For each instance of the 5-fold CV, a signature of a particular writer from one of the testing sets in the 5-fold CV is fed to the system. The GF, ASF and TFFHE features are extracted from it. The classification is performed as follows: three RF classifiers are trained, each of them using GF, ASF and TFFHE features. Each classifier is trained by a genuine class consisting of the current writer's genuine class in the training set of the 5-fold CV, and a forged class consisting of the genuine signatures of all the remaining writers in the same set. The result of the verification process is then the combination of the outputs of these three classifiers computed as in (7.1).

7.3.2 Pre-classification

For each instance of the 5-fold CV, a signature of a particular writer from one of the testing sets in the 5-fold CV is fed to the system. The global features (T , T_{pd} , T_{vx} , \bar{P} , P_M and T_{PM}) or the FHE based features (v_T , θ , ρ and dp) are computed for the input signature, depending on the pre-classification scheme to be performed, in order to construct the GF or the TFFHE feature vectors (\mathbf{g}_{test}), respectively. Then, the pre-classification is performed as follows: the distance in (5.1) between \mathbf{g}_{test} and $\bar{\mathbf{g}}_{train}$ (sample mean computed over the current writer's genuine signatures available in the training set of the 5-fold CV) is computed. If this distance is larger than the threshold (α^2), the signature is declared to be a forgery. If this is not the case, the signature is subjected to the subsequent classification stage, as follows: the DWT approximation coefficients are computed for the considered automatically selected time functions associated with the input signature, in order to construct the ASF feature vector. Then, two RF classifiers are trained, one of them with the ASF features and the other one with the GF or the TFFHE, depending on the pre-classification scheme being performed. Each RF classifier is trained by a genuine class consisting of the current writer's genuine class in the training set of the 5-fold CV, and a forged class consisting of the genuine signatures of all the remaining writers in the same set. A decision level fusion is performed over the two classifiers outputs, giving the final output of this classification stage. The result of the verification process is then either the result of the pre-classification (the input signature is declared a forgery), or the result of the decision level fusion of the RF classifiers.

7.4 Results and Discussion

For both approaches, the meta-parameters of the RF classifier are set like in Subsection 4.4.1.2 and the meta-parameters corresponding to the **db4** wavelet approximation of the time functions are set like in Subsection 6.5.4.

For the multivariate pre-classification approach, parameters α and λ have to be chosen. The former is computed resorting to (5.2) over the Training Sets, while the latter was also optimised over the Training Sets, being set to $\lambda = 5$ for both pre-classification approaches, *i.e.* for the pre-classification performed on the basis of the GF and the pre-classification performed on the basis of the TFFHE features, and both datasets.

For the decision level fusion approach, parameters β and γ have to be chosen. Their optimal values are obtained by minimising the \hat{C}_{ur}^{min} over the corresponding Training Sets, and they are: $\beta^{Dutch} = 0.2$ and $\gamma^{Dutch} = 0.5$ for the Dutch dataset, and $\beta^{Chinese} = 0.1$ and $\gamma^{Chinese} = 0.8$ for the Chinese dataset. Using the optimised values of β and γ the fusion rule in (7.1) will become:

$$P_{fused} = P_{GF}^{0.2} P_{ASF}^{0.5} P_{TFFHE}^{0.3}, \quad (7.2)$$

for the Dutch data and:

$$P_{fused} = P_{GF}^{0.1} P_{ASF}^{0.8} P_{TFFHE}^{0.1}, \quad (7.3)$$

for the Chinese data.

The verification results for the proposed combination schemes are presented in Table 7.1, for the Dutch (left) and Chinese (right) datasets, respectively. Regarding the pre-classification approach, the results corresponding to the case of the pre-classification approach based on GF (PC-GF) are presented in the first row of Table 7.1. On the other hand, the results obtained in the case of the pre-classification approach based on TFFHE (PC-TFFHE) are not good and, therefore, it is considered that it does not make sense to include them in Table 7.1. Regarding the decision level fusion approach (DLF), the results are presented in the second row of Table 7.1.

As already mentioned, the idea here is to evaluate if the combination of the different types of features could be more powerful than the different types of features used individually. In order to compare the verification results of the combination approaches with the individual performances, the verification results corresponding to the ASF and TFFHE features used individually shown in Table 6.1 are included in the third row of Table 7.1. State-of-the-art results from the SigComp2011 Competition as reported in [LMd⁺11], over the same datasets, are also included in the last two rows of Table 7.1, for comparison purposes. In particular, the ones corresponding to the best commercial and non-commercial verification systems are included.

Table 7.1: Verification results for the Dutch (left) and Chinese (right) Datasets

| | Dutch Dataset | | | Chinese Dataset | | |
|---------------------------------|---------------|-------------------|-------------------------|-----------------|-------------------|-------------------------|
| | EER | \tilde{C}_{llr} | \tilde{C}_{llr}^{min} | EER | \tilde{C}_{llr} | \tilde{C}_{llr}^{min} |
| PC-GF | 3.55 | 0.172 | 0.133 | 5.1 | 0.194 | 0.162 |
| DLF | 6.95 | 0.261 | 0.228 | 7.31 | 0.268 | 0.218 |
| ASF [PGLA13a] | 6.58 | 0.243 | 0.205 | 7.455 | 0.296 | 0.248 |
| Comm. [LMd ⁺ 11] | – | 0.259 | 0.123 | – | 0.413 | 0.218 |
| Non-comm. [LMd ⁺ 11] | – | 0.493 | 0.237 | – | 0.565 | 0.351 |

Regarding the experimental results for the pre-classification approach mentioned above, it is important to note that GF features resulted to be a better choice than the TFFHE features for pre-classification purposes. Global features can roughly represent signatures considering them as a whole. Then, gross forgeries can be correctly detected on the basis of global features and a simple distance based classifier. On the other hand, TFFHE features do not represent the signatures as a whole, since they are computed from the time functions associated with the signatures. Then, TFFHE features can also model details of the signatures,

which are probably not that easy to analyse, making it necessary the use of a more powerful classifier to separate forgeries from genuine signatures.

From Table 7.1, it can be observed that the pre-classification based on global features (PC-GF) obtains better verification results than the ones corresponding to the decision level fusion (DLF) approach, for both datasets. This shows that, among the combination approaches proposed in this Chapter for combining the considered feature sets, the pre-classification is the most suitable one. This is probably due to the nature of the considered feature sets. As already mentioned, GF are very useful for pre-classification purposes, then, the PC-GF approach exploits the intrinsic characteristics of GF in a better way than the DLF approach.

The results in Table 7.1 also show that, for both datasets, the PC-GF approach outperforms the ASF approach, outperforming, in this way, the TFFHE and GF approaches individually. The fact that the proposed PC-GF approach outperforms the ASF approach is a very important result. It is reasonable to expect that features selected out from a meaningful original feature set by a widely used automatic feature selection technique would reach such good results that it would be difficult to outperform them. Nevertheless, in this Chapter it is shown that a proper combination of ASF features with another meaningful (but not that discriminative when used individually) features such as GF and TFFHE ones can improve the performance. On the other hand, the DLF approach improves the results corresponding to the ASF approach only for the Chinese dataset.

It can also be observed from Table 7.1 that the verification results of the two proposed combination approaches are comparable to those of the state-of-the-art reported in [LMd⁺11]. In particular, the best results of the non-commercial verification system in [LMd⁺11] are improved for both datasets, while only for the Chinese dataset improvements are achieved with respect to the best commercial verification system in [LMd⁺11]. In any case, the verification results obtained with the PC-GF approach for the Dutch dataset, are comparable with the corresponding to the best commercial verification system in [LMd⁺11]. As has already been pointed out in this Thesis, results reported over the same datasets in the state-of-the-art (like, for instance, [LMd⁺11]), show better performances for Dutch data than for Chinese data, meaning that Chinese data is more complex. In this Chapter, the results obtained for the Chinese data with the proposed feature combinations are particularly promising since they are even better than the best commercial verification system in [LMd⁺11].

Finally, it can be concluded, from the above comments on the results in Table 7.1, that the best combination strategy is to use GF features for pre-classification and to perform decision level fusion with the additional information regarding ASF features and TFFHE features. In addition, it is important to note that the verification results obtained with this combination (the most discriminative one) outperform the verification results obtained with the multivariate pre-classification proposed in Subsection 5.5.2 (shown in Table 5.2), which were

up to this Chapter, the best verification results obtained in this Thesis. This means that the proposed combination strategy based on the use of GF features for pre-classification and performing decision level fusion with the additional information regarding ASF features and TFFHE features, yields to the best verification results for online signature verification reached in the whole Thesis. It is in this sense that it can be said that the whole is greater than the sum of the parts as the title of the Chapter states.

