Universidad
Carlos III de Madrid

TESIS DOCTORAL

Online Signature Verification Algorithms and Development of Signature International Standards

Autor:
Óscar Miguel Hurtado

Director/es:
Raúl Sánchez Reíllo
Luis Mengibar Pozo

Departamento de Tecnología Electrónica

Leganés, Septiembre 2011

TESIS DOCTORAL

Online Signature Verification Algorithms and Development of Signature International Standards

Autor: **Óscar Miguel Hurtado**

Director/es: **Raúl Sánchez Reíllo, Luis Mengibar Pozo**

Firma del Tribunal Calificador:

Firma

Presidente: **Marcos Faúndez Zanuy**

Vocal: **Richard Guest**

Vocal: **Olaf Henniger**

Vocal: **Enrique Cabello Pardos**

Secretaria: **Judith Liu Jimenez**

Calificación:

Leganés, de de

ABSTRACT

The science of biometrics is based on discovering the identities of human beings by investigating their physical and behavioural traits. Of the many different biometric traits, i.e. fingerprint, iris, vascular, etc... the handwritten signature is still one of the most accepted techniques. Advancing progress in identification applications has led to widespread demand for new generation ID documents, such as electronic passports and citizen cards, which contain additional biometric information required for more accurate user recognition. This can be achieved by embedding dynamic signature features within the documentation. However, this would result in two significant drawbacks that must be addressed, these are: Memory Capacity and Computational Load. These problems and the increasing demand for standardized biometric verifications systems have motivated the research work performed in this Thesis.

In order to achieve this, an attempt to reduce the information involved in verification processes is performed using feature selection criteria of the signature biometric data. Such reduced information content not only satisfies the memory capacity restrictions but also provides much more efficient use of the verification algorithms. In particular, two novel methods in the signature context, based on Principal Component Analysis and Hellinger Distance, are proposed here. The performance of the optimized features set obtained has been analyzed using two different verification algorithms. By reducing the sample size it has been observed that the error rates are maintained sufficiently low and the results obtained are in agreement with the current state of the art for signature techniques. It will be shown that in some cases that feature selection does not provide an adequate reduction solution, where a different strategy has been analyzed to achieve the aforementioned problems.

A direct consequence of the widespread nature of biometric verification has led to demands for standardized protocols to improve interoperability. The work presented throughout this Thesis has considered current ISO/IEC signature standard data formats. It has been observed that the current compact data formats, 19794-7 Compact Format and 19794-11, do not meet the requirements of modern data formats. In particular, 19794-7 Compact Format, although having good compression ratios, has been found to imply an inadmissible loss in information. This problem has been solved by defining a new near-lossless compression data format based on lossless compression algorithms, and proposing different enhanced strategies to store signature data. This new data format achieves the same compression ratio, but without losing any relevant information. In addition, the problems found in the 19794-11CD2 regarding the lack of compression and information loss have been addressed. A new data format structure has been proposed, where the lack of compression is solved by reducing the data stored, avoiding duplicated data and providing a new singular point definition. This new structure has provided improved compression ratios, and, at the same time, carries more information. The two new data format definitions were presented to the ISO/IEC SC37 WG3 experts and accepted as the new third subformat “Compression Format” within the 19794-7 and the new committee draft for the 197974-11 CD3.

RESUMEN

En la sociedad actual existe la necesidad de verificar la identidad de usuarios de una manera automática y segura, sobre todo teniendo en cuenta las nuevas posibilidades que el comercio electrónico ha originado. Desgraciadamente todas estas nuevas posibilidades electrónicas de acceso a distintos servicios, también han incrementado las probabilidades de actividades delictivas como la usurpación de identidad.

La biometría ha demostrado ser una tecnología válida para la verificación de identidades, ya que ofrece un alto nivel de seguridad a la vez que resulta cómoda al usuario. De hecho su uso ya ha sido probado con éxito para tales fines en distintos contextos, siendo uno de los más comunes y conocidos su aplicación en la nueva generación de documentos de identidad electrónicos, tales como el Documento Nacional de Identidad Electrónico (DNIe) así como en los nuevos pasaportes electrónicos. Estas nuevas generaciones de documentos de identidad incorporan técnicas biométricas que permiten a los usuarios la autenticación de su identidad en procesos remotos.

Junto con estas ventajas de la tecnología biométrica, la capacidad de almacenamiento y procesado de datos por parte de los nuevos documentos de identidad hace posible la incorporación de la información dinámica que posee la firma manuscrita. Esta información puede ser utilizada para la verificación de la identidad de los usuarios de una manera muy familiar, ya que el uso de la firma manuscrita para la verificación de identidades está muy extendido. No obstante, a la hora de incluir esta información dentro de este tipo de dispositivos, se deben tener en cuenta dos limitaciones significativas. En primer lugar, hay que examinar las necesidades de almacenamiento indispensables para guardar los datos obtenidos de la firma manuscrita capturada así como para el patrón del usuario. En segundo lugar, hay que considerarla baja potencia de cálculo de estos dispositivos a la hora de desarrollar algoritmos de verificación. Del mismo modo, se debe tener en cuenta que los documentos de identidad se diseñan para ser usados en una gran variedad de escenarios, tanto a nivel nacional como internacional. Por esta razón el uso de normas internacionales que garanticen su interoperabilidad se hace un requisito indispensable.

Partiendo de lo expuesto anteriormente, la presente Tesis Doctoral se ha centrado en mejorar la viabilidad de sistemas automáticos de verificación de firma dinámica manuscrita en entornos con fuertes limitaciones tanto en capacidad de almacenamiento de datos como en capacidad de computación. A su vez, se ha llevado a cabo un análisis exhaustivo de los actuales formatos de datos definidos en las norma internacional “19794 Biometric data interchange formats” existentes para firma manuscrita dinámica (parte 7 y 11 de esta norma), para contrastar como pueden llegar a afectar dichos formatos al rendimiento de los algoritmos de verificación.

Los aspectos anteriormente indicados sobre las necesidades de almacenamiento y de computación han sido abordados a través de técnicas de selección de características

probadas en dos implementaciones de algoritmos de verificación de firma basados en Modelado de Mezcla de Gausianas (designado por sus siglas en inglés “GMM”) y Alineamiento Dinámico Temporal (designado por sus siglas en inglés “DTW”). En concreto, las técnicas de selección de características empleadas han sido el Ratio de Fisher (cuyas siglas en inglés son FR), el Análisis de Componentes Principales (cuyas siglas en inglés son PCA), la combinación de ambas y por último, la distancia de Hellinger (cuyas siglas en inglés son HD). La primera de ellas es una técnica muy extendida en la literatura de firma manuscrita, mientras que las otros dos, PCA y HD, no se ha encontrado ninguna constancia de haber sido utilizada anteriormente en entornos de firma manuscrita. Los resultados han desvelado que la técnica PCA genera una selección de características más óptima que la técnica FR, mejorando las tasas de error de los algoritmos utilizados. Además, la combinación de esta técnica (PCA) con la técnica FR ha obtenido mejores resultados que aplicadas de manera individual. Por su parte, HD también ha demostrado su utilidad en el ámbito de la firma manuscrita dinámica, obteniendo mejores resultados que las técnicas expuestas anteriormente sobre todo en el caso del algoritmo DTW en el que el solapamiento de distribuciones de las características entre firmas genuinas y las firmas falsas es bajo.

A la vista de estos resultados, con las técnicas de selección de características propuestas se ha logrado cumplir con los objetivos de reducir las necesidades tanto de espacio de almacenamiento como de capacidad computacional, manteniendo tasas de error acordes con el estado del arte. Cabe destacar que para el algoritmo GMM desarrollado se han propuesto dos vectores de características, uno formado por 28 elementos y otro de tan solo 13 elementos para entornos con limitaciones más extremas. A su vez, el algoritmo GMM implementado también ha demostrado ser robusto frente al número de funciones Gausianas que lo forman, obteniendo resultados en línea con el estado del arte para combinaciones de sólo cuatro funciones Gausianas. Estos dos resultados (el bajo número de elementos en el vector de características y el bajo número de funciones Gausianas) conllevan que tanto el modelo de usuario, como las firmas capturadas, requieran un mínimo espacio de almacenamiento. Del mismo modo, hacen que la carga computacional sea mucho menor que la de los algoritmos basados en GMM publicados con anterioridad.

Con respecto al algoritmo DTW planteado, se ha propuesto un vector de características formado tan solo por seis elementos, obteniendo de nuevo bajas tasas de error tanto para falsificaciones aleatorias, como, especialmente, para falsificaciones entrenadas. Estos resultados una vez más muestran que las técnicas de selección de características han respondido satisfactoriamente. Pero a pesar de que el número de elementos del vector de características es muy bajo, no se han podido reducir las necesidades ni de espacio, ni de complejidad de cálculo, dado que para el algoritmo DTW todavía se incluye información de la presión. Sin embargo, estos objetivos han sido cubiertos mediante el análisis efectuado en relación con el número de puntos que se requieren para el almacenamiento tanto de las firmas capturas como para el del patrón de usuario. Las pruebas realizadas han puesto de manifiesto que submuestreando las firmas capturadas de manera que estén formadas sólo por 256 puntos, es suficiente para asegurar que los niveles de error obtenidos por los algoritmos se mantengan en niveles dentro del estado del arte de los algoritmos DTW.

Incluso, bajando el número de puntos hasta la cifra de 128 se ha visto que aún se consiguen tasas de error aceptables.

Además del estudio a nivel algorítmico de la viabilidad de implementación de algoritmos de firma manuscrita dinámica, esta Tesis Doctoral se ha también se ha enfocado en la mejora de las actuales normas internacionales de formato de datos existentes para firma manuscrita dinámica, teniendo por objetivo incrementar sus posibilidades de uso en dispositivos tales como documentos de identidad.

Inicialmente, se ha realizado un estudio de la viabilidad del uso de estas normas internacionales (proyectos 19794-7 y 19794-11 del subcomité SC37 dentro de la organización ISO/IEC) en cuanto a tamaño de la muestra examinando varias bases de datos públicas de firma dinámica. De este análisis se ha concluido que el formato compacto definido en el proyecto 19794-7 presenta un ratio de compresión del 56% comparado con el formato completo. Por otro lado, el proyecto 19794-11 que se definía como un formato de compresión de datos para firma manuscrita, presentó ratios de compresión negativos, indicando que en lugar de tener un menor tamaño de muestra, este formato incrementa el tamaño en comparación con las firmas almacenadas siguiendo el formato completo 19794-7. A su vez, se ha mostrado como la compresión de datos, tanto en el formato compacto 19794-7 como en el formato 19794-11, tiene un impacto en el rendimiento de los algoritmos, incrementando sus tasas de error. Esto es debido a la información que se pierde en el proceso de compresión de los datos.

Para resolver la perdida de rendimiento de los algoritmos cuando se usa el formato de datos compacto definido dentro del proyecto 19794-7, se han presentado dos nuevos formatos de datos. Estos formatos , denominados formatos de datos comprimidos, se basan en algoritmos de compresión de datos sin pérdida de información. Se ha llevado a cabo la evaluación de distintos algoritmos de estas características, así como distintas opciones de reordenación de los datos de la firma manuscrita para maximizar la compresión obtenida gracias a los algoritmos de compresión. Dentro de los formatos de datos sugeridos, se ha planteado un formato de datos comprimido que presenta los mismos ratios de compresión que el formato compacto 19794-7, pero sin incurrir en ninguna pérdida de datos, por lo que no presenta ningún impacto en las tasas de error de los algoritmos de verificación. Asimismo, también se ha propuesto un formato de datos comprimido con mínima perdida de información, mejorando las tasas de compresión, sin influir de nuevo en el rendimiento de los algoritmos de verificación. Este formato comprimido de datos con reducidas pérdidas tiene además la capacidad de ajustar el nivel de información perdida, lo que resulta una importante característica teniendo en cuenta las espectaculares resoluciones (tanto espaciales como temporales) que los dispositivos de captura presentan en la actualidad. Estas altas resoluciones conllevan un aumento importante en el tamaño de las muestras capturas, que puede ser minimizado con el uso de este formato comprimido con pérdidas.

Ambos formatos de datos comprimidos, con y sin perdidas, fueron presentados a la comunidad internacional dentro del subcomité ISO/IEC SC37, proponiendo su inclusión en el proyecto 19794-7. Esta petición fue aceptada por los expertos internacionales de firma

manuscrita, convirtiéndose el formato de datos comprimidos en el tercer subformato dentro de esta norma internacional. La publicación de esta norma con la inclusión de las contribuciones mencionadas está planificada para el año 2012.

Con respecto al proyecto 19794-11CD2, se analizó el uso de una nueva estructura de datos que solucionara los problemas de la falta de compresión a través de la eliminación de información duplicada, almacenando menos datos y redefiniendo los puntos singulares en los que está basada la segmentación. Además, para aumentar aún más las tasas de compresión obtenidas, diferentes estrategias de eliminación de puntos espurios fueron tratadas. A su vez, para mejorar la calidad de la información almacenada dentro de este formato de datos, se ha estudiado la posibilidad de recrear los datos contenidos en el formato completo partiendo de los datos almacenados en esta parte 19794-11. Mediante estos análisis, se han obtenido tasas de compresión menores que los presentados por el formato compacto 19794-7. Esta nueva definición para el proyecto 19794-11 también se presentó al subcomité SC37, siendo igualmente aceptada por los expertos internacionales en firma manuscrita y adoptada en la nueva revisión del proyecto 19794-11CD3. La publicación de este proyecto como norma internacional se espera para 2013.

CONTENT

Chapter 1 Introduction 1

1.1 Motivation **1**

1.2 Document Structure **3**

Chapter 2 Biometrics and International Biometric Standards 5

2.1 Biometric Technology **5**

2.1.1 Brief History of biometrics 6

2.1.2 Biometrics Modalities 7

2.1.3 General biometric system 8

2.2 Biometric System Evaluation **11**

2.2.1 Types of Biometric Performance Testing evaluations 12

2.2.2 Performance Measures: Error Rates 13

2.2.3 Graphical Presentation Of Results 15

2.3 International Standards **17**

2.3.1 International Standards And Organizations 17

2.3.2 ISO/IEC JTC1 18

2.3.3 ISO/IEC JTC1 SC37 Biometric Standards 20

2.3.4 Stages on the development of International Standards 21

2.3.5 ISO/IEC Project 19794 – Biometrics Data Interchange Format 23

2.4 Conclusions **27**

Chapter 3 Automatic Signature Verification 29

3.1 Introduction 29

3.1.1 Automatic Signature Verification: Online and Offline 31

3.1.2 Signature Forgeries 32

3.2 Capture Devices 33

3.3 Comparison Algorithms 36

3.3.1 Distance Based Approaches 36

3.3.2 Model Based Approaches 37

3.4 Evaluation Databases 39

3.4.1 MCyT Signature Subcorpus 39

3.4.2 SVC'2004 Development Set 40

3.4.3 MyIDea Signature Subcorpus 40

3.4.4 Databases Summary 41

3.5 Evaluation Campaigns 44

3.5.1 SVC 2004 44

3.5.2 BioSecure Reference Benchmarking Framework For Signature Verification 47

3.5.3 BioSecure Signature Evaluation Campaign 2009 (BSEC'2009) 49

3.6 Current Data Format Standards For Signature Biometrics 56

3.6.1 2nd Working Draft 19794, Biometric Data Interchange Formats – Part 7: Signature/Sign Time Series Data 56

3.6.2 2nd Committee Draft 19794-11, Biometric data interchange formats – Part 11: Signature/Sign Processed Dynamic Data 62