7.5 Some Concluding Remarks

The idea in this last Chapter was to evaluate if the combination of the different types of features could be more powerful (in the sense of having a better discriminative capability) than the different types of features used individually. Among the different combination proposed, the pre-classification outperformed the decision level fusion strategy. The obtained results show that, for both datasets, the PC-GF approach outperforms the ASF approach and, in this way, it also outperforms the TFFHE and GF approaches individually. The fact that the proposed PC-GF approach outperforms the ASF approach is a very important result, since it shows that a proper combination of ASF features with another meaningful (but not that discriminative when used individually) features such as GF and TFFHE ones, can improve the already good results obtained when using only ASF features.

The best combination strategy is to use GF features for pre-classification and to perform decision level fusion with the additional information regarding ASF features and TFFHE features. The verification results obtained with this combination are the best ones for online signature verification reached in this Thesis.

7.6 Some Ideas for Future Work

Studying the stability of the signatures is still an open challenge for the author of this Thesis. This analysis would be undoubtedly useful to improve the selection of features and the combination strategies.

Conclusions and Future Work

In the present Thesis some of the most important actual challenges in the field of automatic signature verification has been addressed. Contributions towards offline and online signature verification were done.

Regarding the offline case, a new feature extraction technique was proposed on the basis of a circular grid scheme and the computation of graphometric features. In addition, the FFT was used in order to achieve the rotation invariance property. The proposed new circular grid scheme proved to have several advantages with respect to other gridding techniques widely used in the literature. Among them, the better adaptability to the shape of the signatures can be highlighted. Moreover, the incorporation of the rotation invariance property, makes the proposed feature extraction more robust and useful. Finally, the obtained verification results are technically sound since they are comparable to similar ones reported in the state-of-the-art.

Regarding the online case, contributions towards feature extraction and selection were done. New techniques for feature extraction based on approximating the time functions associated with the signing process by orthogonal polynomials series, was proposed. In particular, Legendre polynomials and wavelet decomposition were used. Both approximations proved to be well suited for modeling the time functions, having important advantages such as achieving fixed-length feature vectors and allowing for a dimensionality reduction with respect to the case of using all the points in the time functions. In addition, the obtained verification results are among the best ranked ones reported in the state-of-the-art over the same datasets. An in depth analysis of different time function combinations was carried out in order to give some insight of their actual discriminative power. It is expected that the results obtained from these studies could make some contributions towards the long-term discussion about which are the best features for online signature verification. Following this direction, a new consistency measured was also defined to quantify the discriminative power of the features. This consistency factor proved to be well correlated with the verification results, meaning that it could be used for feature selection purposes.

To incorporate a pre-classification stage based on global based features to an online signature verification system for the purposes of improving its performance, has also been proposed. This approach proved to be a good alternative to combine global and time function based features, since showed to be capable of exploiting the discriminative power of both types of features. The proposed pre-classification approach has the advantage of being very simple, since it is based only on global based features, but proved to be powerful, allowing significant improvements regarding verification errors, process speed and simplicity of the whole signature verification system.

Some contributions towards bridging the gap between FHE and PR communities were presented. The discriminative power of some features which seem to be relevant to signature analysis by FHEs is studied. Based on this analysis, the feasibility of developing a system which could complement the FHEs work is evaluated. The results are promising, since the verification error rates obtained when using the FHE based features are comparable to the ones reported in the state-of-the-art over the same datasets. Moreover, they show that the proposed signature verification system could be integrated into toolkits that could be used by FHEs for helping them to analyse and further understand the signatures. Based on these promising results, a fusion between automatically selected features and FHE based features is proposed in order to improve the verification results. The experimental results, obtained by combining these two types of features on the basis of two different information fusion schemes (feature level fusion and decision level fusion), show that it is possible to combine both types of features to improve the verification performance, without necessarily increasing the system complexity. Finally, the feasibility of using only FHE based features for automatic online signature verification is evaluated. Two combination approaches of global and time function features which are relevant to FHEs, are proposed. One of them is based on the pre-classification approach and, the other one is based on a decision level fusion. Experimental results show that automatic online verification systems using exclusively FHE based features achieve verification performances comparable to those of the state-of-the-art over the same datasets. These results are promising since they could make automatic signature verification techniques useful tools for the FHE community.

Finally, the question arises whether it is possible to combine all the proposed online features in order to achieve better verification results. In an attempt to give some answers to this question, automatically selected features, FHE based features and global features are combined based on two different fusion strategies, namely, pre-classification and decision level fusion. The verification results show improvements when combining the features in a suitable way. Achieving the best verification results in this Thesis and showing that, in some cases, *the whole is greater than the sum of the parts*.

There are several interesting and open challenges in the field of automatic

signature verification. This Thesis was an attempt to address some of them.

To analyse the influence of the cultural origin of the signatures in the verification performance is an important issue in the field. In the present Thesis, Dutch and Chinese signatures have been analysed in order to make some contributions towards this direction. Based on the this study, some interesting observations could be done that could constitute a promising starting point to further research in this area.

Another important problem to solve in the field is the lack of sufficient amount of data in practical applications. One option, would be to develop verification systems that could deal with small size training sets. Other option, and the most studied until now, would be to generate artificial training samples. The latter is an interesting challenge for future work.

A

Notation

| | |
|-------|---------------------------------------------------------------|
| ANN | Artificial Neural Network |
| AFHA | Automated Forensic Handwriting Analysis |
| ASF | Automatically Selected Features |
| DET | Detection Error Tradeoff |
| DBN | Dynamic Bayesian Network |
| DFT | Discrete Fourier Transform |
| DFL | Decision level Fusion |
| DTW | Dynamic Time Warping |
| DWT | Discrete Wavelet Transform |
| EER | Equal Error Rate |
| FAR | False Acceptance Rate |
| FFT | Fast Fourier Transform |
| FHE | Forensic Handwriting Expert |
| FPR | False Positive Rate |
| FRR | False Rejection Rate |
| HMM | Hidden Markov Model |
| ICDAR | International Conference on Document Analysis and Recognition |

| | |
|---------|------------------------------------------------------------------|
| ICFHR | International Conference on Frontiers in Handwriting Recognition |
| k-NN | k-Nearest Neighbour |
| PC | Pre-classification |
| PDA | Personal Digital Assistant |
| PDF | Personal Density Function |
| PR | Pattern Recognition |
| RBF | Radial Basis Functions |
| RF | Random Forest |
| ROC | Receiver Operating Characteristic |
| SIFT | Scale-Invariant Feature Transform |
| SURF | Speed up Robust Features |
| SVC2004 | 2004 Signature Verification Competition |
| SVM | Support Vector Machine |
| TPR | True Positive Rate |

B

Datasets

B.1 GPDS300Signature CORPUS Database

The GPDS300Signature CORPUS is an offline database which is public and freely available for research purposes. A detailed description of this database can be found in [VFTA07]. There are 160 writers enrolled in the database. For each writer, there are 24 genuine signatures, and 30 forged signatures, that is, a total of $160 \times 24 = 3840$ genuine and $160 \times 30 = 4800$ forged signatures. The database authors state that forgeries in the database are simple and skilled forgeries. By using the term simple forgery authors mean that the signature has not been practised by the forger, while in the case of a skilled forgery, forgers are allowed to practise the signature for as long as they deem it necessary. Signatures in the database were collected as follows: Writers were provided with A4 sheets of paper and black or blue ink. They were asked to fit their signatures (genuine or forged samples) into two different types of boxes, one of them a 5 cm wide and 1.8 cm high box, and the other one a 4.5 cm wide and 2.5 cm high box. Half of the samples were written in each type of box. All the genuine samples of each writer were collected in single day writing sessions. Forgeries were collected in a single day writing session asking each forger to imitate three signatures of five signers. Forgers were given a static image of the three genuine signatures (scanned at 300dpi). For each signer in the database there are 30 forgeries done by 10 different forgers. A standard scanner with 75 dpi was used to acquire the signatures in an 8-bit gray scale image. A freely distributed version of the described database is available from: <http://www.gpds.ulpgc.es/download/index.htm>.

B.2 SigComp2011 Database

The SigComp2011 Dataset is one of the most recent databases presented in the state-of-the-art, and it is public and freely available for research purposes. This database has been used for some of the most recent signature verification com-

petitions [LMd⁺11] and [MLA⁺13b]. A detailed description of the database can be found in [LMd⁺11]. There are two separate datasets in the SigComp2011 database, containing offline as well as online samples of Western (Dutch) and Chinese signatures, respectively. Since the database was created to be used in several competitions, each dataset is divided into a Training and a Testing Set. The forgeries in the database are skilled forgeries. The signatures were acquired using a ballpoint pen on paper over a tablet, which is the natural writing process. This is in contrast to the approach of other researchers who tested signatures produced on a PDA or with a Wacom-stylus on a glass or plastic surface. The offline samples of the signatures were scanned at 400dpi, resulting in RGB coloured PNG images. The online samples were collected with a WACOM Intuos3 A3 Wide USB Pen Tablet together with the collection software MovAlyzer at a sampling rate of 200Hz and consist of ascii files with the format: x, y, z , representing the pen coordinates x and y , and the pen pressure. The Dutch (left) and the Chinese (right) datasets are described in Table B.1¹.

Table B.1: Online Dutch (left) and Chinese (right) datasets.

| Dutch Dataset | | | Chinese Dataset | | |
|---------------|----------|-----------|-----------------|----------|-----------|
| Training Set | | | Training Set | | |
| Authors | Genuines | Forgeries | Authors | Genuines | Forgeries |
| 10 | 240 | 119 | 10 | 230 | 429 |
| Testing Set | | | Testing Set | | |
| Authors | Genuines | Forgeries | Authors | Genuines | Forgeries |
| 54 | 1296 | 611 | 10 | 219 | 461 |

¹The amount of genuine and forged signature samples may differ from those in [LMd⁺11] since when making signatures available for the research community some of them were missing [Sig11].

Support Vector Machine

Support Vector Machines is a quite recent technique of statistical learning theory developed by Vapnik ([Vap95], [Vap98]).

Suppose a separable classification problem in a two-dimensional input space. There are several separating hyperplanes that can separate the two data classes (Fig. C.1a). Nevertheless, a unique separating hyperplane has to be chosen.

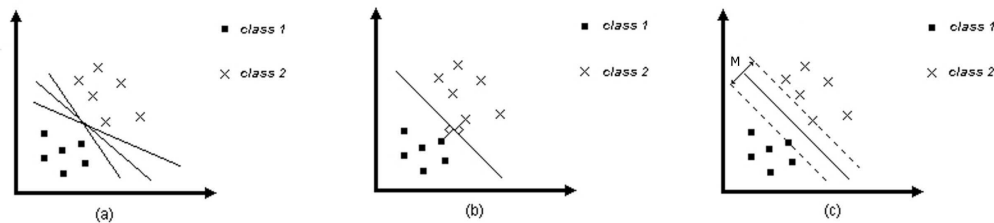


Figure C.1: Separable classification problem example: (a) Possible separating hyperplanes; (b) Selection of a unique hyperplane maximizing the distance between the nearest point of each class; (c) Optimal separating hyperplane that maximizes the margin.

The separating hyperplane is chosen so that the nearest points to the hyperplane satisfy $|\omega^T x + b| = 1$. Then, the points satisfying $\omega^T x + b = 1$ will be the support vectors of one of the classes while the points satisfying $\omega^T x + b = -1$ will be the support vectors of the other class. Defining the “margin” (M) as the sum of the distances between the hyperplane and the closest point of each class (Fig. C.1b), the separating hyperplane will be optimal if it maximizes this margin (Fig. C.1c). In this case, the margin M equals $2/\|\omega\|_2$ and the problem is solved by minimizing $\|\omega\|_2$ subject to the restrictions imposed by the data.

Formally, the problem can be presented as follows. Consider a given training set $\{x_k, y_k\}_{k=1}^N$, with input data $x_k \in \mathbb{R}^n$, output data $y_k \in \{-1, +1\}$ and the

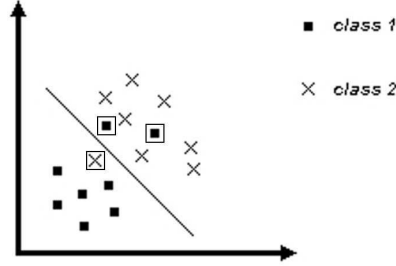


Figure C.2: Non-separable classification problem example.

linear classifier $y(x) = \text{sign}[\omega^T x + b]$. In a separable case, the classifier can be rewritten as

$$y_k[\omega^T x_k + b] \geq 1, \quad k = 1, \dots, N, \quad (\text{C.1})$$

and it is easy to notice that no mistakes are committed in the classification procedure (Fig. C.1).

However, in a more general case of non-separable data, one cannot avoid misclassifications (Fig. C.2). In this case, it is necessary to introduce additional slack variables (ξ_k) in the formulation problem to represent the classification error. Then, the set of inequalities takes the following form

$$y_k[\omega^T x_k + b] \geq 1 - \xi_k, \quad k = 1, \dots, N. \quad (\text{C.2})$$

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 can be written as

$$y_k[\omega^T \varphi(x_k) + b] \geq 1 - \xi_k, \quad k = 1, \dots, N. \quad (\text{C.3})$$

Fortunately, no explicit construction of the nonlinear mapping $\varphi(x)$ is needed. This is possible 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 $k, \ell = 1, \dots, N$.

Finally, the SVM solution can be found by solving the following optimisation problem

$$\begin{aligned} \min_{\omega, b, \xi} \quad & J_P(\omega, \xi) = \frac{1}{2} \omega^T \omega + c \sum_{k=1}^N \xi_k & (\text{C.4}) \\ \text{s.t.} \quad & y_k[\omega^T \varphi(x_k) + b] \geq 1 - \xi_k, \quad k = 1, \dots, N \\ & \xi_k \geq 0, \quad k = 1, \dots, N. \end{aligned}$$

Resorting to the dual of problem (C.4), the solution of the Quadratic Programming (QP) problem is the set of the real positive constants α_k , and the SVM

classifier takes the following form

$$y(x) = \text{sign}\left[\sum_{k=1}^N \alpha_k y_k K(x, x_k) + b\right]. \quad (\text{C.5})$$

Different Kernels have been proposed, being Linear, Polynomial and RBF Kernels among the most popular ones in the literature. Each one of them are respectively defined as follows:

$$\begin{aligned} K_{\text{linear}}(x_k, x_\ell) &= x_k^T x_\ell, \\ K_{\text{polynomial}}(x_k, x_\ell) &= (\tau + x_k^T x_\ell)^d, \\ K_{\text{RBF}}(x_k, x_\ell) &= \exp(-\|x_k - x_\ell\|_2^2 / \sigma^2). \end{aligned}$$

D

Random Forest

D.1 Decision trees

Decision trees [LB84] are widely used in many botany, taxonomy or medical diagnosis, for example, because of their interpretability and accuracy. Basically, a decision tree is a hierarchical set of nodes, starting from a root node, each one containing a decision involving the comparison of an attribute with a given threshold, which then leads to another node or to a leaf (a terminal node with an associated product). Figure 1 shows an example of a decision tree for an artificial problem with two attributes and three products. To evaluate a given sample, it is passed down the tree until a leaf is reached, which gives the associated product. Several automatic procedures were developed to build (or grow) decision trees from a set of samples. All of them are recursive processes with a common strategy: at each step the procedure evaluates all available attributes and possible thresholds, selecting the combination that maximize a given fitness measure. The dataset is then splitted according to the selected decision, and the procedure is recursively applied to both subsets. The different procedures can differ, for example, in the fitness measure (RF maximizes the GINI index ([LB84], C4.5 ([CPfMLbJRQMKP93]) maximizes the entropy), the selection of the thresholds, or the stopping criteria for the recursive growing process.