3.7 Conclusions 66

Chapter 4 Improvement in Automatic Signature Verification 67

4.1 Introduction 67

4.2 Signature Verification Algorithms 69

4.2.1 Gaussians Mixture Models 69

4.2.2 Dynamic Time Warping 78

4.3 Features Selection Techniques For Dimensionality Reduction 85

4.3.1 Introduction 85

4.3.2 Fisher Ratio 86

4.3.3 Principal Component Analysis For Feature Selection 87

4.3.4 Fisher Ratio Combined with Principal Component Analysis 87

4.3.5 Hellinger Distance 88

4.4 Feature Selection Applied to The GMM 89

4.4.1 Based on Identification Error Rates 89

4.4.2 Features Selection Based on Hellinger Distance 92

4.4.3 Verification Error Rates For the Selected Features Subsets 94

4.5 Feature Selection Applied to DTW 97

4.5.1 Based on Identification Error Rates 97

4.5.2 Features Selection Based on Hellinger Distance 98

4.5.3 Verification Error Rates For The Selected Features Subsets 100

4.6 Analysis of the GMM Algorithm Parameters 103

4.6.1 Number of Gaussians in the GMM model 103

4.6.2 Number of signatures Samples taken for the enrolment process 104

4.6.3 User Model Size 105

4.6.4 Computational Load 106

4.7 Analysis of the DTW Algorithm Parameters 108

4.7.1 Number Of Equi-Spaced Points Used 108

4.7.2 Number of signatures Samples taken for the enrolment process 110

4.7.3 User Model Size 111

4.7.4 Computational Load 112

4.8 Conclusions 114

Chapter 5 Viability Analysis of Signature Standard Data Formats 117

5.1 Introduction 117

5.2 BDIR Size Analysis 118

5.2.1 Implementation Details 119

5.2.2 Results Obtained 121

5.3 Algorithm Performance Achieved with Current Data Formats 124

5.4 Conclusions 127

Chapter 6 Interoperability of Signature Biometrics at Signal Level 129

6.1 Introduction 129

6.2 Compressed Data Format 131

6.3 Compression Algorithms Tested 136

6.3.1 7 Zip 136

6.3.2 GZip 136

6.3.3 BZip2 137

6.3.4 LZW06 137

6.4 BDIR Size Analysis 138

6.4.1 Results Obtained 138

6.5 Algorithm Performance Results 142

6.6 Conclusions 144

Chapter 7 Interoperability of Signature Biometrics at Data Processed Level 147

7.1 Introduction 147

7.2 Proposal for a New Data Format.....149

7.3 Proposal Analysis.....152

7.3.1 Data Format Versions Analyzed 152

7.3.2 Analysis of Pen-Stroke Data Options..... 155

7.4 Biometric Data Interchange Record Size For Signature/Sign Data Formats within the 19794-11.....157

7.5 Interpolation Error For Reconstructing 19794-7 from 19794-11 Data 162

7.5.1 Interpolation from 19794-11 CD2 and the New Proposal..... 163

7.5.2 PCHIP And Spline Study 165

7.5.3 Enhancing Interpolation 167

7.6 Conclusions.....170

Chapter 8 Conclusions and Future Work 171

8.1 Conclusions.....171

8.2 Future Work 175

Chapter 9 References 179

LIST OF TABLES

Table 2-1 ISO/IEC JTC1 Subcommittees 19
Table 2-2 ISO/IEC JTC1 SC37 Working Groups 20
Table 2-3 ISO/IEC Project Stages and associated documents [26] 22
Table 2-4 Publication Date for Biometric Data Interchange Format Standards, first generation 25
Table 2-5 Stage of Biometric Data Interchange Format Standards, second generation 26
Table 3-1 Active Impostor Attempts Levels for Biometric Signature Systems 32
Table 3-2 Summary of Signature Databases 41
Table 3-3 Capture Details for the Signature Databases 41
Table 3-4 Average Features of the Signature Datasets 42
Table 3-5 SVC2004 Participating Teams [22] 45
Table 3-6 SVC'2004 EER Results for Task1 46
Table 3-7 SVC'2004 EER Results for Task2 46
Table 3-8 Systems evaluated on BSEC'2009 [20] 47
Table 3-9 EERS of the systems on the MCyT-100 database and their Confidence Interval (CI) of 95% 48
Table 3-10 Systems evaluated on BSEC'2009 50
Table 3-11 EERS of the systems for Session 1 for the DS2 and DS3 datasets, for skilled and random forgeries 51
Table 3-12 EERS of the systems for Session 1 and 2 using the DS2 dataset, skilled and random forgeries 52
Table 3-13 EERS of the systems for Session 1 and 2 using the DS3 dataset, skilled and random forgeries 53
Table 3-14 EERS of the systems in Session 1 and 2 for the DS2 where pressure information is added, skilled and random forgeries 53
Table 3-15 EERS of the systems for Session 1 and 2 using the DS2 where pressure and tilts information is added, skilled and random forgeries 54
Table 3-16 EERS of the systems in Session 1 for DS2 and for each writer category, skilled and random forgeries 55
Table 4-1 On-Line Signature Verification Systems based on GMM 70
Table 4-2 Features analyzed from the signals captured directly using the tablet input device, GMM algorithm 75
Table 4-3 Features analyzed for the velocity and accelerations, GMM algorithm 76

Table 4-4 Global features analyzed, GMM algorithm 77
Table 4-5 On-Line Signature Verification Systems based on DTW 78
Table 4-6 Derived signals used for DTW algorithm 82
Table 4-7 Kogure and Sato features used for the DTW algorithm 84
Table 4-8 GMM 13 Features Vector Subset selected using FRPCA 91
Table 4-9 GMM 28 Features Vector Subset selected using FRPCA 91
Table 4-10 GMM 44 Features Vector Subset selected using FRPCA 92
Table 4-11 GMM 16 Features Vector Subset selected using HD 93
Table 4-12 GMM 28 Features Vector Subset selected using HD 93
Table 4-13 GMM 60 Features Vector Subset selected using HD 94
Table 4-14 Equal Error rates for skilled forgeries and different features subsets 95
Table 4-15 Equal Error rates for random forgeries and different feature subsets ... 96
Table 4-16 DTW 6 Features Vector Subset selected using FRPCA 98
Table 4-17 DTW 8 Features Vector Subset selected using FRPCA 98
Table 4-18 DTW 6 Features Vector Subset selected using HD 99
Table 4-19 DTW 11 Features Vector Subset selected using HD 100
Table 4-20 DTW 17 Features Vector Subset selected using HD 100
Table 4-21 DTW Equal Error rates for skilled forgeries and different features subsets 101
Table 4-22 DTW Equal Error rates for Random forgeries and different features subsets 102
Table 4-23 GMM equal error rates, number of Gaussians analysis 104
Table 4-24 GMM equal error rates, training samples analysis 104
Table 4-25 User model size for different feature subsets and different numbers of Gaussian functions 106
Table 4-26 Comparison time for different feature subsets and different number of Gaussian functions 107
Table 4-27 DTW equal error rates, number of Equi-Spaced points analysis 109
Table 4-28 DTW Equal Error rates, training samples analysis 110
Table 5-1 Subsets used for BDIR Size Study 118
Table 6-1 Compression Algorithms Tested 136
Table 7-1 Summary of New Part 11 Data Format Versions analyzed 154
Table 7-2 Allowed Values for type of dynamic event 156
Table 7-3 Allowed Values for the type of dynamic event for a pressure-stroke .. 156

LIST OF FIGURES

Figure 2-1 Biometric Market by Application, 2006 [3] 6
Figure 2-2 Components of a general biometric system [1] 9
Figure 2-3 Example of ROC curves for different biometric systems [18] 15
Figure 2-4 Example of DET curves for different biometric systems [18] 15
Figure 2-5 Example of CMC curves for different biometric systems [18] 16
Figure 2-6 ISO and IEC Logos 18
Figure 2-7 General interrelation model of biometric issues [28] 24
Figure 3-1 Signature Sets in a Disputed Will Case around 1900 [29] 30
Figure 3-2 Biometric Smart Pen [43] 33
Figure 3-3 Accelerometer Pen [45] 33
Figure 3-4 IMUPEN [46] 33
Figure 3-5 Genius Tablet [47] 34
Figure 3-6 Wacom Tablets [48] 34
Figure 3-7 Signals acquired by digital tablets 34
Figure 3-8 Wacom Singature Input Devices [48] 35
Figure 3-9 Touch-sensitive screen devices than can be used for signature acquisition 35
Figure 3-10 Average Dataset Lengths 43
Figure 3-11 Average Dataset Total Times 43
Figure 3-12 Average Dataset Velocity 43
Figure 3-13 Average Dataset Number of Strokes 43
Figure 3-14 Structure of a multiple representation BDIR defined in 19794-1.2FDIS. 57
Figure 3-15 BDB General Header for 19794-7.2WD2 Full Format 58
Figure 3-16 BDB Body of 19794-7.2WD2 Full Format 59
Figure 3-17 BDB Representation Header of 19794-7.2WD2 Full Format 59
Figure 3-18 Channel descriptions of 19794-7.2WD2 Full Format 60
Figure 3-19 BDB Representation Body of 19794-7.2WD2 Full Format 60
Figure 3-20 Sequence of Sample Points of 19794-7.2WD2 Full Format 61
Figure 3-21 BDB of 19794-7.2WD2 Compact Format 61
Figure 3-22 BDB Header of 19794-11CD2 62
Figure 3-23 BDB Body of 19794-11CD2 63
Figure 3-24 Signature Representation of 19794-11CD2 63
Figure 3-25 Pen-Stroke defined in 19794-11CD2 64

Figure 3-26 Pressure-Stroke defined in 19794-11CD2 64
Figure 3-27 Overall Data defined in 19794-11CD2 65
Figure 4-1 Gaussians Mixture Model Representation 70
Figure 4-2 Mixture of Gaussian Probabilistic functions 70
Figure 4-3 Time Alignment between two sequences 79
Figure 4-4 Warping Path aligning the pattern and the sample 80
Figure 4-5 Example of feature distribution. Dark line shows the genuine distribution, dot line shows the forgery distribution 88
Figure 4-6 Identification Error Rates vs. Number of Features 90
Figure 4-7 Identification Error Rates vs. Number of Features, Fisher Ratio combined with PCA Analysis, Detail 90
Figure 4-8 Overlap for GMM features analyzed 92
Figure 4-9 GMM Signature Verification System Results for Skilled Forgeries, length analysis of the feature vector 95
Figure 4-10 GMM Signature Verification System Results for Random Forgeries, length analysis of the feature vector 96
Figure 4-11 Identification Error Rates vs. Number of Features for the DTW Algorithm 97
Figure 4-12 Common Distribution Area for Pseudo-distances 99
Figure 4-13 DTW Signature Verification Results for Skilled Forgeries, different feature vector Analysis 101
Figure 4-14 DTW Signature Verification Results for Random Forgeries, feature vector Analysis 101
Figure 4-15 GMM Signature Verification System Results, number of Gaussians Analysis 103
Figure 4-16 GMM Signature Verification System Results, Training Samples Analysis 105
Figure 4-17 DTW error rates, number of Equi-Spaced points analysis 109
Figure 4-18 DTW Equal Error rates, training samples analysis 110
Figure 4-19 User's model size for DTW algorithm and different number of equi-spaced points 111
Figure 4-20 Comparison time for DTW algorithm and different number of equi-spaced points 112
Figure 5-1 BDIR Sizes for different Signatures Data format Standards 121
Figure 5-2 Compression Ratios for different Signatures Data format Standards 121
Figure 5-3 Average Signature Points for Different Datasets 122
Figure 5-4 Average Signature Total Time for Different Datasets 122
Figure 5-5 Average Signature Strokes for Different Datasets 122

LIST OF FIGURES

Figure 5-6 Average Pen and Pressure Strokes for different Signatures Data Format Standards 123
Figure 5-7 Error Rates for GMM and Different Data Formats 125
Figure 5-8 Error Rates for DTW and Different Data Formats 126
Figure 6-1 BDB Body Data Compression in Compressed Data Format 131
Figure 6-2 Sample Point Values order for Compression Data Format Version 1 ... 131
Figure 6-3 Sample Point Values order for Compression Data Format Version 2 ... 132
Figure 6-4 Sample Point Values order for Compression Data Format Version 3 ... 132
Figure 6-5 Sample Point Values order for Compression Data Format Version 4 ... 133
Figure 6-6 Block Diagram for Compressed Data Format Version 4, without any error. 133
Figure 6-7 Block Diagram for Compressed Data Format Version 4, with error. 134
Figure 6-8 Average BDIR Sizes for Different Data Formats and datasets..... 138
Figure 6-9 Compression Ratios for Different Data Formats and Datasets 139
Figure 6-10 Compression Ratio for Different Compression Format Versions and Compression Algorithms 140
Figure 6-11 Compression Ratio for different Datasets and Compression Algorithms amongst all versions tested 141
Figure 6-12 Error Rates for GMM and Different Data Formats 142
Figure 6-13 Error Rates for DTW and Different Data Formats 143
Figure 7-1 Signature Representation Data Block for 19794-11CD2 149
Figure 7-2 Signature Representation Data Block Proposed 150
Figure 7-3 Average BDIR sizes for different data formats and datasets 157
Figure 7-4 Average Compression Ratios for different data formats and datasets 158
Figure 7-5 Average number of Pen Dynamic Events for different data format versions and datasets 159
Figure 7-6 Average number of Pen Dynamic Events Reduction for different data format versions and datasets 159
Figure 7-7 Average number of Pressure Dynamic Events for different data format versions and datasets 160
Figure 7-8 Average number of Pressure Dynamic Events Reduction for different data format versions and datasets 160
Figure 7-9 Average compression ratio for the different data formats proposed and the different options for storing singular point types. 161
Figure 7-10 Error estimation for interpolated signal 163
Figure 7-11 Maximum Error for the interpolation of the X-Y Position Signal, comparison between the 19794-11CD2 and the new Proposal 164

Figure 7-12 Mean Error for the interpolation of the X-Y Position Signal, comparison between the 19794-11CD2 and the new Proposal 164
Figure 7-13 Maximum Error for the interpolation of the Pressure Signal, comparison between the 19794-11CD2 and the new Proposal 165
Figure 7-14 Mean Error for the interpolation of the Pressure Signal, comparison between the 19794-11CD2 and new Proposal 165
Figure 7-15 Maximum Error for the interpolation of the X-Y Position Signal, comparison between SPLINE and PCHIP interpolation methods..... 166
Figure 7-16 Mean Error for the interpolation of the X-Y Position Signal, comparison between SPLINE and PCHIP interpolation methods 166
Figure 7-17 Maximum Error for the interpolation of the Pressure Signal, comparison between SPLINE and PCHIP interpolation methods 166
Figure 7-18 Mean Error for the interpolation of the Pressure Signal, comparison between SPLINE and PCHIP interpolation methods 166
Figure 7-19 Stroke maintaining a certain constant value..... 167
Figure 7-20 Maximum Error for the interpolation of X-Y Position Signal, comparison between SPLINE and PCHIP interpolation methods 168
Figure 7-21 Mean Error for the interpolation of X-Y Position Signal, comparison between SPLINE and PCHIP interpolation methods 168
Figure 7-22 Maximum Error for the interpolation of the Pressure Signals, comparison between SPLINE and PCHIP interpolation methods 168
Figure 7-23 Mean Error for the interpolation of the Pressure Signals, comparison between SPLINE and PCHIP interpolation methods 168

ACRONYMS

API	Application Program Interface
ASVS	Automatic Signature Verification System
BDB	Biometric Data Block
BDIR	Biometric Data Interchange Record
CBEFF	Common Biometric Exchange Formats Framework
CD	Committee Draft
CDV	Committee Draft for Voting
CEN	Comité Européen de Normalisation (European Committee for Standardisation)
CMC	Cumulative Match Characteristic
DAPS	Distributed Application Platforms and Services
DET	Detection Error Trade-off
DIS	Draft International Standard
DTW	Dynamic Time Warping
EER	Equal Error Rate
EM	Expectation-Maximization algorithm
ETSIT	European Telecommunications Standards Institute
FDIS	Final Draft International Standard
FMR	False March Error
FMR-RF	False Match Rate for Random Forgeries
FMR-SF	False Match Rate for Skilled Forgeries
FNIR	False-Negative Identification Rate
FNMR	False Non-Match Error
FR	Fisher Ratio
FRPCA	Fisher Ratio and Principal Component Analysis combination
GMM	Gaussian Mixture Model
HD	Hellinger Distance
HMM	Hidden Markov Model
IC	Integrated Circuits
ICT	Information and Communications Technologies
IEC	International Electrotechnical Commission
IS	International Standard
ISO	International Organization for Standardization
IT	Information Technology
JTC	Joint Technical Committee
MCA	Minor Component Analysis
N/R	Not Reported
NN	Neural Network

NP	New work item Proposal
PCA	Principal Component Analysis
PCHIP	Piecewise Cubic Hermite Interpolation Polynomial
PDA	Personal Digital Assistant
PWI	Preliminary Work Item
RF	Random Forgery
ROC	Receiver Operating Characteristic
SBH	Standard Biometric Header
SC	SubCommittee
SF	Skilled Forgery
SVM	Support Vector Machine
SWG	Special Working Group
TC	Technical Committee
TPIR	True-Positive Identification Rate
TR	Technical Report
WD	Working Draft
WG	Working Group

Chapter 1 INTRODUCTION

1.1 MOTIVATION

The need to accurately and automatically verify claimed identities of users has become an important issue when considering new and upcoming techniques of performing electronic transactions. Unfortunately, such transactions have also increased the opportunities for fraudulent claims and "identity theft".

Biometrics is employed to accomplish current verification requirements. This technology provides high levels of security and is both convenient and comfortable for the user. Biometrics has already been deployed in many different scenarios, where one of the most common applications is in new generation identification documents, such as Citizen ID Cards and Electronic Passports.

Several biometric modalities are currently being tested for identity verification, but amongst all the possible biometric modalities, the handwritten signature has been used for the longest period of time as a means of identification. It is commonly found in commerce and banking transactions, credit card payments and, in general, all types of legal documents. Therefore, considering all the different biometric modalities, the signature is undoubtedly the most accepted for the majority of different scenarios. The image of the user's handwritten signature is already incorporated into ID documents. However, current error rates in verifying signature images are not yet sufficient for massive deployment.

The possibility of incorporating dynamic features, which are unique characteristics of every user, during the act of signing can provide additional verification mechanisms to be embedded into modern ID documents. By improving the error rates using these added characteristics, the handwritten signature will become a viable verification option for users of online processes such as e-banking and e-commerce.

However, to embed dynamic signature verification within ID documents, two limitations must be faced: memory capacity required for the users samples and template, and the computational load of the algorithms involved.