D.2 Ensemble methods

Ensemble methods combine multiple individual discriminant functions to solve a single problem. The idea is that using multiple models would obtain better predictive performance than using any of the constituent models. Whenever a new sample is fed to the ensemble, each of the individual discriminant function classifies the input and their individual outputs are combined in order to reach the ensemble final decision. Usually, a majority-vote rule is used, that is, each function votes for one of the possible outputs, and the one with more votes is

given as the response of the ensemble. The level of diversity among the ensemble individual members is important and it is desirable to be high. Different ensemble construction strategies can be used. The desirable diversity can be generated by using different kind of individual models (artificial neural networks, decision trees, etc), choosing different subsets from the whole dataset for fitting them or fitting the procedure with different initial conditions. The bias-variance trade-off is a well-known statistical principle being the basis of one of the most successful lines of research among the ensemble techniques. If an statistical methodology has low bias (intrinsic capacity of accurately modeling any distribution) it is likely that this methodology has high variance (diversity in models fitted to different samples of the same distribution). Ensembles should be constructed using several models with low-bias. Of course, these models would have high-variance. Nevertheless, when combining these low-bias high-variance models, the variance is reduced. This results in a discrimination method which is more accurate than its members in average. Furthermore, the accuracy of the ensemble is related to the level of diversity of the individual members [GRC⁺05].

D.3 Random Forest

Random Forest is an ensemble of decision trees. The ensemble construction strategy is focused in increasing the diversity among the trees. Since decision trees are very unstable (generally a small change in the dataset results in large changes in the developed model [Bre96]), the diversity among the trees in the ensemble can be increased by fitting each tree on a bootstrap replicate (random subset of the available data, of the same length, taken with replacement) of the whole data. In addition, more diversity is introduced during the growing of each tree. For each node the method selects a small random subset of m attributes (from the total number of attributes available) and use only this subset to search for the best split. The combination of these two sources of diversity produces an ensemble with good predicition performance. This performance will depend on the correlation between any two trees in the forest and on the strength of each individual tree. The more strong the individual trees are and the less correlated they are, the best error rate the classifier will achieve. The parameters to adjust for a Random Forest classifier are the number of trees to grow and the number of randomly selected splitting variables to be considered at each node. The number of trees to grow does not strongly influence the results as long as it is kept large (generally, 500 trees are enough). Then, in practice, the only tunning parameter of the model is the number of randomly selected splitting variables to be considered at each node. Nevertheless, the results do not strongly depend on this parameter either. Usually, the square root of the total number of attributes is a good choice.

In addition to its very good discriminative capability, Random Forests also

have several advantages. Among them, the following can be mentioned:

- They run efficiently on large databases
- They can handle thousands of input variables avoiding the need for variable selection
- They are fast and can grow as many trees as it is necessary without overfitting
- They only have one relevant tuning parameter and usually there is no need to adjust it since the default value (square root of the total number of attributes) is a good choice.

Bibliography

- [AFFAOG05] F. Alonso-Fernandez, J. Fierrez-Aguilar, and J. Ortega-Garcia. Sensor interoperability and fusion in signature verification: A case study using tablet pc. In *Proc. Intl. Workshop on Biometric Recognition Systems, IWBRIS*, volume 3781 of *LNCIS*, pages 180–187. Springer, October 2005.
- [AFFFOG07] F. Alonso-Fernandez, M. C. Fairhurst, J. Fierrez, and J. Ortega-Garcia. Impact of signature legibility and signature type in off-line signature verification. In *Biometrics Symposium, BSYM*, pages 1–6. IEEE, September 2007.
- [Alh12] M.J. Alhaddad. Multiple classifiers to verify the online signature. *J. World of Computer Science and Information Technology*, 2(2):46–50, 2012.
- [ASVE13] F.A. Alsulaiman, N. Sakr, J.J. Valds, and A. El Sadik. Identity verification based on handwritten signatures with haptic information using genetic programming. *ACM Transactions on Multimedia Computing, Communications, and Applications*, 9(2):11:1–11:21, May 2013.
- [AYF86] M. Ammar, Y. Yoshida, and T. Fukumura. A new effective approach for automatic off-line verification of signatures by using pressure features. In *Proc. 8th Int. Conf. Pattern Recognition*, pages 556–569, 1986.
- [AYF90] M. Ammar, Y. Yoshida, and T. Fukumura. Structural description and classification of signature images. *Pattern Recognit.*, 23:697–710, 1990.
- [BC97] R. Bajaj and S. Chaudhury. Signature verification using multiple neural classifiers. *Pattern Recognit.*, 30(1):1–7, 1997.
- [Bd06] N. Brümmer and J. du Preez. Application-independent evaluation of speaker detection. *Computer Speech & Language*, 20:230–275, 2006.

- [BFBR11] C.L. Bird, B. Found, K.N. Ballantyne, and D. Rogers. Dynamic features of naturally written, disguised and forged handwritten text. In *Proc. of 15th Conf. of the Int. Graphonomics Society*, Cancun, Mexico, June 12-15 2011.
- [BGS10] L. Batista, E. Granger, and R. Sabourin. A multi-classifier system for off-line signature verification based on dissimilarity representation. In *Multiple Classifier Systems*, volume 5997 of *LNCS*, pages 264–273. Springer, 2010.
- [Bis06] C. M. Bishop. *Pattern Recognit. and Mach. Learning*. Springer, Berlin, NY, 2006.
- [BP93] J.J. Brault and R. Plamondon. Segmenting handwritten signatures at their perceptually important points. *IEEE Trans. Pattern Anal. Mach. Intell.*, 15(9):953–957, 1993.
- [BP10] I. Bărbăntan and R. Potolea. Enhancements on a signature recognition problem. In *Proc. of IEEE 6th Int. Conf. on Intell. Comput. Communic. and Process.*, pages 141–147, Cluj-Napoca, Romania, August 2010.
- [Bre96] L. Breiman. Bagging predictors. *Machine Learning*, 26(2):123–140, 1996.
- [Bre01] L. Breiman. Random forests. Technical report, Stat. Dep., U. of California, Berkeley, 2001.
- [Bur98] C. Burges. A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery*, 2:121–167, 1998.
- [BvdHCKL09] V.L. Blankers, van den Heuvel C.E., Franke K.Y., and Vuurpijl L.G. The ICDAR 2009 signature verification competition. In *Proc. of 10th Int. Conf. on Doc. Anal. and Recognit.*, pages 1403–1407, Barcelona, Spain, July 2009.
- [BZZE13] K. Barkoula, E.N. Zois, E. Zervas, and G. Economou. Off-line signature verification based on ordered grid features: An evaluation. In *Proc. 2nd International Workshop On Automated Forensic Handwritten Analysis (AFHA 2013)*, pages 36–40, 2013.
- [CDIP93] G. Congedo, G. Dimauro, S. Impedovo, and G. Pirlo. Offline signature verification by fundamental components

- analysis. In S. Impedovo, editor, *Proceedings of the 7th International Conference on Image Analysis Processing (ICIAP-7)*, pages 331–337. World Scientific, 1993.
- [CDW⁺12] Hao Chang, Daoming Dai, Pingshui Wang, Yong Xu, Fengshan Si, and Shuqin Huang. Online signature verification using wavelet transform of feature function. *J. of Informat. & Computat. Sci.*, 9(11):3135–3142, 2012.
- [CM12] M. Caligiuri and L. Mohammed. *The Neuroscience of Handwriting*. CRC Press, Florida, 2012.
- [CNPM04a] A. Chalechale, G. Naghdy, P. Premaretne, and A. Mertins. Cursive signature extraction and verification. In *Proc. 2nd International Workshop on Information Technology & Its Disciplines (WITID 2004)*, pages 109–113, 2004.
- [CNPM04b] A. Chalechale, G. Naghdy, P. Premaretne, and A. Mertins. Cursive signature extraction and verification. In *2nd Int. Workshop on Information Technology & Its Disciplines (WITID 2004)*, 2004.
- [CPfMLbJRQMKP93] 1993 C4.5: Programs for Machine Learning by J. Ross Quinlan. Morgan Kaufmann Publishers, Inc. Legendre polynomials based feature extraction for online signature verification. consistency analysis of feature combinations. *Machine Learning*, 16(3):235–240, 1993.
- [CS07] W. Chang and J. Shin. Modified dynamic time warping for stroke-based on-line signature verification. In *Proceedings of the International Conference on Document Analysis and Recognition*, pages 724–728, Curitiba, Brazil, Sept. 2007.
- [CVGR05] S. Canu, Y. Grandvalet V., Guigue, and A. Rakotomamonjy. *SVM and Kernel Methods Matlab Toolbox*. Perception Systemes et Information, INSA de Rouen, Rouen, France, 2005.
- [Dau92] I. Daubechies. *Ten Lectures on Wavelets*. SIAM: Society for Industrial and Applied Mathematics, Pennsylvania, 1992.

- [DG14] H.V. Davda and S.K. Gonsai. Offline signature verification system using energy on grid level. *International Journal of Engineering Research*, 3(2):104–107, Feb. 2014.
- [DIPS97] G. Dimauro, S. Impedovo, G. Pirlo, and A. Salzo. A multi-expert signature verification system for bankcheck processing. *Int. J. Pattern Recogn. and Artif. Intell.*, 11(5):827–844, August 1997.
- [DNP98] S. Djeziri, F. Nouboud, and R. Plamondon. Extraction of signatures from check background based on a filiformity criterion. *IEEE Trans. Image Process.*, 7(10):1425–1438, October 1998.
- [DSG94] J.P. Drouhard, R. Sabourin, and M. Godbout. A neural network approach to offline signature verification using directional pdf. *Pattern Recognit.*, 29(3):415–424, 1994.
- [FAT05] M.A. Ferrer, J.B. Alonso, and C.M. Travieso. Offline geometric parameters for automatic signature verification using fixed-point arithmetic. *IEEE Trans. on Pattern Anal. and Machine Intell.*, 27(5):993–997, 2005.
- [FGSD02] M. Fuentes, S. Garcia-Salicetti, and B. Dorizzi. On-line signature verification: Fusion of a hidden markov model and a neural network via a support vector machine. In *Proc. 8th Int. Workshop Front. Handwriting Recognit.*, pages 253–258, Niagara-on-the-Lake, ON, Canada, Aug. 2002.
- [FNL⁺05] J. Fierrez-Aguilar, L. Nanni, J. Lopez-Peñalba, J. Ortega-Garcia, and D. Maltoni. An on-line signature verification system based on fusion of local and global information. In *Proc. of IAPR Int. Conf. on Audio- and Video-Based Biomet. Person Authentication*, pages 523–532, Rye Brook, NY, USA, 2005.
- [FORG07] J. Fierrez-Aguilar, J. Ortega-Garcia, D. Ramos-Castro, and J. Gonzalez-Rodriguez. HMM-based on-line signature verification: Feature extraction and signature modelling. *Pattern Recognit. Letters*, 28:2325–2334, 2007.
- [FSV06] E. Frias-Martinez, A. Sanchez, and J. Velez. Support vector machines versus multi-layer perceptrons for efficient off-line signature recognition. *Engineering Appl. of Artif. Intell.*, pages 693–704, 2006.

- [GC03] J.D. Gibbons and S. Chakraborti. *Nonparametric Statistical Inference*. Marcel Dekker, Inc., New York, 2003.
- [GD98] S. Garcia-Salicetti and B. Dorizzi. On using the viterbi path along with HMM likelihood information for online signature verification. *IEEE Trans. Syst., Man, Cybern. B*, 37(5):1237–1247, October 1998.
- [GFRO05] J. Gonzalez-Rodriguez, J. Fierrez-Aguilar, D. Ramos-Castro, and J. Ortega-Garcia. Bayesian analysis of fingerprint, face and signature evidences with automatic biometric systems. *Forensic Science Int.*, 155:126–140, 2005.
- [GGKS10] C. Gruber, T. Gruber, S. Krinninger, and B. Sick. On-line signature verification with support vector machines based on LCSS kernel functions. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 40(4):1088–1100, 2010.
- [GMH12] R. Guest and O. Miguel-Hurtado. Enhancing static biometric signature verification using speeded-up robust features. In *Proc. IEEE International Carnahan Conference of Security Technology (ICCST)*, pages 213–217, Boston, USA, 2012.
- [GPFMD11] J. Galbally, R. Plamondon, J. Fierrez, and M. Martinez-Diaz. Quality analysis of dynamic signature based on the sigma-lognormal model. In *Proc. IAPR Intl. Conf. on Document Analysis and Recognition, ICDAR*, pages 633–637, September 2011.
- [GRC⁺05] P.M. Granitto, A. Rebola, U. Cervino, F. Gasperi, F. Biasoli, and H.A. Ceccatto. Cascade classifiers for multiclass problems. In *34th International Conference of the Argentine Computer Science and Operational Research Society*, 2005.
- [GS13] S. Garhawal and N. Shukla. SURF based design and implementation for handwritten signature verification. *International Journal of Advanced Research in Computer Science and Software Engineering*, 3(8):3539–3541, 2013.
- [GW08] R.C. Gonzalez and R.E. Woods. *Digital Image Processing*. Prentice Hall, 3rd edition, 2008.

- [GW10] O. Golubitsky and S. M. Watt. Distance-based classification of handwritten symbols. *Int. J. of Doc. Anal. and Recognit.*, 13(2):133–146, 2010.
- [GWE09] R.C. Gonzalez, R.E. Woods, and S.L. Eddins. *Digital Image Processing*. Gatesmark Publishing, 2nd edition, 2009.
- [HC13] J. Hu and Y. Chen. Offline signature verification using real adaboost classifier combination of pseudo-dynamic features. In *Proc. International Conference on Document Analysis and Recognition*, 2013.
- [HGD09] N. Houmani, S. Garcia-Salicetti, and B. Dorizzi. On assessing the robustness of pen coordinates, pen pressure and pen inclination to time variability with personal entropy. In *Proc. of IEEE 3rd Int. Conf. on Biomet.: Theory, Application, and Systems*, Washington DC, USA, 2009.
- [HL77] N. Herbst and C. Liu. Automatic signature verification based on accelerometry. *IBM J. Res. Develop.*, pages 245–253, 1977.
- [HMM05] Y. Hongo, D. Muramatsu, and T. Matsumoto. Adaboost-based on-line signature verifier. In A.K. Jain and N.K. Ratha, editors, *Proc. SPIE, Biometric Technology for Human Identification II*, Orlando, USA, April 2005.
- [HY97] K. Huang and H. Yan. Offline signature verification based on geometric feature extraction and neural network classification. *Pattern Recognit.*, 30(1):9–17, January 1997.
- [HYW04] W. Hou, X. Ye, and K. Wang. A survey of off-line signature verification. In *Proc. International Conference on Digital Object Identifier, Intelligent Mechatronics and Automation*, 2004.
- [IG00] M.A. Ismail and Samia Gad. Off-line Arabic signature recognition and verification. *Pattern Recognit.*, 33(10):1727–1740, 2000.
- [IGHG04] J.J. Igarza, L. Gómez, I. Hernández, and I. Goirizelaia. Searching for an optimal reference system for on-line signature verification based on (x, y) alignment. In D. Zhang and A.K. Jain, editors, *Proceeding of the International*

- Conference on Biometric Authentication (ICBA 2004)*, volume 3072 of *LNCS*, pages 519–525. Springer, 2004.
- [IP92] G. Dimauro S. Impedovo and G. Pirlo. A stroke-oriented approach to signature verification. In S. Impedovo and J.C. Simon, editors, *From Pixels to Features III Frontiers in Handwriting Recognition*, pages 371–384. Elsevier, 1992.
- [IP08] D. Impedovo and G. Pirlo. Automatic signature verification: The state of the art. *IEEE Trans. on Systems, Man, and Cybernetics – Part C: Application and Reviews*, 38(5):609–635, 2008.
- [IPB12] D. Impedovo, G. Pirlo, and D. BarbuZZi. Multi-classifier system configuration using genetic algorithms. In *Proc. of International Conference on Frontiers in Handwriting Recognition (ICFHR)*, pages 560–564, Bari, Italy, 2012.
- [IPP12] D. Impedovo, G. Pirlo, and R. Plamondon. Handwritten signature verification: New advancements and open issues. In *Proc. of 13th Int. Conf. on Frontiers in Handwriting Recognit.*, pages 367–372, Bari, Italy, September 2012.
- [JBP99] A.K. Jain, R.M. Bolle, and S. Pankanti. *Biometrics: Personal Identification in Networked Society*, volume 479 of *The International Series in Engineering and Computer Science*, chapter Introduction to biometrics,. Kluwer Academic Press, 1999.
- [JBS05] E. Justino, F. Bortolozzi, and R. Sabourin. A comparison of SVM and HMM classifiers in the offline signature verification. *Pattern Recognit. Letters*, 26(9):1377–1385, July 2005.
- [JC09] Jun-Wen Ji and Xiao-Su Chen. Off-line Chinese signature verification segmentation and feature extraction. In *Proc. of Int. Conf. on Comput. Intell. and Software Engineering*, pages 1–4, 2009.
- [JFOP12] J.Galbally, J. Fierrez, J. Ortega-Garcia, and R. Plamondon. Synthetic on-line signature generation. part ii: Experimental validation. *Pattern Recognition*, 45:2622–2632, July 2012.