Furthermore, ID documents are designed with the idea of being used for a broad range of applications, in both national and international scenarios. To achieve this, the requirement for interoperability is of high importance, making the development and application of international standards imperative. One signature standard is already published: "ISO/IEC 19794 – Biometric Exchange Data Format – Part 7 - Signature/sign time series data". This standard provides two different data formats, a full format which stores the row data acquired by input devices with minimal transformation, and a compact format to be used in those systems involving smart cards or others tokens, which may require a smaller signature representation size. This standard is currently under revision and is planned to be released in 2013. There is another signature international standard under development, "ISO/IEC 19794 – Biometric Exchange Data Format – Part 11 – Signature/sign Processed Dynamic Data". This standard is defined for the storage of non-raw data acquired by signature input devices. This processed stored data is based on reducing the sample size by segmenting the signature into components of pen-strokes and pressure-strokes. The 19794-11 is defined as a compression of the part 7, and aims to provide reduced memory requirements when compared to the full format described in the part 7.

The main objective of this Thesis has been to investigate the information required for signature verification systems with limited resources in terms of both data size and computational load. The research work presented also analyses how the use of signature international standard data formats impacts on the size of the data involved in the process and the way this affects the performance of recognition algorithms in terms of their error rates.

The first two issues, user data size and computational load, are tackled by applying feature selection techniques. This allows the most effective features for verification to be identified, while at the same time discarding the least, minimizing the number of features. This implies smaller user samples and references and reduces the computational load.

In addition, and in accordance with international standards, this Thesis proposes solutions to assist the current standardization processes and to encourage future applications of handwritten signatures in electronic transactions and ID documents.

1.2 DOCUMENT STRUCTURE

This document has been divided into eight different chapters which will guide the reader through the State of the Art of the technique being investigated and to the improvements proposed by the author of this Thesis. The document has the following chapter structure

- ***Chapter 2 "Biometrics and International Biometric Standards":*** In this chapter the concept of Biometrics and its different modalities will be introduced. Several of the fundamental ideas on how a biometric system is evaluated are also explained here, defining commonly used evaluation biometric terms which will be used in later chapters. This chapter finalizes with an introduction to the international standards organisation which is leading the development of biometric standards, detailing the process that a standard must follow to get international consensus and reach its final publication. The most relevant biometric standard for this Thesis, "ISO/IEC project 19794 – Biometrics Data Interchange Format", will also be introduced.

- ***Chapter 3 "Automatic Signature Verification":*** The state-of-the-art for automatic signature verification is presented in this chapter. An overview introduction of this biometric modality is provided, followed by a deeper description of the capture devices available and the different comparison techniques that have been used in specialized literature. After this introduction, the databases used in this Thesis are described along with the results of the two evaluation campaigns that have been undertaken to date. As in chapter 2, this chapter concludes with a discussion on international standards, focusing on the specific signature standards within the ISO/IEC project 19794: Part 7 "Signature/sign time series data" and "Part 11 Signature/Sign Processed Dynamic Data", which are currently under development.

- **Chapter 4 "Improvement in Automatic Signature Verification":** This chapter focuses on the work done in signature verification algorithms to reduce the user model size and the computational requirements of the algorithms. Here, the two algorithms that have been analysed are explained in detail, as well as the feature selection techniques proposed. The results of applying these techniques to the signature verification algorithms are disclosed, along with an analysis of the impact of such selected feature sets to the requirements of data size and computational load.

- **Chapter 5 "Viability Analysis of Signature Standard Data Formats":** in this chapter an in-depth discussion on the international signature standard data formats is presented. A viability analysis of each of these standard data formats (19794-7 and 19794-11) is provided in terms of the sample size and performance. The results

obtained are disclosed, demonstrating several limitations of the current versions, and suggesting possible solutions that will be covered in later chapters.

- **Chapter 6 "Interoperability of Signature Biometrics at Signal Level":** The problems found in ISO/IEC 19794-7 "Signature/sign time series data" and the solutions proposed within this Thesis are analysed in this chapter. A new compressed compact format is proposed and is analysed and compared to the data formats already included in the 19794-7: full and compact data formats.

- **Chapter 7 "Interoperability of Signature Biometrics at Data Processed Level":** the focus of this chapter is on the ISO/IEC 19794-11 international standard. The drawbacks discovered within this standard, presented in chapter 5, are faced and different strategies are proposed to solve them. A new definition for the data format is defined and takes into consideration distinct ways of solving the limitations pointed out in chapter 5. Also the influence of the stored information on this data format is analysed. This work shows how these different options reflect on the sample size. The chapter ends with an analysis of the errors introduced when attempting to reconstruct the original temporal signal acquired from the input devices using the processed dynamic data stored within the 19794-11.

- **Chapter 8 "Conclusions and future work":** finally, this chapter summarizes the findings, contributions and conclusions described throughout this Thesis. Also included here is a description of several areas of further research within this area of the signature biometric modality.

Chapter 2 BIOMETRICS AND INTERNATIONAL BIOMETRIC STANDARDS

2.1 BIOMETRIC TECHNOLOGY

Biometric technology attempts to imitate the way humans recognise each other and is based on their physical or behavioural characteristics. We recognise our friends/relatives/colleagues in many different ways:

- by the way he/she looks (face recognition),
- by the way he/she walks (gait recognition),
- by the way he/she speaks (speech recognition),
- by the way he/she smells (scent recognition).

Human recognition is based on one or several of these traits.

A more formal and concise definition of *biometrics* can be found in "ISO/IEC TR 24741 Information Technology – Biometrics Tutorial" [1], where it is defined as:

"*automated recognition of individuals based on their behavioural and biological characteristics*"

Ted Dunstone and Neil Yager discuss, in their book, "Biometrics Systems and Data Analysis" [2] the multidisciplinary aspects of biometrics: "What other area of science or engineering combines aspects of biology, statistics, forensics, human behaviour, design, privacy and security - and also spans everything from the simple door lock to huge

government systems?". Apart from the lucrative investment, Biometric technology is currently a very exciting research field [3].

Biometrics has experienced impressive growth in the last few decades. In modern day recognition systems, biometric technology is commonly used to authenticate identities in diverse daily activities. Biometrics can be found in a wide range of applications, such as; Physical Access Control Systems (access to physical spaces such as enterprise buildings, houses, etc.), Logical Access Control Services (access to electronic systems such as networks, personal computers, etc.), Consumer Identification (health care, banking, online shopping, etc.), border-control (e-gates), etc. The distribution of biometrics for different applications in 2006 is shown in Figure 2-1, here it can be seen that Civil and Criminal identification occupies more than half of the market.

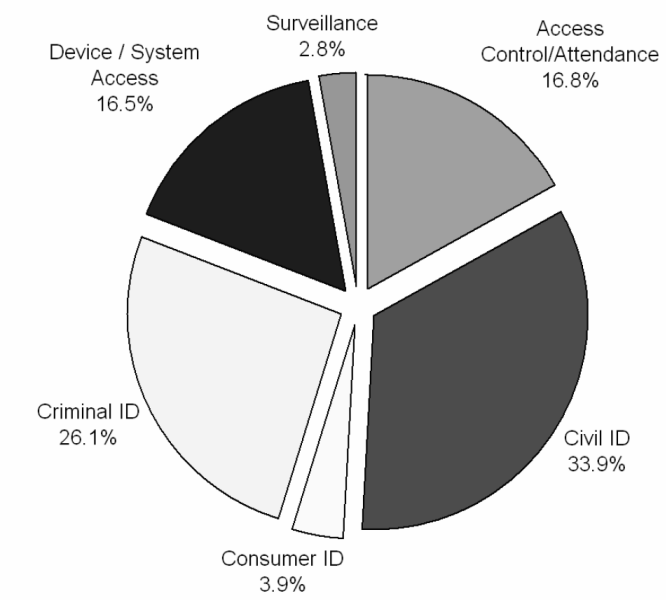


Figure 2-1 Biometric Market by Application, 2006 [3]

2.1.1 BRIEF HISTORY OF BIOMETRICS

The history of using biometric traits for identification purposes is discussed extensively in many publications [4-6]. Using biometric traits and biometrics to recognize different individuals dates back to ancient history, some of the many examples are listed below [7]:

- Fingerprints on legal documents dates back more than 8000 years to ancient Assyria and Babylonia,
- A Portuguese explorer in China reported that children's palm and foot prints were stamped on paper as a means of distinguishing between children,
- Fingerprint identification was used on government documents in Persia in the 14th Century.

The first documented work regarding the scientific approach to Biometric Recognition came from Sir William Herschel and Alphonse Bertillon. Fingerprint research by Sir William Herschel began in the late 1850's, and the Bertillon System of Anthropometric Identification proposed by Alphonse Bertillon is dated to the 1860's. The Bertillon Method (based on a number of bodily measurements includes, amongst others, height, weight, the length and width of the head, width of the cheeks, length of the trunk, feet, and ears) was used in France and the United States in the late 19th century and at the beginning of the 20th century. At the beginning of the 20th century, fingerprinting became the main method of assessing identity and was based on the works of Sir Francis Galton, a few years previous to the 1900's he provided the first scientific demonstration on the uniqueness of fingerprints for identification and indicated that they do not change throughout a person's life. F. Galton also identified the basic fingerprint features, or minutiae, which are currently in use.

Biometrics as used above does not meet the definition given at the beginning of this chapter, this is because it does not involve automation. Research on biometrics as an automatic process started in 1960 with the development of automated recognition systems in different modalities: speech [8], face [9] and fingerprint [10]. In the 1970s the first operational fingerprint [11] and hand geometry systems [12] were fielded. In the 1980s the first steps in biometric standardisation began (first version of the fingerprint exchange standards [13]). Both iris [14] and face recognitions systems [15] appeared in the 1990s. Nowadays, commercially-available systems are based on modalities such as fingerprint, iris, vascular, voice, handwritten signature, key-strokes, as well as many others.

2.1.2 BIOMETRICS MODALITIES

Not all biometric traits are suitable for the verification process. Experts in biometrics have studied the characteristics which a biometric trait must conform to [6] [16]. The following characteristics are generally acknowledged to provide adequate guidance for assessing the suitability of a biometric trait:

- **Universality**: this describes how common a biometric trait is among users, if it is present in all, or nearly all, members of the relevant population,
- **Uniqueness**: this is directly related to how well the biometric features can differentiate between users, making clear distinctions between any two members of the population,
- **Permanence**: measures how well a biometric trait can resist aging, disease, injury and any other temporal conditions,
- **Collectability**: ease of biometric data acquisition, depends on sensor technology and environmental conditions,
- **Performance**: covers a wide area, such as; accuracy, speed and robustness,

- **Acceptability**: level of user approval (willingness to use) of the technology. This characteristic may be affected by social, cultural and legal circumstances,
- **Circumvention**: this refers to how easy it is to deceive the identification system using fraudulent techniques.

Unfortunately, not all biometric traits fulfil all of the above characteristics, but good biometric traits comply with most of them to a certain degree. The perfect biometric trait does not exist, and the choice of trait to be used is highly dependent on the application and target population.

Biometric traits are normally divided into two main groups: physical or behavioural. Physical modalities are those referred to a biological characteristic of the user. The most commonly used physical biometric modalities are:

- Fingerprint,
- Face,
- Iris,
- Finger-vein,
- Vascular (palm-vein, finger-vein, etc.),
- Palm Prints,
- Hand geometry.

Behavioural modalities are those related to the way a user does something. Examples of behavioural modalities are:

- Voice recognition,
- Signature,
- Handwriting,
- Keystroke,
- Gait.

Details for these modalities (how they work, their basis, pros and cons) can be found in specialized literature [1-2, 5-6, 10, 16].

2.1.3 GENERAL BIOMETRIC SYSTEM

Although there are many different modalities and applications, most biometric systems share the same general scheme, shown in Figure 2-2. This scheme has been agreed on by biometric experts and is used as a reference scheme for standardisation procedures [1].

From this figure, the different phases of a biometric system are presented, these are: *Enrolment, Identification and Verification*. The following subsections will provide a brief overview of each of these. For a more detailed and deeper analysis, specialized literature [1-2, 6, 10, 16] dealing with each subsection can be reviewed.

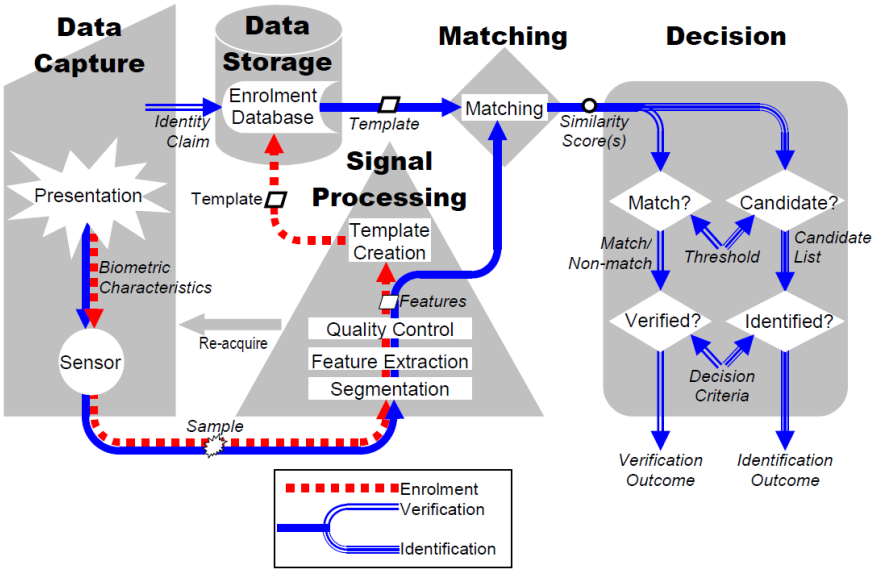


Figure 2-2 Components of a general biometric system [1]

2.1.3.1 Enrolment Phase

The first phase of a biometric system is always the enrolment. During this phase, a biometric reference model is created for each user. This reference is used in the rest of the biometric systems stages as a reference for comparison. The enrolment phase (see Figure 2-2) involves the acquisition of at least one biometric sample (data capture subsystem), which is then processed (signal processing subsystem) in order to obtain the features or mathematical model (feature extraction) which is used to generate a biometric reference for each user (template creation). This reference is stored (data storage subsystem) and used, when required, for a later comparison stage during either identification or verification. During the enrolment phase, other information, such as name, may be stored along with the biometric reference data.

2.1.3.2 Recognition Phase

Once the user is enrolled in a biometric system, the biometric reference data is available to perform recognition tasks, i.e. verification or identification. These phases are clearly defined in [17-19].

2.1.3.2.1 Verification

During verification (see Figure 2-2) the user presents the biometric data to the system (data capture subsystem) and at the same time, his/her claimed identity. This biometric raw data captured is then processed (signal processing subsystems) while the biometric reference data, for the identity claimed, is retrieved from the data storage subsystem. Both elements, the features that represent the biometrics data presented by the user and the biometric reference data retrieved from the Data Storage, are compared (comparison subsystem)

obtaining a similarity degree between them, generally referred to as "comparison score". This score is taken by the decision subsystem which verifies, based on a predetermined threshold level, if the claim regarding the user's identity is positive.

The verification decision outcome will be successful if a true claim is accepted and a false claim is rejected. The outcome will be considered erroneous if either a false claim is accepted (false non match error) or a true claim is rejected (false match error). The rating of these two errors, i.e. false non-match error rate (FNMR) and false match error rate (FMR), are used to compare the performance of different algorithms for verification tasks.

When the system is designed for use as an authentication system, the submission of biometric data during the enrolment phase is accompanied by other necessary additional identification data (ID, passport, etc.) in order to ascertain the validity of the identity.

The reference data may be stored in the followings ways:

- Locally in the data acquisition device,
- Locally on a personal computer,
- Centrally on a remote server,
- On a portable medium, token, such as a smart card.

2.1.3.2.2 IDENTIFICATION

Alike the verification process, during identification the user also supplies his/her biometric data to the data capture subsystem, however, in this case the claimed identity is not provided. The biometric system processes the raw data coming from the sensor and extracts the features (signal processing subsystem) and compares it to all the biometric references stored in the data storage subsystems. The biometric system attempts to locate the identifier for the users, providing a candidate list of enrolees based on the comparison scores achieved.

The outcome of this process is successful when the user is enrolled in the biometric system and his/her identity is included on the candidate list of enrolment records (true-positive identification). Otherwise, the identification process outcome will be considered erroneous, i.e. when the user is not enrolled and the candidate list is not empty (false-positive identification error), or the user identity is not included on the candidate list (false-negative identification error).

These errors are used to measure the performance of biometric systems for identification tasks. The *true-positive identification rate* (TPIR) of *rank r* is the proportion of identification transactions by a user enrolled in the system, for which the user's true identifier is included in the candidate list returned, composed for a maximum of *r* users (the *r* users with greatest scores over a certain threshold). This list will be empty if none of the identities obtain a score greater than the threshold. The false-negative identification-error rate is the proportion of identification transactions by users enrolled in the system, for which the user's correct identifier is not included in the candidate list returned.

2.2 BIOMETRIC SYSTEM EVALUATION

Biometric Systems rely on user data captured by an input device. Every biometric sample taken from this interaction between users and input devices are different. These samples can be used for template generation during enrolment or for a comparison score calculation in verification or identification processes. The intrinsic variability of biometric samples, especially for biometric systems based on behaviour traits, are thus that biometric decisions are based on probability, which can lead to mismatch errors. Measuring error rates in Biometrics is very important as these systems are subject not only to errors from pattern recognition but also from the capture process.

There are many guidelines for evaluating biometric systems [16-17, 20], as well as several ISO/IEC standards [18-19]. In order to avoid conflicting definitions, this document will use the ISO/IEC guidelines related to "Biometrics Performance Testing and Reporting" [18-19], which is based on the previous work "Best Practices in Testing and reporting Performance of Biometric Devices" [17].