- [JGC02] A.K. Jain, F.D. Griess, and S.D. Connell. On-line signature verification. *Pattern Recognit.*, 35:2963–2972, 2002.
- [JPFO12] J. Galbally, R. Plamondon, J. Fierrez, and J. Ortega-Garcia. Synthetic on-line signature generation. part i: Methodology and algorithms. *Pattern Recognition*, 45:2610–2621, July 2012.
- [JS01] E. Justino and F. Bortolozzi and R. Sabourin. Off-line signature verification using HMM for random, simple and skilled forgeries. In *Proc. of Int. Conf. on Doc. Anal. and Recognit.*, pages 105–110, Seattle, USA, 2001.
- [JYBS00] E. Justino, A. El Yacoubi, F. Bortolozzi, and R. Sabourin. An off-line signature verification system using HMM and graphometric features. In *Proc. of 4th IAPR Int. Workshop on Doc. Anal. Syst.*, pages 211–222, Rio de Janeiro, Brazil, 2000.
- [KD12] B. Kar and P.K. Dutta. SVM based signature verification by fusing global and functional features. *International Journal of Computer Applications*, 60(16):34–39, 2012.
- [KFGMD13] Ram P. Krish, Julian Fierrez, Javier Galbally, and Marcos Martinez-Diaz. Dynamic signature verification on smart phones. In *Proc. Workshop on User-Centric Technologies and Applications, PAAMS*, May 2013.
- [KHDM98] J. Kittler, M. Hatef, R.P. Duin, and J. Matas. On combining classifiers. *IEEE Trans. on Pattern Anal. and Mach. Intell.*, 20(3):226–239, March 1998.
- [KHH03] H.Y. Kwon, E.Y. Ha, and H.Y. Hwang. Segmentation based on dynamic features for on-line signature verification. In *Int. 12th Turkish Symp. Artif. Intell. Neural Netw. (TAINN)*, pages 371–384, 2003.
- [KJZ13] N.D. Kapale, M.A. Joshi, and R.G. Zope. Robust identity verification using signature of a person. *International Journal of Computer Applications*, pages 17–24, 2013.
- [KP10] A. Kumar and A. Passi. Comparison and combination of iris matchers for reliable personal authentication. *Pattern Recognit.*, 43:1016–1026, 2010.

- [KRD05] H. Ketabdar, J. Richiardi, and A. Drygajlo. Global feature selection for on-line signature verification. In *Proceedings of 2005 Conference International Graphonomics Society*, pages 625–629, 2005.
- [KSX04] M.K. Kalera, S. Srihari, and A. Xu. Offline signature verification and identification using distance statistics. *Int. J. Pattern Recognit. Artif. Intell. (IJPRAI)*, 18(6):1339–1360, 2004.
- [KTN96] R.S. Kashi, W. Turin, and W. Nelson. On-line signature verification using stroke direction coding. *Optical Engineering Journal*, 35(9), 1996.
- [KY05a] A. Kholmatov and B. Yanikoglu. Identity authentication using improved online signature verification method. *Pattern Recognit. Letters*, 26:2400–2408, 2005.
- [KY05b] A. Kholmatov and B.A. Yanikoglu. Identity authentication using improved online signature verification method. *Pattern Recognition Letters*, 26:2400–2408, 2005.
- [KY09] A. Kholmatov and B. Yanikoglu. Susig: an on-line signature database, associated protocols and benchmark results. *Pattern Anal. and Applications*, 12:227–236, 2009.
- [Lak13] L. Lakshmanan. Efficient person authentication based on multi-level fusion of ear scores. *IET Biomet.*, 2(3):97–106, 2013.
- [LB84] C.J. Stone R.A. Olshen L. Breiman, J. Friedman. *Classification and Regression Trees*. Wadsworth International Group, Michigan, 1984.
- [LBA96] L.L. Lee, T. Berger, and E. Aviczer. Reliable on-line human signature verification systems. *IEEE Trans. on Pattern Anal. and Mach. Intell.*, 18(6):643–647, 1996.
- [Lee95] L.L. Lee. Neural approaches for human signature verification. In *Proc. 3rd Int. Conf. Doc. Anal. Recognit.*, pages 1055–1058, Montreal, Canada, Aug. 1995.
- [LG05] H. Lei and V. Govindaraju. A comparative study on the consistency of features in on-line signature verification. *Pattern Recognit. Letters*, 26:2483–2489, 2005.

- [LMA⁺12] M. Liwicki, M.I. Malik, L. Alewijnse, E. van den Heuvel, and B. Found. ICFHR2012 competition on automatic forensic signature verification (4NsigComp 2012). In *Proc. of 13th Int. Conf. on Frontiers in Handwriting Recognit.*, pages 373–378, Bari, Italy, September 2012.
- [LMd⁺11] M. Liwicki, M.I. Malik, C.E. den Heuvel, X. Chen, C. Berger, R. Stoel, M. Blumenstein, and B. Found. Signature verification competition for online and offline skilled forgeries (SigComp2011). In *Proc. of 11th Int. Conf. on Doc. Anal. and Recognit.*, Beijing, China, September 19-21 2011.
- [LP94] F. Leclerc and R. Plamondon. Automatic signature verification: The state of the art–1989-1993. *Int. J. on Pattern Recognit. and Artif. Intell.*, 8(3):643–660, 1994.
- [LvdHFM10] M. Liwicki, C.E. van den Heuvel, B. Found, and M.I. Malik. Forensic signature verification competition 4NSigComp2010 - detection of simulated and disguised signatures. In *Proc. of International Conference on Frontiers in Handwriting Recognition (ICFHR)*, pages 715–720, Kolkata, India, 2010.
- [LW02] A. Liaw and M. Wiener. Classification and regression by randomforest. *R News*, 2:18–22, 2002.
- [LWWZ05] H. Lv, W. Wang, C. Wang, and Q. Zhuo. Off-line chinese signature verification based on support vector machines. *Pattern Recognit. Letters*, 26:2390–2399, 2005.
- [MALD13] M.I. Malik, S. Ahmed, M. Liwicki, and A. Dengel. FREAK for real time forensic signature verification. In *Proc. of Int. Conf. on Doc. Anal. and Recognit.*, pages 971–975, Washington DC, USA, 2013.
- [MC96] R. Martens and L. Claesen. On-line signature verification by dynamic time-warping. In *Proceedings of the 13th International Conference on Pattern Recognition*, pages 38–42, Vienna, Aug. 1996.
- [MDFKG14] M. Martinez-Diaz, J. Fierrez, R. P. Krish, and J. Galbally. Mobile signature verification: Feature robustness and performance comparison. *IET Biometrics*, 2014.

- [MDFOG14] Marcos Martinez-Diaz, Julian Fierrez, and Javier Ortega-Garcia. *Automatic Signature Verification on Handheld Devices*, pages 321–338. IGI Global, May 2014.
- [MDK+97] A. Martin, G. Doddington, T. Kamm, M. Ordowski, and M. Przybocki. The DET curve in assessment of detection task performance. In *Proc. of 5th European Conf. on Speech Communication and Technology*, pages 1895–1898, 1997.
- [ML12] M.I. Malik and M. Liwick. From terminology to evaluation: Performance assessment of automatic signature verification systems. In *Proc. of International Conference on Frontiers in Handwriting Recognition*, pages 613–618, Bari, Italy, 2012.
- [MLA+13a] M.I. Malik, M. Liwicki, L. Alewijnse, W. Ohyama, M. Blumenstein, and B. Found. ICDAR 2013 competitions on signature verification and writer identification for on- and offline skilled forgeries (SigWiComp 2013). In *12th International Conference on Document Analysis and Recognition*, pages 1477–1483, 2013.
- [MLA+13b] M.I. Malik, M. Liwicki, L. Alewijnse, W. Ohyama, M. Blumenstein, and B. Found. ICDAR2013 competitions on signature verification and writer identification for on- and offline skilled forgeries (SigWiComp2013). In *Proc. of 13th Int. Conf. on Doc. Anal. and Recognit.*, Washington DC, USA, September 25-28 2013.
- [MLD11] M.I. Malik, M. Liwicki, and A. Dengel. Evaluation of local and global features for offline signature verification. In *Proceedings of the 1st International Workshop on Automated Forensic Handwriting Analysis*, pages 31–35, Beijing, China, Sept. 2011.
- [MLD13a] M.I. Malik, M. Liwicki, and A. Dengel. Local features for forensic signature verification. In A. Petrosino, L. Madalena, and P. Pala, editors, *New Trends in Image Analysis and Processing (ICIAP2013)*, volume 8158 of *LNCS*, pages 103–111. Springer, 2013.
- [MLD13b] M.I. Malik, M. Liwicki, and A. Dengel. Part-based automatic system in comparison to human experts for forensic signature verification. In *Proc. of Int. Conf. on Doc.*

- Anal. and Recognit.*, pages 872–876, Washington DC, USA, 2013.
- [MLDF14] M.I. Malik, M. Liwicki, A. Dengel, and B. Found. Man vs. machine: A comparative analysis for signature verification. *Int. J. Forensic Document Examination*, 24, 2014.
- [MM07] D. Maramatsu and T. Matsumoto. Effectiveness of pen pressure, azimuth, and altitude features for online signature verification. In *Proc. of Int. Conf. on Biomet.*, pages 503–512, 2007.
- [MM14] R.P. Krish Javier Galbally Marcos Martinez-Diaz, Julian Fierrez. Mobile signature verification: feature robustness and performance comparison. *IET Biometrics*, pages 1–11, 2014.
- [MMLL07] O. Miguel-Hurtado, L. Mengibar-Pozo, M.G. Lorenz, and J. Liu-Jimenez. On-line signature verification by dynamic time warping and gaussian mixture models. In *41st Annual IEEE International Carnahan Conference on Security Technology*, pages 23–29, Ottawa, Oct. 2007.
- [MMO07] M. R. Freire M. Martinez-Diaz, J. Fierrez and J. Ortega-Garcia. On the effects of sampling rate and interpolation in HMM-based dynamic signature verification. In *Proc. of 9th Int. Conf. on Doc. Anal. and Recognit.*, pages 1113–1117, Curitiba, Brazil, 2007.
- [NP94] F. Nouboud and R. Plamondon. Global parameters and curves for offline signature verification. In *Proc. 4th Int. Workshop Front. Handwriting Recognit. (IWFHR-4)*, pages 145–155, Taipei, Taiwan, 1994.
- [NR77] R.N. Nagel and A. Rosenfeld. Computer detection of free-hand forgeries. *IEEE Trans. Computers*, 26(9):895–905, 1977.
- [OJFS05] L.S. Oliveira, E. Justino, C. Freitas, and R. Sabourin. The graphology applied to signature verification. In *Proc. of the 12th Conf. of the Int. Graphonomics Soc.*, 2005.
- [OSK05] E. Özgündüz, T. Sentürk, and M. E. Karşligil. Off-line signature verification and recognition by support vector machine. In *European Signal Processing Conf.*, Turkey, 2005.

- [Ots94] N. Otsu. A threshold selection method from grey level histogram. *IEEE Transactions on Systems, Man and Cybernetics*, 9:62–66, 1994.
- [PBP11a] S. Pal, M. Blumenstein, and U. Pal. Off-line signature verification systems: a survey. In *Proc. of ICWET'11 Int. Conf. & Workshop on Emerging Trends in Technol.*, Mumbai, Maharashtra, India, February 2011.
- [PBP11b] S. Pal, M. Blumenstein, and Umapada Pal. Non-English and Non-Latin signature verification systems: A survey. In *Proc. of 1st Int. Workshop on Automated Forensic Handwriting Anal.*, Beijing, China, 2011.
- [PFM13] A. Parziale, S.G. Fuschetto, and A. Marcelli. Exploiting stability regions for online signature verification. In A. Petrosino, L. Maddalena, and P. Pala, editors, *New Trends in Image Analysis and Processing (ICIAP2013)*, volume 8158 of *LNCS*, pages 112–121. Springer, 2013.
- [PFV11] J.M. Pascual-Gaspar, M. Faundez-Zanuy, and C. Vivaracho. Fast on-line signature recognition based on vq with time modeling. *Eng. Appl. of Artific. Intell.*, 24(2):368–377, 2011.
- [PG10a] M. Parodi and J.C. Gómez. A circular grid-based feature extraction approach for off-line signature verification. In *Proceedings of the XI Argentine Symposium on Artificial Intelligence (ASAI2010)*, pages 163–174, CABA, Argentina, 2010.
- [PG10b] M. Parodi and J.C. Gómez. Off-line signature verification: A circular grid-based feature extraction approach. In *Proc. 36th Latinamerican Conf. on Informatics (CLEI 2010)*, Asuncion, Paraguay, October 2010.
- [PG12] M. Parodi and J.C. Gómez. Online signature verification based on legendre series representation. consistency analysis of different feature combinations. In L. Alvarez et al., editor, *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications (CIARP 2012)*, volume 7441 of *LNCS*, pages 715–723. Springer, 2012.
- [PG13] M. Parodi and J.C. Gómez. Online signature verification: Improving performance through pre-classification based

- on global features. In A. Petrosino, L. Maddalena, and P. Pala, editors, *New Trends in Image Anal. and Processing (ICIAP2013)*, volume 8158 of *LNCS*, pages 69–76. Springer, 2013.
- [PG14a] M. Parodi and J.c. Gomez. *Emerging Aspects in Handwritten Signature Verification*, chapter Using Global Features for Pre-classification in Online Signature Verification Systems, pages 39–52. World Scientific, 2014.
- [PG14b] M. Parodi and J.C. Gómez. Legendre polynomials based feature extraction for online signature verification. consistency analysis of feature combinations. *Pattern Recognition*, 47(1):128–140, 2014.
- [PGA14] M. Parodi, J.C. Gómez, and L. Alewijnse. Automatic online signature verification based only on fhe features: an oxymoron? In *Proc. of 14th Int. Conf. on Frontiers in Handwriting Recognit.*, Creta Island, Greece, September 2014.
- [PGAL14] M. Parodi, J.C. Gómez, L. Alewijnse, and M. Liwicki. On-line signature verification: Automatic feature selection vs. fhe’s choice. *Journal of Forensic Document Examination*, 2014.
- [PGB11] M. Parodi, J.C. Gómez, and Abdel Belaïd. A circular grid-based rotation invariant feature extraction approach for off-line signature verification. In *Proc. of the Int. Conf. on Doc. Anal. and Recognit.*, pages 1289–1293, Beijing, China, Sept. 2011.
- [PGL12] M. Parodi, J.C. Gómez, and M. Liwicki. Online signature verification based on Legendre series representation. robustness assessment of different feature combinations. In *Proc. of 13th Int. Conf. on Frontiers in Handwriting Recognit.*, pages 379–384, Bari, Italy, September 2012.
- [PGLA13a] M. Parodi, J.C. Gómez, M. Liwicki, and L. Alewijnse. Orthogonal function representation for online signature verification: which features should be looked at? *IET Biomet.*, 2(4):137–150, 2013.
- [PGLA13b] M. Parodi, J.C. Gomez, M. Liwicki, and L. Alewijnse. Wavelet based representation for online signature verification. automatic feature selection vs. fhe’s choice. In