Before describing the different biometric performance evaluations, both online and offline generation of matching scores needs to be explained. Testing a biometric system involves the collection of input biometric data samples that are then used to generate the template during the enrolment phase and for calculations of the matching score attempts.

The biometric traits captured can be used immediately for an *online* enrolment, verification or identification, or may be stored and used later for *offline* enrolment, verification or identification. The terms online and offline are defined as [17]:

- **Online**: Enrolment or calculation of comparison scores is said to be "*online*" when it is done at the time the image or signal is submitted. This has the advantage that the biometric sample can be immediately discarded, saving the need for storage and for the system to operate in a manner different from usual.
- **Offline**: Enrolment or calculation of matching scores is said to be "offline" when it is based on images or signals collected earlier. Collecting a database of images for offline enrolment and calculation of matching scores allows greater control over which attempts and template images are to be used in any transaction.

2.2.1 TYPES OF BIOMETRIC PERFORMANCE TESTING EVALUATIONS

In the guidelines laid out in "ISO/IEC 19795-1: Biometric Performance Testing and Reporting – Part 1: Principles and Framework" [18] and "Best Practices in Testing and reporting Performance of Biometric Devices" [17] three kinds of Biometric Performance Testing evaluations can be defined:

- **Technology evaluation**: The goal of a technology evaluation is to compare competing algorithms of a single technology. Testing of all algorithms is carried out on a standardised database collected by a "universal" sensor (i.e. a sensor that collects samples equally suitable for all algorithms tested). Nonetheless, performance against this database will depend upon both the environment and the population in which it is collected. Consequently, the "Three Bears" rule might be applied, attempting to create a database that is neither too difficult nor too easy, but is "just right" for the algorithms to be tested. Although sample or example data may be distributed for developmental or tuning purposes prior to the test, the actual testing must be done on data that has not previously been seen by algorithm developers. Testing is carried out using offline processing of the data. Because the database is fixed, the results of technology tests are repeatable.
- **Scenario evaluation**: The goal of scenario testing is to determine the overall system performance in a prototype or simulated application. Testing is carried out on a complete system in an environment that models a real-world target application of interest. This includes not only the algorithm, but also the data capture sub-system. The evaluation is carried on in a Laboratory, where the conditions of the final deployment are modelled.
- **Operational evaluation**: The goal of operational testing is to determine the performance of a complete biometric system in a specific application environment with a specific target population, at the location and conditions of the final deployment. Depending upon data storage capabilities of the tested device, offline testing might not be possible. In general, operational test results will not be repeatable because of unknown and undocumented differences between operational environments.

In this Thesis technology evaluations are applied to different algorithms and the different settings for those algorithms using offline signatures collected from public databases. The different public signature databases will be discussed in the following chapter. In the technological evaluation the offline enrolment and calculation of matching scores are compiled so that different algorithms can be tested using the same databases, allowing a comparative analysis.

These offline collected signatures are not to be confused with offline signature algorithms. In the first case, the term offline refers to the fact that the collection process and

the enrolment or verification/identification tasks are carried out at a different time, as described above. In the second case, the term offline refers to the type of signature algorithms that deal with signature images instead of the temporal signals captured by the input devices. These two different kinds of signature algorithms are described in Chapter 3.

2.2.2 PERFORMANCE MEASURES: ERROR RATES

Once the type of evaluation has been decided upon and executed, the algorithm's performance will be assessed in terms of the error rates. The rates used in this Thesis are described below. Other error rates exist, such as; Failure to Acquire, Failure to Enrol, etc., however these have not been considered as they are not relevant to the work performed for this Thesis. For more information on all the different error rates that can be used to measure the performance of a biometric system the reader should refer to the relevant literature, e.g. [2, 17-20].

The most common error rates, as defined by the ISO/IEC 19795-1 [18] and "Best Practices in Testing and reporting Performance of Biometric Devices" [17] are:

- **False Match Rate (FMR)**: proportion of zero-effort impostor attempt samples falsely declared to match the compared non-self template.
  The false match rate is the expected probability that a sample will be falsely declared to match a single randomly-selected "non-self"[1] template. (A false match is sometimes called a "false positive" in the literature.)
- **False Non-Match Rate (FNMR)**: proportion of genuine attempt samples falsely declared not to match the template of the same characteristic from the same user supplying the sample.
  The false non-match rate is the expected probability that a sample will be falsely declared not to match a template of the same measure from the same user supplying the sample. (A false non-match is sometimes called a "false negative" in the literature.)
- **(True-Positive) Identification Rate (TPIR):** proportion of identification transactions by users enrolled in the system in which the user's correct identifier is among those returned.
- **False-Negative Identification-Error Rate (FNIR):** proportion of identification transactions by users enrolled in the system in which the user's correct identifier is not among those returned.

[1] Non-self: Genetically different. It has been noted in the literature that comparison of genetically identical biometric characteristics (for instance, between a person's left and right eyes or across identical twins) yields different score distributions than comparison of genetically different characteristics. Consequently, such genetically similar comparisons should not be considered in computing the false match rate.

- **Equal Error Rate:** is the rate when both FMR and FNMR are equal. The equal error rate is commonly used when a rapid comparison of two systems is required. The lower the EER value the more accurate the system is considered to be.

There is also a need to clearly define "genuine" and "impostor" transactions, as they are continuously used in biometric systems based on handwritten signature. These definitions are independent to the type of test being performed, and are defined in [17] as:

- **Genuine attempt**: A "genuine" attempt is a single good faith attempt by a user to match his or her own stored template.
- **Impostor attempt**: An "impostor" attempt is a single "zero-effort" attempt, by a person "unknown to the system", to match a stored template.
- **Zero-effort attempts**: An impostor attempt is classed as "zero-effort" if the individual submits their own biometric feature as if they were attempting successful verification against their own template. In the case of dynamic signature verification, an impostor would therefore sign his or her own signature in a zero-effort attempt. In such cases, where impostors may easily imitate aspects of the required biometric, a second impostor measure based on "active impostor attempts" may be required.

As pointed out in [17], the biometric signature verification systems require more than one impostor attempt definition. Zero-effort impostor attempts are commonly referred to as "random forgery attempt", where an impostor attempts to impersonate another user identity by signing with his/her own signature. The second impostor measure based on "active impostor attempts" is commonly referred to as "skilled forgery attempt" and indicates that the impostor would try to imitate the aspects (shape and also even dynamic aspects) of the signature belonging to the claimed identity. Normally, Signature Corpus Databases [21-22] contains impostor signatures of users where the impostor has a static image of the user signature to forge and can practice the forgery signature until he/she are confident to imitate it. As a result, in signature biometric systems two different FMRs are generally provided. In this Thesis, these are referred to as:

- **False Match Rate for Random Forgeries (FMR-RF):** proportion of zero-effort impostor attempt samples falsely declared to match the compared non-self template.
- **False Match Rate for Skilled Forgeries (FMR-SF):** proportion of active impostor attempt samples falsely declared to match the compared non-self template.

2.2.3 GRAPHICAL PRESENTATION OF RESULTS

To analyse and compare the performance of different biometrics systems, Receiver Operating Characteristic (ROC) curves and the Detection Error Trade-off (DET) are used for the verification systems. In the case of identification systems, cumulative match characteristic curves (CMC) are used. These curves are defined [18] as:

- **Receiver Operating Characteristic (ROC) Curve**: plot of the rate of false positives (i.e. impostor attempts accepted) on the x-axis versus the corresponding rate of true positives (i.e. genuine attempts accepted) on the y-axis, these are plotted parametrically as a function of the decision threshold.

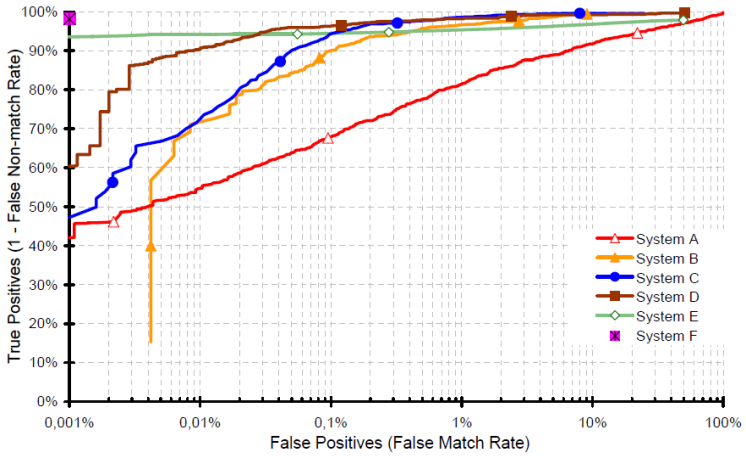


Figure 2-3 Example of ROC curves for different biometric systems [18]

- **Detection Error Trade-off (DET) Curve**: modified ROC curve which plots error rates on both axes (false positives on the x-axis and false negatives on the y-axis).

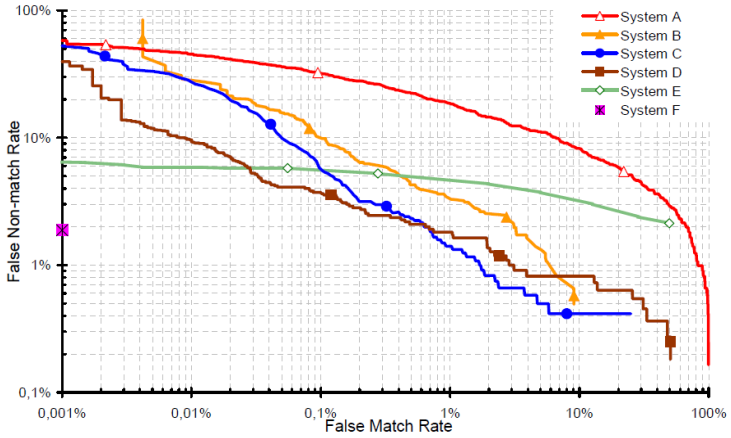


Figure 2-4 Example of DET curves for different biometric systems [18]

- **Cumulative Match Characteristic (CMC) Curve**: graphical presentation of results from an identification task test, where the rank values are represented on the x-axis and the probability of correct identification at or below that rank on the y-axis.

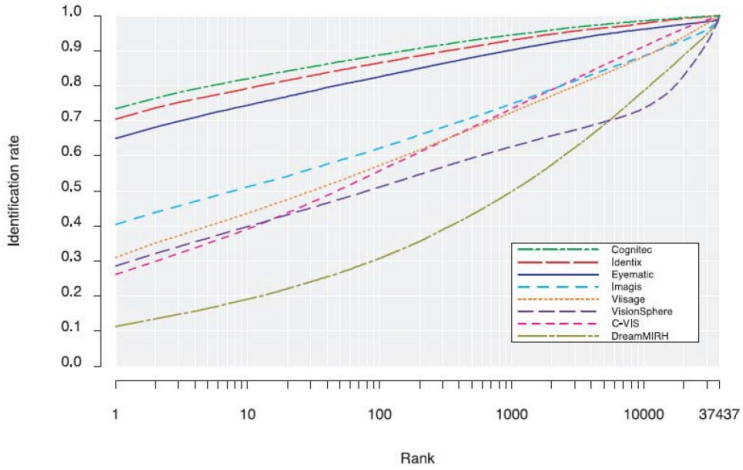


Figure 2-5 Example of CMC curves for different biometric systems [18]

2.3 INTERNATIONAL STANDARDS

The work performed in this Thesis has been designed to help develop signature international standards. This fact has motivated the current section which is dedicated to International Standards, and introduces the ISO/IEC JTC1 Subcommittee 37, that forms the major international forum for biometrics standardisation. Also in this section, the standardization project ISO/IEC 19794 will be described, where this defines the biometric data interchange format standards.

2.3.1 INTERNATIONAL STANDARDS AND ORGANIZATIONS

A standard is defined as a "document, established by consensus and approved by a recognized body that provides, for common and repeated use, rules, guidelines or characteristics for activities or their results, aimed at the achievement of the optimum degree of order in a given context" [23]. This document "should be based on the consolidated results of science, technology and experience, and aimed at the promotion of optimum community benefits" [23]. When this document is adopted by international standard organizations, it becomes an International Standard.

The use of international standards provides a means of overcoming technical barriers that may be caused by differences among technical requirements from different countries, markets and/or organizations. International Standards allow vendors to supply their products to a worldwide market, assuring that their products are compatible within their specific markets. The use of the International Standards and their conformity assessment enhance confidence in products for users, assure quality, safety, reliability, efficiency and interoperability among different products.

International Standards also represent an international consensus on the state of the art for specific areas.

Standards are generally developed by non-profit organizations. Within these organizations the most relevant are ISO, IEC, ETSI or CEN. Several of these are based on transnational interest (e.g. ISO, IEC), and others are based on specific markets/sectors (e.g. ETSI).

The work performed in this Thesis focuses on both ISO and IEC. In the following paragraphs of this section their structure and characteristics are described.

The International Organization for Standardization (ISO) is the world's largest standards developing organization. It is made up of more than 160 members, with more than 18,000 international standards and other types of published normative documents. ISO was founded in 1947 and its work programme ranges from standards for traditional activities, such as

agriculture and construction, through to mechanical engineering, manufacturing and distribution, transport, medical devices, information and communication technologies, as well as standards for good management practice and services [24].

The International Electrotechnical Commission (IEC) is the leading global organization that prepares and publishes international standards for all electrical, electronic and related technologies. IEC was founded in 1904 and is currently composed of 71 members. IEC has published more than 6000 documents (5000+ international standards, 350+ technical reports and 200+ technical specifications amongst others) [25].

2.3.2 ISO/IEC JTC1

As it has been explained in the previous section, both ISO and IEC cover a wide range of areas. To deal with several of the common areas covered by both, the organizations have created an alliance that has been implemented by creating joint committees to develop standards within these areas in a cooperative way.

Figure 2-6 ISO and IEC Logos

This has been the case for the Information Technology (IT) area, where in 1987 ISO and IEC created a joint technical committee, named ISO/IEC JTC1 to deal with all issues related to the IT sector.

The JTC1 develops worldwide Information and Communication Technologies (ICT) standards for business and consumer applications. Additionally, the JTC 1 provides the standards approval environment for integrating diverse and complex ICT technologies. The JTC1 defines its mission as “Develop, maintain, promote and facilitate IT standards required by global markets, meeting business and user requirements and concerns:

- Design and development of IT systems and tools,
- Performance and quality of IT products and systems,
- Security of IT systems and information,
- Portability of application programs,
- Interoperability of IT products and systems,
- Unified tools and environments,
- Harmonized IT vocabulary,
- User friendly and ergonomically designed user interfaces.”

From the beginning of JTC 1, 14 years ago, a number of successful and relevant ICT standards have been implemented in many fields, such as; multimedia (e.g.: MPEG), Integrated Circuits (IC) cards ("smart cards"), ICT security, database query and programming languages as well as character sets.

Within JTC1, different working groups and subcommittees have been created to deal with specific areas related to IT. These groups are listed in the following table:

Table 2-1 ISO/IEC JTC1 Subcommittees

Subcommittee/ Working Group	Title
SWG	Special working group on planning
SWG 1	Accessibility (SWG-A)
WG 6	Corporate Governance of IT
WG 7	Sensor networks
SC 2	Coded character sets
SC 6	Telecommunications and information exchange between systems
SC 7	Software and systems engineering
SC 17	Cards and personal identification
SC 22	Programming languages, their environments and system software interfaces
SC 23	Digitally Recorded Media for Information Interchange and Storage
SC 24	Computer graphics, image processing and environmental data representation
SC 25	Interconnection of information technology equipment
SC 27	IT Security techniques
SC 28	Office equipment
SC 29	Coding of audio, picture, multimedia and hypermedia information
SC 31	Automatic identification and data capture techniques
SC 32	Data management and interchange
SC 34	Document description and processing languages
SC 35	User interfaces
SC 36	Information technology for learning, education and training
SC 37	Biometrics
SC 38	Distributed application platforms and services (DAPS)

2.3.3 ISO/IEC JTC1 SC37 BIOMETRIC STANDARDS

Biometrics can be found in national defence applications (i.e. the use of biometrics in machine readable travel documents where the passport is accepted worldwide, law enforcement biometric devices etc.) and in commercial fields ranging from financial transactions to visitor authentication.

With the rapid dissemination of biometric technologies it is important to recognize that systems and applications based upon consensus-based biometric standards are more likely to be interoperable, scalable, usable, reliable, secure, and in many cases more economical than proprietary systems.

The SC37 was formed in 2002, and is responsible for the standardization of generic biometric technologies pertaining to human beings to support interoperability and data interchange. It has currently developed 59[2] standards, and is composed of 28 members.

SC37 is subdivided into 6 working groups:

Table 2-2 ISO/IEC JTC1 SC37 Working Groups

Subcommittee/ Working Group	Title
JTC 1/SC 37/WG 1	Harmonized biometric vocabulary
JTC 1/SC 37/WG 2	Biometric technical interfaces
JTC 1/SC 37/WG 3	Biometric data interchange formats
JTC 1/SC 37/WG 4	Biometric functional architecture and related profiles
JTC 1/SC 37/WG 5	Biometric testing and reporting
JTC 1/SC 37/WG 6	Cross-Jurisdictional and Societal Aspects of Biometrics

The SC37 WG1 is responsible for defining the harmonized biometric vocabulary used for the complete SC37 and assures that all projects and parts use the same definitions for the same concepts.

The SC37 WG2 specifies interfaces and interactions between biometric components, systems and subsystems. Among the defined interfaces, the BioAPI specification (ISO/IEC 19784-1) defines an open system standard application program interface (API) that allows software applications to communicate with a broad range of biometric technologies. Another important interface that has been identified is the Common Biometric Exchange Formats Framework (CBEFF) defined in ISO/IEC 19785-1. This defines meta-data to describe biometric data in the structure (e.g., identification of the data format and modality), enabling applications to obtain the data of interest without having to decode it.

[2] Up to June 2011

The SC37 WG3 project 19794 specifies the content, meaning, and representation of formats for the interchange of biometric data and provides platform independence. This project is of special relevance to this Thesis and will be detailed later in this chapter. The SC37 WG3 also specifies the methodology used to perform conformance testing (project 29109 for first generation of data formats), and in some parts has begun the task of defining metrics for biometric sample quality (project 29194).

The SC37 WG4 defines profiles for biometric applications which specify the base standards that apply, identifying which classes, conforming subsets, options, and parameters of those base standards or standardized profiles are required to achieve a particular function and/or application.

The SC37 WG5 defines standard testing methodologies to evaluate the performance of systems and devices.

Finally, the SC37 WG6 develops technical reports on the "Cross-Jurisdictional and Societal Aspects of Implementations of Biometrics Technologies", which deal with legal, social, cultural, and ethical issues that are related to biometric methods. The aim is to achieve an internationally harmonized and practical assessment of biometrics, which above all, provides developers and users with biometric guidelines for legally binding and socially acceptable applications that go beyond the purely technical point of view.

2.3.4 STAGES ON THE DEVELOPMENT OF INTERNATIONAL STANDARDS

An International Standard represents an extensive review of applications from a large number of international experts. These experts represent different national bodies and express the interest of their countries. As a result, an International Standard is put in place when an agreement and consensus has been reached from all nationals bodies involved in its development.

As mentioned above, Biometric International Standards within the ISO/IEC are developed by the subcommittee 37, who follow a seven-stage process defined by the ISO/IEC in "ISO/IEC Directives – Part 1: Procedures for the technical work" [26]. Table 2-3 summarizes these different stages along with the titles of the documents associated with each project stage:

Table 2-3 ISO/IEC Project Stages and associated documents [26]

Project Stage	Associated Document	
	Name	Abbreviation
Preliminary Stage	Preliminary work Item	PWI
Proposal Stage	New Work Item Proposal	NP
Preparatory Stage	Working Draft(s)	WD
Committee Stage	Committee draft(s)	CD
Enquiry Stage	Enquiry Draft	DIS (ISO) or CDV (IEC)
Approval Stage	Final Draft International Standard	FDIS
Publication Stage	International Standard	ISO, IEC or ISO/IEC

The preliminary stage has been put in place to analyze issues which are not yet sufficiently mature for processing to further stages. This stage is used to elaborate on new work item proposals. The document elaborated in this stage is named “preliminary work item” (PWI).

When the issue proposed is sufficiently mature, it can be submitted as a “new work item proposal”, and is generally launched by a national body as a “national proposal” (NP) to the relevant technical committee (TC) or subcommittee (SC). The need for the International Standard has to be confirmed by the international community. The NP is circulated to be voted upon by the TC or SC body members. If the NP obtains sufficient support and commitment for active participation in its development, the NP is then accepted for inclusion in the TC or SC work programme.

After the NP is accepted, the proposal stage commences where a working group of experts is set up by the TC or SC. The TC or SC define the task(s) and set the target date(s) for submission of the working draft(s) (WD). This stage concludes when the developed working draft contains the best technical solution to the problem being addressed. At this point, the draft is forwarded to the working group's parent committee for the consensus-building stage which is referred to as the *committee stage*.

The *committee stage* is the principal stage where comments from national bodies are taken into consideration, although national body comments may have been included during the development of the WD. The committee drafts (CD) are circulated for voting and comments by the TC or SC body members. The comments are taking into consideration for each round of the CD, and once consensus has been reached, the text is finalized for submission as a “draft international standard” (DIS), and goes to the *enquiry stage*.

During the *enquiry stage* the Draft International Standard (DIS) is circulated to all ISO member bodies by the ISO Central Secretariat to be voted on and for commenting. Again, successive DISs may have to be considered for the DIS to be approved for submission as a Final Draft International Standard (FDIS). During the enquiry stage no technical comments are expected, but in the case where they are submitted, they can be taken into consideration.

The Final Draft International Standard (FDIS) at the approval stage is circulated to all ISO member bodies by the ISO Central Secretariat for a final Yes/No vote. At this stage no technical comments can be submitted. If the FDIS obtains enough support from the TC or SC body members, the document is approved for publication as an International Standard (ISO/IEC IS).

Once a final draft International Standard has been approved, only minor editorial changes can be introduced in the final text for the publication stage. The final text is sent to the ISO Central Secretariat which publishes the International Standard (IS).

All International Standards are reviewed at least three years after publication and every five years after the first review by all ISO member bodies.

2.3.5 ISO/IEC PROJECT 19794 – BIOMETRICS DATA INTERCHANGE FORMAT

The ISO/IEC project 19794 “Biometrics Data Interchange Format” specifies the content, meaning, and representation of formats for the interchange of biometric data and assures platform independency and interoperability among biometric systems. This Biometrics Data Interchange Records can be embedded within a superior header structure named Common Biometric Exchange Formats Framework (CBEFF) which is defined in the ISO/IEC 19785 “Information technology - Common Biometric Exchange Formats Framework” [27], or can also be used by itself. The CBEFF structure consists of three parts: the standard biometric header (SBH), the biometric data block (BDB), and the security block (SB). The SBH fields are defined in the ISO/IEC 19785 “Information technology - Common Biometric Exchange Formats Framework” project, whereas the BDB structure is defined in the ISO/IEC 19794 “Information technology -- Biometric data interchange formats” project. It may be seen from Figure 2-7 that the Project 19794 represents the core component of biometric interoperability among biometric systems.

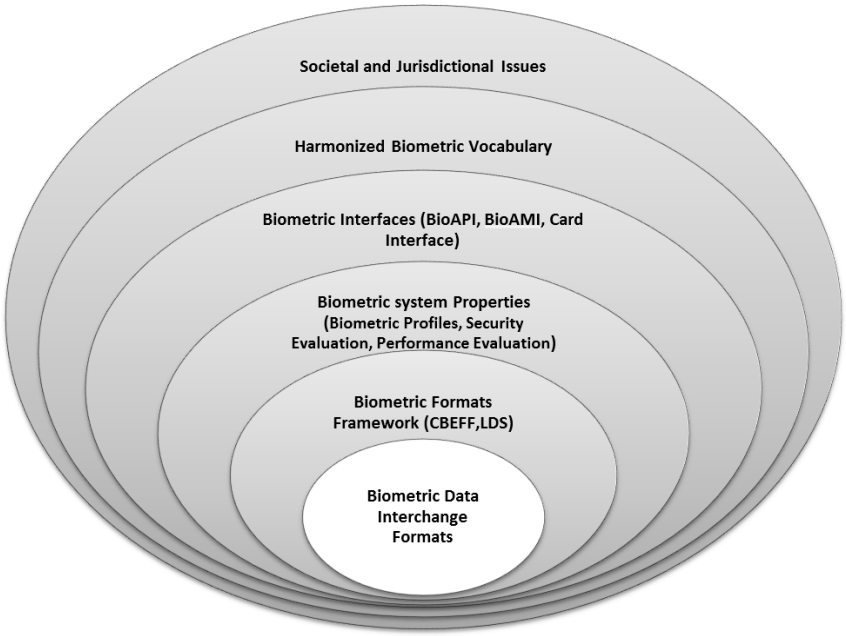


Figure 2-7 General interrelation model of biometric issues [28]

The project 19794 is split in different parts. Part 1: Framework [28] sets the context for the standardisation of BDBs and their use in other biometrics data structures (Part 2-N). Parts 2-N are specific for each biometric modality. Currently, the parts that are considered are:

- Part 2: Finger minutiae data,
- Part 3: Finger pattern spectral data (although this has been removed from the work plan),
- Part 4: Finger image data,
- Part 5: Face image data,
- Part 6: Iris image data,
- Part 7: Signature/sign time series data,
- Part 8: Finger pattern skeletal data,
- Part 9: Vascular image data,
- Part 10: Hand geometry silhouette data,
- Part 11: Signature/sign processed dynamic data,
- Part 12: Face Identity Data (although this has been removed from the work plan),
- Part 13: Voice data,
- Part 14: DNA data.

The first generation of these biometric data interchange format standards were released between 2005 and 2007 as shown in the following table. All Parts used binary encoding to describe the data formats.

Table 2-4 Publication Date for Biometric Data Interchange Format Standards, first generation

ISO/IEC Project	Publication Date	Title
19794-1	2006	Framework
19794-2	2005	Finger minutiae data
19794-3	2006	Finger pattern spectral data
19794-4	2005	Finger image data
19794-5	2005	Face image data
19794-6	2005	Iris image data
19794-7	2007	Signature/sign time series data
19794-8	2006	Finger pattern skeletal data
19794-9	2007	Vascular image data
19794-10	2007	Hand geometry silhouette data
19794-11	Withdrawn	Signature/sign processed dynamic data
19794-12	Withdrawn	Face Identity Data

It may be observed from Table 2-4 that part 11 and 12, dealing with Signature/Sign processed dynamic data and Face Identity Data, were withdrawn. This is due to the absence of significant comments from the National Bodies.

The second generation of these biometric data formats began after the publication of their first generation, and has attempted to incorporate all new advances in biometric systems.

The second generation framework attempts to achieve harmonization amongst all the different parts (Part 2-N), as this has not been achieved in the first generation. The new framework includes within its 3rd clause all the definitions used amongst multiple parts (Part2 – N), leaving specific biometric modality definitions to their respective parts. The framework provides a new common description for the general header along with a common structure for the representation header, which can be completed with the specific needs of different modalities. These definitions are used across all subsequent parts.

Most of the parts published in the first generation of biometric data formats are currently under revision to become their second generation, except Part 3 “Finger Pattern Spectral data” and Part 10 “Hand geometry silhouette data”. The Part 11 “Signature/sign processed dynamic data” was presented again as a new work item and accepted. Another two modalities were added in this second generation: Voice (Part 13) and DNA (Part 14).

Table 2-5 summarizes the stage[3] of all the parts within the second generation of the 19794 project:

Table 2-5 Stage of Biometric Data Interchange Format Standards, second generation

ISO/IEC Project	Stage	Title
19794-1	FDIS	Framework
19794-2	FDIS	Finger minutiae data
19794-3	Withdrawn	Finger pattern spectral data
19794-4	FDIS	Finger image data
19794-5	FDIS	Face image data
19794-6	FDIS	Iris image data
19794-7	CD	Signature/sign time series data
19794-8	FDIS	Finger pattern skeletal data
19794-9	FDIS	Vascular image data
19794-10	Not Initiated	Hand geometry silhouette data
19794-11	FCD	Signature/sign processed dynamic data
19794-13	CD	Voice Data
19794-14	CD	DNA Data

[3] Up to May 2011

2.4 CONCLUSIONS

This chapter has introduced the fundamental ideas required to understand the function of biometric systems and has also explained several concepts that will allow the reader to understand the findings and contributions of this Thesis.

A general biometric system overview has been presented along with the basis for its evaluation. These concepts are used throughout chapter 4, where the contribution of this Thesis to the improvement of automatic signature verification is presented.

Finally, the international biometric standards and their development have been discussed, and will assist the understanding of the research work on signature standard developments detailed in chapters 5 to 7.

Chapter 3 AUTOMATIC SIGNATURE VERIFICATION

3.1 INTRODUCTION

For centuries the most commonly used and established techniques for personal verification in daily activities is the handwritten signature. It is commonly used in commerce and banking transactions, credit card payments and, in general, all types of legal documents. When considering non-automated systems, amongst all the possible biometric modalities the user's signature is almost certainly the most accepted method.

Furthermore, a handwritten signature has one major advantage over other modalities when considering legal issues in the majority of countries. The act of signing a document can legally prove that the signer has read it, understood it and, therefore, the signature is used to bind the individual with the disposition contained in the document.

Another advantage of a handwritten signature as a biometric modality is that its acquisition is straightforward, relatively low cost when compared with other biometric modalities, and user friendly. Moreover, new touch-screen devices such as smart-phones and tablets have extended the acquisition process among potential users.

The characteristics of handwritten signatures make them ideal candidates for biometric verification. However, several disadvantages arise when compared with other biometric modalities such as iris and fingerprint modalities. In contrast to the aforementioned biometric modalities, handwritten signatures fall under the category of behavioural biometrics. This behavioural characteristic implies that handwritten signatures have a higher variability than other traits based on physical modalities. Two signatures from the same

individual are never identical, and in some cases the variability in signatures is excessively high to be acceptable. Moreover, signatures are affected by psychological states such as fatigue, stress or distraction. Another disadvantage with signatures is their evolution over time, i.e. change of user signature. Aging is also another important factor as illness can affect the ability to grip the stylus and perform the signature. Additionally, since this modality is classified as behavioural it is vulnerable to imitations; therefore, it is considered not to be as safe as other modalities when considering fraud.

One of the earliest published works that studied handwritten signatures as a means of identity verification came from Osborn [29] and has looked at this modality from a forensic point of view. In this work the author suggested that signature forgery involves a double process where the forger is not only required to copy the features of the written signature but also to hide the writer's own personal writing characteristics. The author studied the conditions that may affect signatures, amongst others are hastily written signatures, strange pen, and unaccustomed location. Osborn discusses the variation of signatures from the same individual, however, on close inspection of the signature a marked and unmistakable individuality is seen. The author presents as an example two sets of signatures from a celebrated case of a contested will in New York (Figure 3-1).

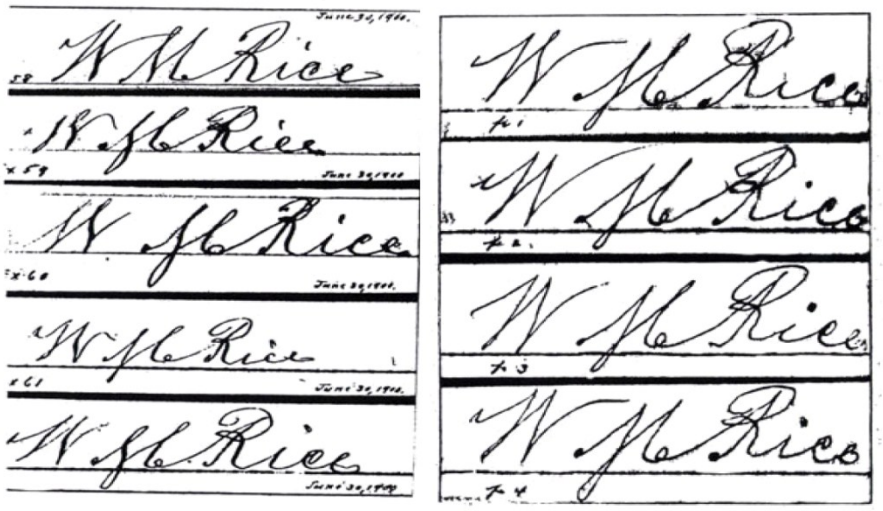


Figure 3-1 Signature Sets in a Disputed Will Case around 1900 [29]

The first automated signature recognition system was developed for North American Aviation in 1965 [30]. In the 1970s a patent was awarded for dynamic signature recognition, using the pressure acquired by a "personal identification apparatus" [31]. Also appearing in the 1970s was the popular published work titled "Automatic Signature Verification Based on Accelerometry" by Herbst and Liu [32] and "Online Signature Verification" by Farag and Chein [33]. Since then, vast work has been carried out on automatic signature verification and has been summarized in several state-of-the-art publications. In 1989, R. Plamondon and G. Lorette published an interesting review on automatic signature verification [34] which was updated 5 years later in 1994 [35] and again in 2000 [36]. Other interesting reviews have been published in 1997 by G. Gupta [37], 2004 by G. Dimauro [38], 2005 by M. Faundez [39],

2006 by Gupta [40] and recently in 2008, a complete and extensive review, by D. Impedovo [41].

All these works reveal the importance of automatic signature verification. It has attracted many researches from universities and companies around the world, whom are interested in both the scientific challenges and the valuable application of this field.

The last few years have seen the release of a great range of capture devices for signatures based on acquiring dynamic signals (i.e. time series channels), this is mainly due to application-led demands. These devices are currently being introduced in our daily lives. For example many shopping centres have begun to use them for credit card transactions to simplify the receipt management process, save paper, energy and money (paperless process). The increased availability of signature systems has presented the biometric community with the opportunity to use their capabilities to improve transactional security via the deployment of automatic signature verification systems.

3.1.1 AUTOMATIC SIGNATURE VERIFICATION: ONLINE AND OFFLINE

Signature Verification Systems are generally split into two main groups: offline (commonly referred to as static) and online (referred to as dynamic).

The difference between these groups is based on the information acquired. In offline systems an image of the signature is used. This is captured by scanning or photographing a signature made on a piece of paper. The signature sample is then reconstructed in the form of a data image that can be black and white, grey scale or colour. The greater the image resolution, the more information experts and systems have to assess the identity of the user.