- Proc. 16th Int. Graphonomics Society Conf. (IGS2013)*, pages 1–6, Nara, Japan, June 2013.
- [PIP⁺13] G. Pirlo, D. Impedovo, R. Plamondon, C. O'Reilly, A. Cozzolongo, R. Gravinese, and A. Rollo. Stability of dynamic signatures: From the representation to the generation domain. In A. Petrosino, L. Maddalena, and P. Pala, editors, *New Trends in Image Analysis and Processing (ICIAP2013)*, volume 8158 of *LNCS*, pages 122–130. Springer, 2013.
- [PL89] R. Plamondon and G. Lorette. Automatic signature verification and writer identification: The state of the art. *Pattern Recognit.*, 22(2):107–131, 1989.
- [PL07] V. Pervouchine and G. Leedham. Extraction and analysis of forensic document examiner features used for writer identification. *Pattern Recognition*, 40(3):1004–1013, 2007.
- [Pla94] R. Plamondon. The design of an on-line signature verification system: from theory to practice. *Pattern Recognit. and Artif. Intell.*, 8(3):795–811, 1994.
- [PLD14] T-A. Pham, H-H. Le, and N-T. Do. Offline handwritten signature verification using local and global features. *Annals of Mathematics and Artificial Intelligence*, 2014.
- [PS00] R. Plamondon and S.N. Srihari. On-line and off-line handwriting recognition: A comprehensive survey. *IEEE Trans. on Pattern Anal. and Mach. Intell.*, 22(1):63–84, 2000.
- [RJI12] J. Rico-Juan and J. Iñesta. Confidence voting method ensemble applied to off-line signature verification. *Pattern Anal. and Applications*, 15:113–120, 2012.
- [RKD05] J. Richiardi, H. Ketabdar, and A. Drygajlo. Local and global feature selection for on-line signature verification. In *Proc. of 8th Int. Conf. on Doc. Anal. and Recognit.*, Seoul, Korea, 2005.
- [SBOJ07] C.R. Santos, F. Bortolozzi, L.S. Oliveira, and E. Justino. Off-line signature verification based on forensic questioned document examination approach. In *Proc. of Sym-*

- posium on Applied Computing*, Seoul, Korea, March 11-15 2007.
- [SD92] R. Sabourin and J.P. Drouhard. Offline signature verification using directional PDF and neural networks. In *Proc. 11th Int. Conf. Pattern Recog., Conf. B: Pattern Recog. Methodol. Syst.*, pages 321–325, The Hague, The Netherlands, 1992.
- [SGGR05] S.Canu, Y. Grandvalet, V. Guigue, and A. Rakotomamonjy. SVM and kernel methods Matlab toolbox. Technical report, Perception Systmes et Information, INSA de Rouen, Rouen, France, 2005.
- [Sha14] P. Sharma. Offline signature verification using surf feature extraction and neural networks approach. *International Journal of Computer Science and Information Technologies*, 5(3):3539–3541, 2014.
- [Sig11] SigComp2011. ICDAR 2011 Signature verification competition. http://www.iapr-tc11.org/mediawiki/index.php/Datasets_List, 2011.
- [SJBS04] C. Santos, E. Justino, F. Bortolozzi, and R. Sabourin. An off-line signature verification method based on the questioned document experts approach and a neural network classifier. In *Proc. of the 9th Int. Workshop on Frontiers in Handwriting Recognition (IWFHR-9)*, 2004.
- [SP86] R. Sabourin and R. Plamondon. Preprocessing of handwritten signatures from image gradient analysis. In *Proc. 8th Int. Conf. Pattern Recognition*, pages 576–579, 1986.
- [SPB94] R. Sabourin, R. Plamondon, and L. Beaumier. Structural interpretation of handwritten signature images. *Int. J. Pattern Recognit. Artif. Intell. (IJPRAI)*, pages 69–108, 1994.
- [SS01] B. Schölkopf and A.J. Smola. *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*. MIT Press, 2001.
- [SS10] N. Shukla and M. Shandilya. Invariant features comparison in hidden markov model and SIFT for offline handwritten signature database. *International Journal of Computer Applications*, 2(7):31–43, 2010.

- [SSR10] M. Trivedi S. Shivappa and B. Rao. Audiovisual information fusion in humancomputer interfaces and intelligent environments: A survey. *Proceedings of the IEEE*, 98(10):1692–1715, 2010.
- [Ste75] J. Sternberg. Automated signature verification using handwriting pressure. *1975 WESCON Technical Papers*, 31(4):1–6, 1975.
- [TAK⁺00] P. Tuyls, A.H.M. Akkermans, T.A.M. Kevenaar, G.-J. Schrijen, A.M. Bazen, and R.N.J. Veldhuis. Practical biomet. authentication with template protection. pages 45–47, Barcelona, Spain, 2000.
- [Ued03] Katsuhiko Ueda. Investigation of off-line Japanese signature verification using a pattern matching. In *Proc. of 7th Int. Conf. of Doc. Anal. and Recognit.*, Edinburgh, Scotland, 2003.
- [Vap95] V. Vapnik. *The Nature of Statistical Learning Theory*. Springer-Verlag, NY, 1995.
- [Vap98] V. Vapnik. *Statistical Learning Theory*. Wiley, 1998.
- [VFTA07] F. Vargas, M.A. Ferrer, C.M. Travieso, and J.B. Alonso. Off-line handwritten SignatureGPDS-960 Corpus. In *Proc. of 9th Int. Conf. on Doc. Anal. and Recognit.*, pages 764–768, Curitiba, Brazil, 2007.
- [VFTA08] F. Vargas, M.A. Ferrer, C.M. Travieso, and J.B. Alonso. Off-line signature verification based on high pressure polar distribution. In *11th Int. Conf. on Front. in Handwriting Recognit. (ICFHR08)*, 2008.
- [WFTZ07a] J. Wen, B. Fang, Y. Tang, and T. Zhang. Offline signature verification:a new rotation invariant approach. In *Proc. IEEE International Conference on Systems, Man and Cybernetics (ISIC 2007)*, pages 3583–3587, 2007.
- [WFTZ07b] J. Wen, B. Fang, Y.Y. Tang, and T.P. Zhang. Offline signature verification:a new rotation invariant approach. In *IEEE Int. Conf. on Syst., Man and Cybernetics (ISIC 2007)*, pages 3583–3587, 2007.
- [Wil12] E.J. Will. Inferring relative speed of handwriting from the static trace. *J. of Forensic Doc. Examination*, 22, 2012.

- [WRV04] M. Wirotius, J.Y. Ramel, and N. Vincent. Improving DTW for online handwritten signature verification. In *Image Analysis and Recognition*, volume 3212 of *LNCS*, pages 786–793. Springer, 2004.
- [XVK⁺08] H. Xu, R.N.J. Veldhuis, T.A.M. Kevenaar, A.H.M. Akkermans, and A.M. Bazen. Spectral minutiae: A fixed-length representation of a minutiae set. In *Proc. of IEEE Computational Soc. Conf. on Computer Vision and Pattern Recognit. Workshops*, 2008.
- [YCX⁺04] D.-Y. Yeung, H. Chang, Y. Xiong, S. George, R. Kashi, T. Matsumoto, and G. Rigoll. SVC2004: First int. signature verification competition. In *Proc. of Int. Conf. on Biomet. Authentication*, pages 16–22, Hong Kong, 2004.
- [YK09] B. Yanikoglu and A. Kholmatov. Online signature verification using fourier descriptors. *EURASIP J. on Advances in Signal Processing*, pages 230–275, 2009.
- [YO77] Makoto Yasuhara and Masatomo Oka. Signature verification experiment based on nonlinear time alignment: A feasibility study. *IEEE Transactions on Systems, Man, and Cybernetics*, pages 212–216, 1977.
- [YWP95] L. Yang, B.K. Widjaja, and R. Prasad. Application of Hidden Markov Models for signature verification. *Pattern Recognit.*, 28(2):161–170, February 1995.
- [YY97] M. Yoshimura and I. Yoshimura. Investigation of a verification system for japanese countersignatures on travelers checks. *Trans. IEICE*, J80-D-II(7):1764–1773, October 1997.
- [YY98] M. Yoshimura and I. Yoshimura. Investigation of a verification system for Japanese countersignatures on traveler’s cheques. *Trans. of the IEICE*, J80-D-II(7):1764–1773, 1998.
- [ZCAM12] J. Zhou, S. Cadavid, and M. Abdel-Mottaleb. An efficient 3-D ear recognition system employing local and holistic features. *IEEE Trans. on Information Forensics and Security*, 7(3):978–991, 2012.
- [ZNA05] E.N. Zois, A.A. Nassiopoulos, and V. Anastassopoulos. Signature verification based on line directionality. In *Proc.*

2005 IEEE Workshop on Signal Processing Systems Design and Implementation, pages 343–346, 2005.

- [ZTLL03] M. Zou, J. Tong, C. Liu, and Z. Lou. On-line signature verification using local shape analysis. In *Proc. of 7th Int. Conf. of Doc. Anal. and Recognit.*, Edinburgh, Scotland, UK, 2003.
- [ZWW11] Z. Zhang, K. Wang, and Y. Wang. A survey of on-line signature verification. In Z. Sun, J. Lai, X. Chen, and T. Tan, editors, *Biometric Recognition*, volume 7098 of *LNCS*, pages 141–149. Springer, 2011.