Dynamic signature systems use input devices which can capture the movement of the stylus during the act of signing. Some devices, such as digital tablets, not only capture the movement of the stylus in the x-y plane, but also the pressure exerted by the tip of the stylus on the writing-surface along with the angles at which the stylus is characteristically held, in general tilt and azimuth angles. Therefore, the dynamic signature systems not only capture the final graph of the signature, but also the characteristics of the event of signing.

Dynamic systems have a lot more information available than static systems, and this fact is reflected on the performance of the verification, achieving lower error rates.

3.1.2 SIGNATURE FORGERIES

As mentioned in chapter 2, one of the greatest challenges of signature verification systems, compared with other biometrics modalities, is dealing with forged signatures which imitate both the shape and the execution.

It has been indicated in [17], biometric signature verification systems need more than one impostor attempt definition. Zero-effort impostor attempts are normally known as "random forgery attempt", where an impostor would sign with his or her own signature attempting to impersonate another user identity.

In section 2.2.2 two different impostor attempts definition were given: "random forgery" and "skilled forgery". The second impostor measure is the "active impostor attempts", this is not a trivial issue and is not contained within the scope of the "ISO/IEC Information technology -- Biometric performance testing and reporting" [18]. In [42], 7 different levels of "active impostor attempts" are defined, all of which depend on prior knowledge of the signature to forge from the impostor, these are listed in Table 3-1.

Table 3-1 Active Impostor Attempts Levels for Biometric Signature Systems

Level	Information Available
0	Impostor has not relevant knowledge of claimed user identity
1	Impostor knows claimed user identity's name
2	Impostor has seen a static image of claimed user identity's signature
3	Impostor can see a static image of claimed user identity's signature at the time of signing
4	Impostor is able to trace a sample of claimed user identity's signature
5	Impostor has recently witnessed of claimed user identity's signature
6	Impostor has repeatedly witnessed claimed user identity's signature

3.2 CAPTURE DEVICES

Even though digital tablets are still the most used signature input device, there have been some attempts at developing a stylus capable of capturing signature dynamics without the requirement for any additional equipment.

For example, the Biometric Smart Pen (BiSP) presented in [43-44] describes a stylus composed of optical sensors for recording the x and y movements, pressure sensors that record the pressure in 3 directions and also tilt sensors to measure angles (Figure 3-2). Another stylus input device was presented in [45] and is composed of accelerometers, a pressure transducer and orientations by sensing gravitational acceleration (Figure 3-3).

Figure 3-2 Biometric Smart Pen [43]

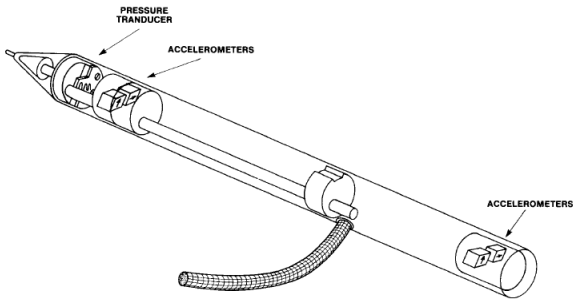


Figure 3-3 Accelerometer Pen [45]

Recently, information on a novel stylus input device has been published [46], and is called the IMUPEN. This device is composed of a triaxial accelerometer, two gyroscopes, a microcontroller, and, an RF wireless transmission module.

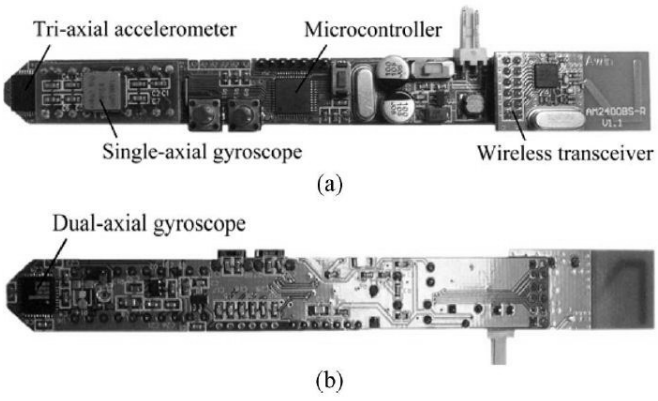


Figure 3-4 IMUPEN [46]

Although stylus-based signature input devices seem to be more user friendly, as they substitute the traditional pen in the act of signing, the most widespread input devices are digital tablets. In Figure 3-5 and Figure 3-6 some examples of digital tablets are shown.

Figure 3-5 Genius Tablet [47]

Figure 3-6 Wacom Tablets [48]

In general, digital tablets are connected to a computer via the USB interface. The tablet has a sensitive surface, which captures the movements from the stylus and transmits them to the computer. Tablets transmit temporal series vectors such as x and y position, pressure and, the more sophisticated tablets include inclination and azimuth. The space resolution, commonly referred to as dots per inch (dpi), range from 1000 to 5000 dpi. The pressure, if present, typically ranges from 256 to 2048 levels. The inclination and azimuth angles have a resolution of approximately +/- 0.5º. These signals are sampled at frequencies ranging from 50Hz to 200Hz.

In Figure 3-7 a graphical description of the different signals captured by a digital tablet used as a signature input device is presented.

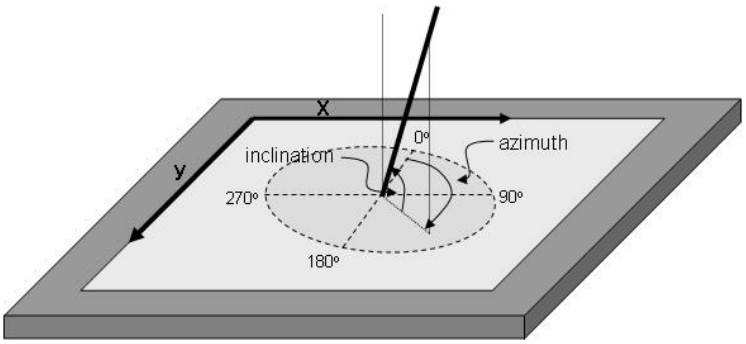


Figure 3-7 Signals acquired by digital tablets

As an example, the tablet used for capturing the MCyT [21] database was a Wacom Intuos 2 A6, where the authors acquired the following signals (value ranges are shown in parenthesis):

- X axis position (0-12700),
- Y axis position (0-9700),
- Pressure (0-1023),
- Azimuth angle (0-360),
- Inclination angle (0-90).

The use of these devices in industry as signature input devices is constantly growing, where their new manufacturing process follows designs which suit online signature verification specifications (Figure 3-8). These devices are widely accepted by users, and several offer interactive information for the user on the built-in screen.

Figure 3-8 Wacom Signature Input Devices [48]

But in the last couple of years, new touch-screen devices have become a reality. These devices have become very popular, reaching a massive portion of the technology market. These products are: smart phones, tablets-pc and tablets (Figure 3-9). All they incorporate touch-sensitive screen.

a) Smart-phone b) Tablet PC c) Tablet

Figure 3-9 Touch-sensitive screen devices than can be used for signature acquisition

These new devices, due to their remarkable widespread coverage, are expected to play an important role in the near future of signature verification systems.

3.3 COMPARISON ALGORITHMS

The intensive work being carried out over the past decades to achieve reliable automatic signature verification systems has been mentioned earlier in this Thesis. Since the first published work on automatic signature verification was released by Mauceri in 1965 [30], many researches from public institutions and industry have contributed to the development of this interesting biometric modality.

Many different techniques have been used as comparison algorithms, from point to point techniques to genetic algorithms. The majority of this work has been summarized in several state-of-the-art publications. Among the many publications, it is worth highlighting those published by Plamondon in 1989 [34], in 1994 [35] and 2000 [36]. Also, a particularly interesting review of the work done on signature biometrics was published by Gupta in 2006 [40]. Recently, in 2008, Donato Impedovo and Giuseppe Pirlo published a complete and extensive state-of-the-art [41] in which the main techniques used for automatic signature verification, both offline and online, are detailed and thoroughly referenced (approx. 400 references).

The signature comparison algorithms used to decide whether a signature sample belongs to the claimed identity are generally split into two main groups, these are: the distance based approach and the model based approach [20].

3.3.1 DISTANCE BASED APPROACHES

In this approach, a reference signature is created from the signature sample enrolment set. Then the distance between this reference signature and the sample signature to verify is computed. To do this, two approaches are generally taken: a point-to-point comparison between signature sample points acquired, or extraction of the features from the temporal signals acquired in order to compare feature vectors.

In the first approach, Dynamic Time Warping (DTW) has generally been used, and is one of the most utilized techniques for signature comparison in published works. The DTW performs an elastic alignment between the user reference signature and the questioned signature. In this way, the DTW minimizes the intrinsic personal variability of the signing process. The first use of the DTW for signature verification came from Sato and Kogure in 1982 [49]. After this work, many researchers have used DTW based comparison algorithms [50-61]. In order to improve the performance of the DTW, different techniques have been used such as Genetic Algorithms [62], Principal Component Analysis (PCA) [59], minor component analysis (MCA) [63] and extreme points [64]. Another technique used is the stroke based DTW [65]. A DTW algorithm has been implemented in this Thesis and will be explained in detail in the next chapter.

Other signal comparison techniques that have been investigated are regional correlation [31, 66] and skeletal tree matching [67-68]. Even though several comparative studies do not show the DTW to perform as well as the aforementioned techniques [69], they have not been used as much as the DTW.

It is worth highlighting that DTW based algorithms were the winning signature systems in two public evaluation campaigns SVC'2004 [22] and BSEC'2009 [70].

Regarding the use of feature vectors to compute the distance between two signatures, classical metrics has been used such as the Euclidean Distance [71-72], the Mahalanobis Distance [73-74] and the City Block Distances [20].

3.3.2 Model Based Approaches

In model based approaches a user reference model is created through statistical methods such as a Hidden Markov Model (HMM), Gaussian Mixture Model (GMM) or by using Neural Networks (NN). These statistical approaches are very common and well referenced for pattern recognition.

The most commonly used statistical model for pattern analysis is the HMM. This is a double stochastic approach. The HMM is made of a hidden stochastic unobservable process which corresponds to the transition between different states and an observable stochastic process whose outputs are a symbol sequence. The basic HMM theory was introduced by Baum [75], where an interesting introduction to HMM can be found in [76]. One of the first and main published works using HMM applied to signature authentication has been presented by Dolfing [77-78]. Different HMM topologies have been adopted, where the left-to-right topology is largely used among researchers [78-80]. The Ergodic topology has also been investigated [81]. In 1995, L. Yang and his collaborators [82] tested different HMM topologies, showing that the best approach for signature characteristics was the left-to-right topology.

Another statistical model successfully used for signature verification is the GMM. Several authors consider this technique as a degenerated version of the HMM with only one state [20]. Jonas Richiardi and Andrzej Drygajlo [83] were the first authors to use the GMM for online signature verification in 2003. Since then, different authors have investigated its use on signature authentication systems [20, 84-85]. This technique has also been used in this Thesis and will be explained in detail in the following chapter.

Apart from these two statistical models, other model approaches based on Neural Networks (NN) have also been used for signature verification tasks. Several NN topologies such as multilayer perceptrons [86-89], time delays [90-91], Bayesian [91-92], input-oriented [91] and radial basis functions [93-94] have been used.

A new promising approach for on-line signature authentication based on Support Vector Machines (SVM) has been introduced in recent years. The SVM is a new classification technique and forms part of the statistical learning theory field, and has been applied successfully for pattern recognition applications. It has also been applied to signature verification [95-99].

3.4 EVALUATION DATABASES

The evaluation of the automatic signature system's performance, presented in this Thesis, has been performed on different databases. Because of the importance of these databases, this section is dedicated to their review. The most well-known and used databases are listed below:

- Philips Signature Database [77-78],
- Biomet Signature Subcorpus (DS2 and DS3) [100],
- SVC2004 Development Set [22],
- MCyT Signature Subcorpus [21],
- MyIDea Signature Subcorpus [101],
- BioSecure Signature Subcorpus [102].

The use of these readily available databases ensures that the results obtained by different researchers may be directly comparable. This also allows the retrieval of contributions from authors that are considered as state-of-the-art. Several of the aforementioned databases are publicly available.

In the following subsections, the signature databases used in this Thesis will be detailed where a comparative summary of their characteristics is presented.

3.4.1 MCYT SIGNATURE SUBCORPUS

The MCyT [21] signature subcorpus was completed in late 2003. It forms part of a bimodal biometric database, where fingerprints were also acquired. The Biometric Research Laboratory from the Universidad Politécnica de Madrid lead the project of collecting this signature, and was carried out in collaboration with 3 other Spanish universities, these are: Universidad de Valladolid, Universidad del País Vasco and Escuela Universitaria Politécnica de Mataró. As a result of this work, the MCyT database is, by in large, composed of Spanish user signatures.

The full database is formed from 330 users (MCyT-330), although the public dataset contains only 100 users (MCyT-100). The database was captured using an Intuos 2 A6 Wacom digital tablet, acquiring the following signals (value ranges are shown in parenthesis) at a sampling rate of 100 Hz:

- X axis position (0-12700),
- Y axis position (0-9700),
- Pressure (0-1023),
- Azimuth angle (0-360),
- Inclination angle (0-90).

The signature subcorpus is comprised of 25 genuine signatures from each user. Also, 25 forgery signatures from different users were collected (5 forgery signatures from 5 different users). The forger had a static image of the signatures to imitate, and had the opportunity to perform several practice attempts, enhancing their confidence.

This database has become the most commonly used in handwritten signature literature, because of the two aforementioned aspects, i.e. publically available and the large amount of signature samples.

3.4.2 SVC'2004 DEVELOPMENT SET

The SVC'2004 database [22] was released in 2003 to assist participants develop and test algorithms for the first signature verification contest (SVC'2004). The capture device was a Wacom Intuos 2 A6 tablet, acquiring x and y position, pressure and 2 angles (inclination and azimuth) from each user. Every user contributed 20 genuine signatures during 2 different sessions. For privacy reasons, the signers were advised not to use their real signatures. Instead, they were recommended to design a new signature specifically for the database. Another 20 forgery signatures were provided by at least four other users. In order to forge the signatures, forgers had the opportunity to see the genuine signatures to be copied using a software application. This software application allowed forgers to replay the writing sequence of a signature. Forgers were advised to practice skilled forgeries until they were confident of their reproduction.

This database contains signatures from Chinese users, who could choose to sign in either Latin characters (chosen by 24 users) or Chinese characters (chosen by 16 users).

3.4.3 MYIDEA SIGNATURE SUBCORPUS

The MyIDea multimodal biometric database [101] was sponsored by the Swiss National Center of Competence in Research during their participation in the Interactive Multimodal Information Management project IM2 [103] and the European IST BioSecure project [104]. Data collection was supervised by the Document, Image and Voice Analysis group from the computer science department at the University of Fribourg.

Data collection started in 2004 and finished in late 2005. The database contains data from 73 users from whom 46 users signed in French and 27 signed in English. For each user a total of 18 genuine signatures were acquired. At the same time 18 static forged signatures and 18 dynamic forged signatures were collected for each user. For the static forge signatures, the forgers were supplied with a paper-copy of the signature to forge and were allowed to train for several minutes to gain confidence. For the dynamic forged signatures, the forgers were able to use dedicated on screen software to study the signatures dynamics.

The capture device used was a Wacom Intuos 2 A4 tablet. Again, five different signals were recorded, x and y position, pressure, elevation and azimuth. The sample rate used was 100 Hz.

3.4.4 DATABASES SUMMARY

In this section, a brief summary of the databases presented above is provided. The 3 databases used in this Thesis: MCyT, SVC 2004 and MyIDea have been split into 5 different subsets, which have considered the language and character set used by the signer.

Table 3-2 Summary of Signature Databases

Dataset	Language	Character Set	# User	# Genuines	# Forgeries	# Signatures
MCyT	Spanish	Latin	100	25	25	5000
SVC 2004 ORI	Chinese	Chinese	24	20	20	960
SVC 2004 OCC	English	Latin	16	20	20	640
MyIDea FR	French	Latin	46	18	36	2484
MyIDea EN	English	Latin	27	18	36	1458

It may be observed in Table 3-2, that the MCyT is the largest dataset in both number of users and number of genuine and forgery signature samples for each user. The SVC datasets has a small number of users; also it must be considered that the signatures are not real. The MyIDea datasets has the smallest number of genuine signature samples and has a low x and y axis resolution (see Table 3-3).

Table 3-3 Capture Details for the Signature Databases

Dataset	Input Device	Sample Rate (Hz)	X Resolution (points/mm)	Y Resolution (points/mm)
MCyT	Intuos 2 A6	100	100	100
SVC 2004	Intuos 2 A6	100	100	100
MyIDea	Intuos 2 A4	100	3.5	3.6

Several of the average basic global features are presented in Table 3-4 and also in Figure 3-10 to Figure 3-13. The features calculated for the datasets are:

- *Length*: average total length of the signature samples. calculated as the sum of the Euclidean distance between all the signature points as indicated in the following equation:

$$Length = \sum_{i=1}^{N-1} \sqrt{(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2}$$

- *Total number of points*: average total number of points acquired and recorded in the database.
- *Total time*: average time elapsed between the first pen-down and the last pen-up.
- *Velocity*: mean signature velocity, calculated as the ratio between the length and the total time:

$$Velocity = \frac{length}{time}$$

- *Number of strokes*: average number of pen-down events of the signature samples.

Table 3-4 Average Features of the Signature Datasets

Dataset	Language	Character Set	Length (mm)	Points	Time (s)	Velocity (mm/s)	Strokes
MCyT	Spanish	Latin	187	350	3.49	68	6.5
SVC 2004 ORI	Chinese	Chinese	339	208	2.86	121	7.6
SVC 2004 OCC	English	Latin	235	168	2.06	117	4.5
MyIDea FR	French	Latin	175	296	2.95	70	4.9
MyIDea EN	English	Latin	209	268	2.67	86	3.8

The average feature values present interesting results that are worth highlighting. All the Latin character set based signatures have similar average lengths, approximately 200 mm (Figure 3-10). The main difference amongst them is the average time required to perform the signature (Figure 3-11). The Spanish MCyT signatures have the greatest performing time, taking an average of 3.49s to perform a signature. On the other hand SVC2004 Latin based signatures are the fastest, with an average of only 2s. Chinese character based signatures collected in the SVC database have the longest average length, 340 mm (Figure 3-10). They also demonstrate the fastest velocity, 120 mm/s (Figure 3-12), and the greatest number of strokes, 7.5 strokes. These facts come from the particularity of Chinese users when signing as they use short and fast small strokes.

In terms of the number of strokes, as mentioned in the previous paragraph, the Chinese character set based signature samples have the greatest number. When observing the Latin character set based signature samples, the MCyT stands out with 6.5 average strokes. This is

due to the fact that Spanish signatures use more pictorial strokes than other Latin based writers, whom normally sign with just their name and do not add pictorial strokes [41].

In this introduction to the public signature databases used in this Thesis has been shown how SVC'2004 and MyIDea were not good candidates to serve as a reference when comparing with other signature systems. SVC'2004 has a lack of quality signatures as its genuine signatures are not real ones. This fact, along with the two very difference sorts of signatures that it contains, make this databases be very sensitive to the algorithm tested. On the other hand MyIDea database is not easy to be found on the signature literature, due to its novelty and also to its lack of resolution on coordinates x and y.

On the contrary, MCyT has the greatest number of users and signature collected and it has been extensively used throughout signature literature. At the same time the quality of the signature acquired is high and the country origin of the signatures match with our researching group interests. Due to these facts, this database will be used to develop and enhance the signature algorithms analyzed within this work.

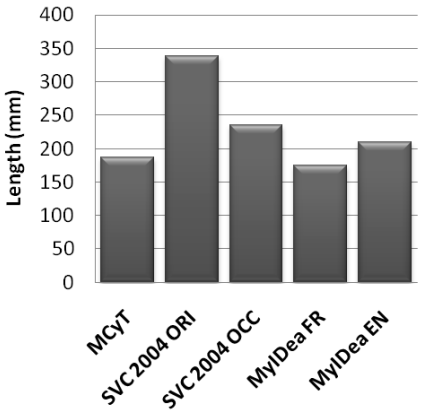


Figure 3-10 Average Dataset Lengths

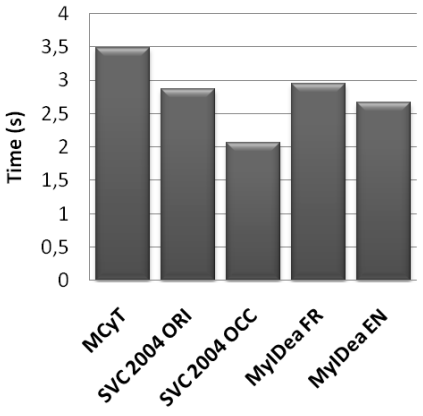


Figure 3-11 Average Dataset Total Times

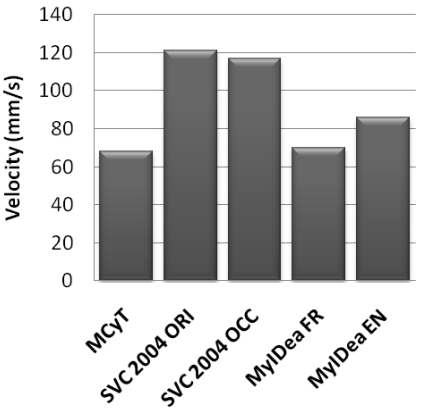


Figure 3-12 Average Dataset Velocity

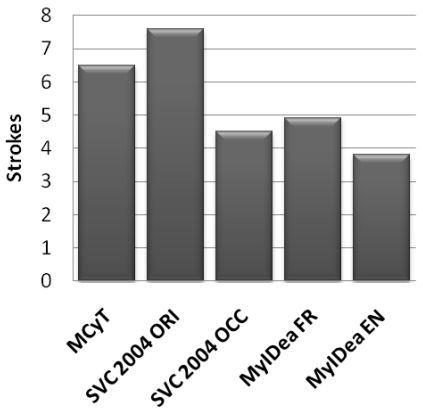


Figure 3-13 Average Dataset Number of Strokes

3.5 EVALUATION CAMPAIGNS

An Evaluation campaign or competition is an independent and reliable way to assess state-of-the-art biometric performance. Major biometric modalities have specific competitions, such as the FVC (held in 2004, 2006 and 2008) and FpVTE (held in 2003) for fingerprint, ICE (held in 2005 and 2006) and IREX for Iris (IREX I held in 2009, IREX II in 2010 and IREX III in 2011), NIST-SRE (held from 1997 to 2010) for speaker recognition, FRGC (held in 2005) and FRVT (2002) for face recognition systems.

There have been two evaluation campaigns for signature verification. The first took place in 2004 and was named Signature Verification Contest (SVC'2004) and the most recent, in 2009, was named the BioSecure Signature Evaluation Campaign (BSEC'2009). Their results and main conclusions will be presented in the following sections.

3.5.1 SVC 2004

The SVC'2004 was the first signature evaluation campaign, held in conjunction with the First International Conference on Biometrics Authentication (ICBA 2004).

The competition was open to both academy and industry. Two scenarios were proposed: Task1 and Task2. The first scenario, Task1, was based on a mobile scenario, using a personal digital assistant (PDA) as the input device. In this scenario, signature data only contained coordinate information (x and y position and time). The second scenario, Task2, was based on a normal office situation using a Wacom Intuos 2 A6 digital tablet as the input device. In this case, the signature data contained additional information on the pressure and pen orientation (pen azimuth and elevation).

The database contained 100 users, of these 40 were released to allow participants to fine tune their algorithms before submission. The other 60 users formed part of the evaluation set. The evaluation methodology proposed used 5 random genuine signature samples from each user during the enrolment process, even though different sets were used in multiple runs (10 trials were run based on 10 different enrolment sets). After each enrolment trial, the program was evaluated using 10 genuine signatures, 20 skilled forgeries, and 20 random forgeries selected from genuine signatures of 20 other users.

A total of 15 teams for Task1 and 12 teams for Task2 submitted their algorithms. The following table contains more information on the teams and Tasks undertaken:

Table 3-5 SVC2004 Participating Teams [22]

Team ID	Institution	County	Member(s)	Task(s)
3	Anonymous	Australia	V. Chandran	1 & 2
4	Anonymous	Turkey		1 & 2
6	Sabanci University		Alisher Kholmatov Berring Yanikoglu	1 & 2
8	Anonymous			2
9	Anonymous			1 & 2
12	Anonymous			1
14	Anonymous			1 & 2
15	Anonymous			1
16	Anonymous			1
17	Anonymous			1 & 2
18	Anonymous			1 & 2
19	Biometrics Research Laboratory, Universidad Politecnica de Madrid	Spain	Julian Fierrez-Aguilar Javier Ortega-Garcia	1 & 2
24	Fraunhofer, Institut Sichere Telekooperation	Germany	Miroslav Skrbek	1
26	State University of New York at Buffalo	USA	Aihua Xu Sargur N. Srihari	1
29	Institut National des Telecommunication	France	Bao Ly Van Sonia Garcia-Salicetti Bernadette Dorizzi	2

The results of both tasks are summarized in Table 3-6 and Table 3-7. The Team ID 6 (Sabanci University) achieved the best EER in both tasks when testing skilled forgeries. In Task1 they achieved a 2.84% EER and in Task2 a 2.89% EER. The algorithm used was based on the DTW [59]. For random forgeries the best EER achieved was 1.85% in Task 1 from Team ID 24 (Franhoufer Institute) and 1.70% from Team ID 19a (Biometric Research Laboratory Universidad Politécnica de Madrid).

There are several areas of these results that should be highlighted. The first relates to the large variability in the EER obtained from the different trials using different enrolment sets, this can be observed when looking at the standard deviation values. This occurrence was explained in the SVC'2004 report [22], where it was stated that the signature samples collected for the database from the testing crew were not real handwritten signatures. Another area which can be highlighted is the slightly better results achieved in Task1 when compared with those in Task2. This result implies that the information provided in the additional information coming from the pressure and pen orientation data are not useful for verification tasks. This fact contradicts several studies on signature features such as [105-106]. Within the competition report, it has been indicated that the reason for this was again due to the way the signature samples were collected.

Table 3-6 SVC'2004 EER Results for Task1

Team ID	10 genuine signatures + 20 skilled forgeries			10 genuine signatures + 20 random forgeries		
	Average	SD	Maximum	Average	SD	Maximum
6	2.84%	5.64%	30.00%	2.79%	5.89%	50.00%
24	4.37%	6.52%	25.00%	1.85%	2.97%	15.00%
26	5.79%	10.30%	52.63%	5.11%	9.06%	50.00%
19b	5.88%	9.21%	50.00%	2.12%	3.29%	15.00%
19c	6.05%	9.39%	50.00%	2.13%	3.29%	15.00%
15	6.22%	9.38%	50.00%	2.04%	3.16%	15.00%
19a	6.88%	9.54%	50.00%	2.18%	3.54%	22.50%
14	8.77%	12.24%	57.14%	2.93%	5.91%	40.00%
18	11.81%	12.90%	50.00%	4.39%	6.08%	40.00%
17	11.85%	12.07%	70.00%	3.83%	5.66%	40.00%
16	13.53%	12.99%	70.00%	3.47%	6.90%	52.63%
4	16.22%	13.49%	66.67%	6.89%	9.20%	48.57%
12	28.89%	15.95%	80.00%	12.47%	10.29%	55.00%

Table 3-7 SVC'2004 EER Results for Task2

Team ID	10 genuine signatures + 20 skilled forgeries			10 genuine signatures + 20 random forgeries		
	Average	SD	Maximum	Average	SD	Maximum
6	2.89%	5.69%	30.00%	2.51%	5.66%	50.00%
19b	5.01%	9.06%	50.00%	1.77%	2.92%	10.00%
19c	5.13%	8.98%	51.00%	1.79%	2.93%	10.00%
19a	5.91%	9.42%	50.00%	1.70%	2.86%	10.00%
14	8.02%	10.87%	54.05%	5.19%	8.57%	52.63%
18	11.54%	12.21%	50.00%	4.89%	6.65%	45.00%
17	12.51%	13.01%	70.00%	3.47%	5.53%	30.00%
4	16.34%	14.00%	61.90%	6.17%	9.24%	50.00%

The last fact to highlight is the high EER obtained with random forgeries. This is particularly true for the best algorithm (team ID 6). Using this algorithm, the EER achieved for skilled and random forgeries are approximately equal. Random forgeries are expected to achieve improved error rates, as a result of the comparative analysis of two signatures from different users, where these signatures are generally very different, at least from a graphical point of view.

3.5.2 BioSecure Reference Benchmarking Framework For Signature Verification

The BioSecure Reference Benchmarking Framework for signature verification aims to become the baseline framework for comparative signature verification systems. It is composed of 3 different publicly available elements [107]:

- Two open-source reference signature verification algorithms. The first open-source reference system is based on the HMM [80] while the second is based on the Levenshtein distance which transforms the signature sample data into a sequence of characters [108].
- Two publicly available databases: BioMet Signature Subcorpus [100] and MCyT-100 Subcorpus Database [21].
- Benchmarking experimental protocols for both databases. The benchmarking framework is explained in the book "Guide to Biometric Reference Systems and Performance Evaluation" [20].

Making all these elements publicly available can serve further comparative analysis of newly proposed research systems.

The BioSecure reference benchmarking framework for signature verification presented the evaluation results of its two open-source reference systems along with other research systems provided by different institutions. The algorithms evaluated are detailed in [20], and are summarized in the following table:

Table 3-8 Systems evaluated on BSEC'2009 [20]

System ID	Description	Algorithm
Ref1	Reference System 1: Fusion of Viterbi path and Likelihood score [80]	HMM
Ref1-Vit	Reference System 1: Viterbi path [80]	HMM
Ref1-Lik	Reference System 1: Likelihood score [80]	HMM
Ref2	Reference System 2: Levenshtein distance [108]	Levenshtein Distance
UAM	Functional Feature extraction and Hidden Markov Model [109]	HMM
GMM	Gaussian Mixture Model with a local feature extraction [110]	GMM
DTWstd	Dynamic Time Warping based system [111]	DTW
DTWnorm	Dynamic Time Warping based system with Score normalization	DTW
Globalappr	Global Features and City Block Distance	City Block Distance

A full description of the results for the BioMet database can be found in [20]. The results of the publicly available MCyT Signature Subcorpus (MCyT-100) will be presented in this Thesis, as this database was used to evaluate the algorithms proposed.

The benchmarking experimental protocol for MCyT-100 is very similar to that used in the SVC'2004. It uses 5 randomly selected genuine signature samples from each user during the enrolment process. 100 different sets are evaluated to reduce the influence of the five enrolment signatures selected. After each enrolment trial, the remaining 20 genuine signatures, 25 skilled forgeries, and 99 random forgeries are selected from the genuine signatures available for each user.

The results obtained for all these systems are presented in Table 3-9. The best system for both skilled and random forgeries was the Ref1 system, based on HMM, obtaining an EER of 3.41% for skilled forgeries and an EER of 0.95% for random forgeries. This system is followed by DTWnorm, which obtained EERs of 3.91% and 1.20% for skilled and random forgeries respectively.

Table 3-9 EERS of the systems on the MCyT-100 database and their Confidence Interval (CI) of 95%

Skilled Forgeries			**Random Forgeries**		
System ID	**EER %**	**CI 95 %**	**System ID**	**EER %**	**CI 95 %**
Ref1	3.41	±0.05	Ref1	0.95	±0.03
DTWnorm	3.91	±0.07	DTWstd	1.20	±0.06
UAM	5.37	±0.08	DTWnorm	1.28	±0.04
Ref1-Vit	5.59	±0.07	Ref1-Lik	2.13	±0.05
Ref1-Lik	5.66	±0.07	UAM	2.34	±0.05
DTWstd	5.96	±0.09	Ref1-Vit	2.44	±0.04
GMM	6.74	±0.09	GMM	2.81	±0.05
Globalappr	7.23	±0.10	Globalappr	3.15	±0.07
Ref2	10.51	±0.13	Ref2	4.95	±0.09

The results of the MCyT-100 Signature Subcorpus are more stable than the results reported in the SVC'2004, this may be concluded by observing the low values of the CI. This fact supports the assumptions made in the SVC'2004 report on the variability introduced by the method of collecting the non-real signatures for the SVC'2004 database. Also, the bigger number of users in the MCyT-100 database increases the reliability of the results.

3.5.3 BIOSECURE SIGNATURE EVALUATION CAMPAIGN 2009 (BSEC'2009)

Following the success of the SVC'2004 campaign, the BioSecure Signature Evaluation Campaign (BSEC'2009) was aimed at expanding the objectives of the evaluation. The BSEC'2009 was executed using a subset from the largest signature database known at that time: BioSecure Signature Corpus [102]. This database contains data acquired from the same user from two different devices, a digital table as an office scenario (Dataset 2, DS2) and a Personal Digital Assistant (PDA) as a mobile scenario (Dataset 3, DS3), and contains the signatures of the same 382 users. The DS2 was acquired using a digital table Wacom Intuos 3 A6 which operates at a frequency of 100Hz and has the following acquisition characteristic: 5080 lines per inch, capture area of 270 mm x 216 mm, 1024 pressure levels, and a tilt accuracy for inclination and azimuth of 2°. The DS3 was acquired using a PDA HP iPAQ hx2790 operating at a frequency of 100Hz and with a touch screen resolution of 1280*960 pixels. Only the x and y coordinates were acquired along with the time elapsed between the acquisition of two successive points. In order to recreate the mobility conditions, the participants were asked to sign standing. A total of 30 genuine and 20 forgery signatures were captured in 2 different sessions spaced by approximately 5 weeks. For further acquisition protocol details see[112].

The main objectives of this evaluation campaign can be found on the website [112], and are the following three:

- To measure the real impact of mobile acquisition conditions on algorithms' performance, by using the same test crew, large enough, for both input devices. These datasets are available through the Association BioSecure [104].
- To evaluate the impact of time variability, making available a subset of 50 users samples acquired only in the first session, while evaluating both sessions separately.
- To measure the impact of the information content of signatures on algorithms' performance, thanks to a protocol categorizing the data in both DS2 and DS3 in subsets. This impact is measured using the notion of Client-Entropy to categorize users depending on the quality of the signature (Complexity, Variability) [113]. Performance will also be measured on the complete databases for comparison purposes.

In order to measure the results of the algorithms for the abovementioned three objectives, three different evaluation protocols were defined. All of these are based on a generic protocol which implies the use of 10 trails of 5 random signature sets from session 1 as enrolment signatures:

- Measurement of the impact of the mobility acquisition conditions on the algorithm's performance. The results of the signature evaluation from session 1

only, were compared for both datasets. In this evaluation, only pen coordinates were considered.

- Measurement of the impact the time variability has on the algorithm's performance. The results of the evaluation signatures were compared for the different sessions, DS2 and DS3 separately. In the case of DS2, the influence of the signature data involved in the acquisition was evaluated. Three different configurations of input time functions were considered: one with x and y coordinates only, others with x and y coordinates and pressure, and the last one with x and y coordinates, pressure and tilt data.
- Measurement of the impact the information content has on the algorithm's performance. The organizer tested all the systems submitted for the previous evaluations of the DS2 and DS3 separately for different user categories. This procedure was dependent on the quality of the signatures [113].

The results of these evaluations are available on the BSEC'2009 website [112]. Nine universities registered for this evaluation campaign, submitting 14 different systems, detailed in Table 3-10. The third column indicates which of the 3 evaluations (numbered 1, 2, 3) each system participated in.

Table 3-10 Systems evaluated on BSEC'2009

University	System	Evals	Matching Algorithm
Escola Universitaria Politecnica de Mataro (Spain)	Sys1	1 2 3	Biometric Dispersion Matcher
	Sys2	1 2 3	Ratio of means and standard deviation of parameters on each piece of signature
U1 Research Institute (Hungary)	Sys3	1 2 3	DTW
Seikei University (Japan)	Sys4	1 2 3	DTW
Ain Shams University (Egypt)	Sys5	1 2 3	DTW
University of Valladolid (Spain)	Sys6	1 2 3	DTW
Sabanci University (Turkey)	Sys7	1 2 3	DTW
Universidad Autonoma de Madrid (Spain)	Sys8	1 2 3	DTW tuned for random forgeries
	Sys9	1 2 3	DTW tuned for skilled forgeries
	Sys10	1 2 3	HMM
	Sys11	1 2 3	Global and Mahalanobis distance
	Sys12	1 2 3	Fusion of 4 systems: Sys8, Sys9, Sys19, and Sys 11
Waseda University (Japan)	Sys14	2	DTW
University of Magdeburg (Germany)	Sys 15	2	Biometric Hash Algorithm, Canberra distance
Reference System TMSP (France)	---	1 2 3	HMM

Evaluation 1: impact of mobility acquisition conditions

The results obtained for all these systems when considering the two different scenarios, office scenario (DS2) and mobile scenario (DS3) are presented in Table 3-11. The best error rates obtained are highlighted in bold.

Table 3-11 EERS of the systems for Session 1 for the DS2 and DS3 datasets, for skilled and random forgeries

System ID	DS2		DS3	
	Skilled	**Random**	**Skilled**	**Random**
Sys1	4.40	1.85	8.18	2.05
Sys2	4.91	2.33	7.38	1.86
Sys3	13.99	8.98	18.32	8.36
Sys4	2.88	1.58	7.87	1.29
Sys5	3.82	2.67	31.57	30.64
Sys6	2.20	0.97	6.58	1.65
Sys7	2.98	2.23	4.99	4.32
Sys8	4.18	0.51	12.20	0.55
Sys9	2.88	1.47	5.77	1.54
Sys10	19.23	24.14	25.85	21.34
Sys11	6.71	3.31	13.26	4.7
Sys12	2.23	0.63	5.47	0.66
Ref Sys	4.47	1.74	11.27	4.8

As it was expected, the office scenario obtains superior results when compared to the mobile scenario. The performance is improved by a factor of approximately 2 for the EER when using skilled forgeries. The random forgery EERs are only slightly inferior in most cases, except for the systems Sys5, Sys7 and the Ref Sys. It is worth pointing out that the best systems for random forgeries are poor at discriminating skilled forgeries.

Comparing these results with the previous public competition, i.e. SVC'2004, it has been observed that the EERs are similar, where the best systems achieved an EER of less than 3% for skilled forgeries. However, for random forgeries there was an improvement in the EERs, where the best systems rated at close to 1% lower than in the SVC'2004 competition. In this new competition there is a noticeable difference between skilled and forgery EERs, which surprisingly, was not appreciable in the SVC'2004.

Evaluation 2: impact of time variability

Here the influence of time variability was evaluated using separate signatures from session 1 and session 2. At the same time, as mentioned above, three different configurations were evaluated to measure the impact the input device has on the signature data acquired. The influence of adding both pressure and tilts was analysed for the office scenario.

In Table 3-12 the results are presented for the different systems in session 1 and 2 separately, using only the x and y coordinate information captured by the digital tablet (office scenario).

Table 3-12 EERS of the systems for Session 1 and 2 using the DS2 dataset, skilled and random forgeries

System ID	DS2 – SESSION 1		DS2 – SESSION 2	
	Skilled	Random	Skilled	Random
Sys1	4.40	1.85	7.15	4.40
Sys2	4.91	2.33	6.20	4.02
Sys3	13.99	8.98	19.03	12.29
Sys4	2.88	1.58	5.99	3.55
Sys5	3.82	2.67	6.61	5.23
Sys6	2.20	0.97	4.21	2.24
Sys7	2.98	2.23	5.13	3.96
Sys8	4.18	0.51	7.26	1.80
Sys9	2.88	1.47	4.08	2.93
Sys10	19.23	24.14	20.47	25.62
Sys11	6.71	3.31	10.92	5.37
Sys12	2.23	0.63	4.18	1.70
Ref Sys	4.47	1.74	5.99	3.16

As expected, considering that the algorithms were tuned during Session 1, the results are seen to be inferior in the presence of time variability. This degradation is more or less homogenous for all the different systems tested.

The same outcome is obtained for the evaluation of the mobile scenario (DS3) as only the x and y coordinates are considered. These results are presented in the following Table 3-13.

Table 3-13 EERS of the systems for Session 1 and 2 using the DS3 dataset, skilled and random forgeries

System ID	DS2		DS3	
	Skilled	Random	Skilled	Random
Sys1	8.71	2.22	14.24	3.94
Sys2	7.38	1.85	11.25	3.76
Sys3	18.32	8.36	24.68	12.40
Sys4	6.37	2.00	9.43	3.72
Sys5	7.55	4.24	11.51	6.78
Sys6	5.69	1.50	8.06	2.90
Sys7	4.98	4.31	7.69	7.02
Sys8	10.40	0.70	14.51	1.67
Sys9	5.24	2.09	7.42	2.83
Sys10	24.79	27.29	23.52	26.81
Sys11	10.49	2.93	15.00	5.01
Sys12	4.93	1.41	7.42	1.93
Sys13	5.98	1.44	9.93	3.48
Ref Sys	11.27	4.8	14.03	6.06

The results obtained by adding the pressure information for the office scenarios (DS2) to the systems is presented in Table 3-14. The EER is only slightly improved when no time variability is present. When the time variability is taken into consideration (session 2), the EERs do not improve with the addition of pressure.

Table 3-14 EERS of the systems in Session 1 and 2 for the DS2 where pressure information is added, skilled and random forgeries

System ID	DS2		DS3	
	Skilled	Random	Skilled	Random
Sys1	4.03	1.70	6.88	4.16
Sys2	4.50	1.96	6.28	3.61
Sys3	13.69	8.62	18.54	11.81
Sys4	2.76	1.33	6.07	3.42
Sys6	2.19	0.97	4.21	2.23
Sys8	3.26	0.42	6.21	1.37
Sys9	2.38	1.17	3.48	2.46
Sys10	27.76	20.51	30.13	21.61
Sys11	5.90	2.02	9.52	3.65
Sys12	1.71	0.65	3.49	1.46
Sys13	2.84	1.38	5.10	3.19
Ref Sys	4.07	1.65	5.32	2.96

In Table 3-15 the results are shown for the office scenario evaluation when all the information acquired by the digital tablet is considered (i.e. x and y coordinates, pressure, elevation and azimuth). In this case, the information provided by the new time series, i.e. elevation and azimuth, is not observed to improve for both possible cases of the time variability, session1 and session 2.

Table 3-15 EERS of the systems for Session 1 and 2 using the DS2 where pressure and tilts information is added, skilled and random forgeries

System ID	DS2		DS3	
	Skilled	Random	Skilled	Random
Sys2	4.52	1.91	5.99	3.53
Sys3	13.41	8.63	17.91	11.77
Sys4	3.02	1.49	6.02	3.52
Sys6	2.19	0.97	4.21	2.23
Sys13	17.94	24.06	19.34	25.61
Sys14	4.82	1.98	8.73	4.24
Ref Sys	4.07	2.39	5.72	3.87

Evaluation 3: impact of information content

The effect that the information content in signatures has on the algorithms performance was only evaluated for the office scenario (DS2). The users contained in this dataset were separated into two different categories (high entropy and low entropy), depending on the signature quality. This process was carried out using the Personal Entropy Measure presented in [113].

The ERRs achieved from the different systems are presented in Table 3-16, where it may be observed that there is a significant difference in the error rates for the two entropy levels. Those users who have a low level of entropy in their signatures (which can be understood as more complex and stable signatures) present lower error rates. However, as expected, users with high levels of entropy show higher error rates. There is only one exception, Sys 10, which obtains much higher error rates for the lower entropy levels.

Table 3-16 EERS of the systems in Session 1 for DS2 and for each writer category, skilled and random forgeries

System ID	DS2 – High Entropy		DS2 – Low Entropy	
	Skilled	Random	Skilled	Random
Sys1	6.50	2.33	3.94	1.50
Sys2	6.58	2.83	4.57	1.80
Sys3	14.00	7.22	14.50	9.98
Sys4	4.08	1.52	2.92	1.38
Sys5	5.67	2.55	3.14	2.47
Sys6	3.75	0.83	1.68	0.87
Sys7	4.00	1.61	2.89	2.27
Sys8	7.83	0.80	2.95	0.27
Sys9	4.17	1.19	2.48	1.42
Sys10	9.92	11.27	21.18	32.43
Sys11	9.00	3.83	6.83	3.14
Sys12	4.17	0.91	1.49	0.62
Ref Sys	6.00	1.52	3.81	1.62

3.6 CURRENT DATA FORMAT STANDARDS FOR SIGNATURE BIOMETRICS

The distributed nature of the market (input device vendors, algorithm providers, integrators and end-users) shows the importance of standardization, enabling all stakeholders to develop systems which easily interact with each other.

The ISO/IEC JC1 SC37 WG3 is already working on the development of a second generation biometric data interchange standards (Project 19794) and within this work there are two standards for handwritten signature biometric data:

- Part 7: Signature/sign time series data
- Part 11: Signature/sign processed dynamic data

The following sections of this chapter will introduce these two signature standards, along with a viability analysis of the work carried out in this Thesis. In the following chapters the contributions made as a result of this viability analysis to develop and improve on these standards will be explained.

3.6.1 2ND WORKING DRAFT 19794, BIOMETRIC DATA INTERCHANGE FORMATS – PART 7: SIGNATURE/SIGN TIME SERIES DATA

The ISO/IEC JTC1 SC37 W3 experts are currently developing a revision of the ISO/IEC IS 19794-7:2007 [114], which will be the second generation of the signature/sign time series data format. This new version is currently in the Working Draft (WD) state [115] (in order to reference this standard, 19794-7.2WD2 notation will be used, ".2" denotes the second generation of this signature international standard, and "WD2" denotes its second working draft). The version under study and presented in this Thesis has been its second working draft and released in 2010. The main changes from the first generation are located within the Full Format, where completely new versions of the General Header and Representation Header have been introduced. This new generation also incorporates conformance test assertions which can be found within Annex A. The conformance testing methodology has been part 7 of the 29109 project [116] for the previous generation of 19794-7 (ISO/IEC IS 19794-7:2007) [114].

The most important modification within this second generation is the inclusion of a new subformat called Compression Format. The decision to incorporate this compressed data format was taken as a result of the work presented in [117], which summarizes several of the outcomes of this PhD Thesis. The viability analysis, introduced in chapter 5, and the

conclusions taken from chapter 6 have both motivated the inclusion of data compression using lossless compressions algorithms for signature data formats.

The following paragraphs provide a brief description of the signature data formats included in this new generation of 19794-7 (19794-7.2WD2) highlighting the key differences regarding the first generation.

Part 7 of the 19794 project defines how data captured by a signature input device, in the form of time-series raw data, has to be stored in order to achieve interoperability amongst different biometric systems and/or applications.

The raw data is divided in different channels. Allowed channels are: x and y position (X, Y), velocity (VX, VY), acceleration (AX, AY), z position (Z), time (T) and time difference (DT), pen tip force (F), switch state (S) and pen orientation (TX, TY, Az, El, R).

Two different formats are already defined within this part 7 [115] (19794-7.2WD2). The first one is the Full Format for general use. The second is the Compact Format, which is used in applications where the size of the biometric record is an important issue, this occurs with smart cards and other token formats. Both of these will be explained in detail in the following subsections.

3.6.1.1 FULL FORMAT

As is defined in the second generation of the part 1 “framework” of the 19794 project [28], which is at the Final Draft International Standard stage (FDIS) being released in 2011 (referenced as 19794-1.2FDIS), the full format for signature time series data is capable of recording data from multiple samples, where each one is contained in a separate “representation” record, following the structure shown in Figure 3-14:

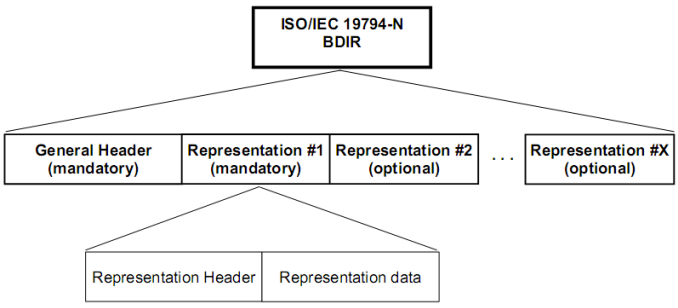


Figure 3-14 Structure of a multiple representation BDIR defined in 19794-1.2FDIS

A Biometrics Data Interchange Record (BDIR) is made by a General Header, defined also within 19794-1.2FDIS, and is called Biometric Data Block (BDB) Header in 19794-7.2WD2. This is followed by the BDB Body, which is composed of at least 1 representation. Each representation is structured in a Representation Header and a Representation Body (also known as Representation Data as shown in Figure 3-14).

As previously mentioned, the main changes within this format and that defined within the first generation is the use of a completely new version of the header, which is now divided into a General Header and one or more Representation Headers.

In Figure 3-15 the fields included in the Record Header for a signature BDIR are shown. In the figures which show data structures, the solid boxes indicate fields that are present whereas the dashed outlined boxes indicate optional fields. Also the length of each field in bytes is indicated in parentheses at the bottom of the corresponding box.

The new Biometric Data Block (BDB) General Header identifies the modality of the BDB ("Format Identifier") as well as its version ("Version Number"). It also indicates the length of the BDB and the number of representations within the BDB Body. In this way, this new generation of 19794-7 Full Format enables a single record to contain multiple signature samples (representations), where this has not been possible in the first generation in 19794-7 [114].

The presence or absence of Certification Blocks at the representation level is indicated by the "Certification Flag" field.

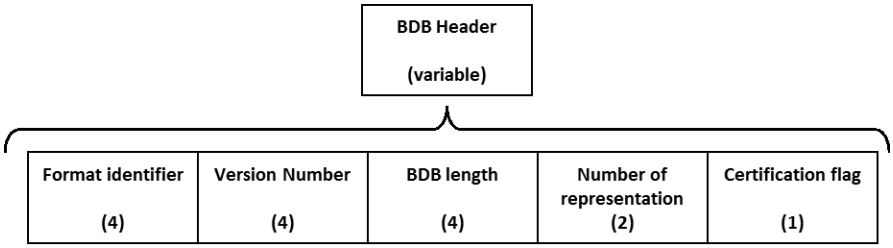


Figure 3-15 BDB General Header for 19794-7.2WD2 Full Format

Following the BDB General Header, the BDB body must contain at least one single signature sample (representation). Each representation consists of a "Representation Header" and a "Representation Body" as shown in Figure 3-16.

When compared to the first generation, the Representation Header (Figure 3-17) incorporates several new fields, these include: "Capture Data and Time" which indicates when the capture of the representation started. The "Capture Device Technology ID" indicates the class of capture device technology used to acquire the biometric sample. The "Capture Device Vendor ID" and "Capture Device Type ID" indicate the vendor and product type as well as the quality blocks which contain the predicted comparison performance of this representation.

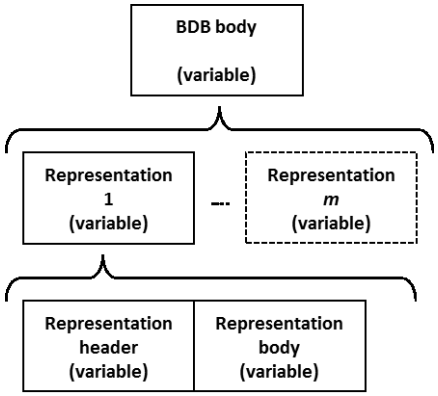


Figure 3-16 BDB Body of 19794-7.2WD2 Full Format

The "Preamble" field in the Representation Header indicates the presence or absence of optional extended data within the BDB Body.

Apart from these fields, the following are the same as described in the first generation: "Channel Description" and "Number of Sample Points".

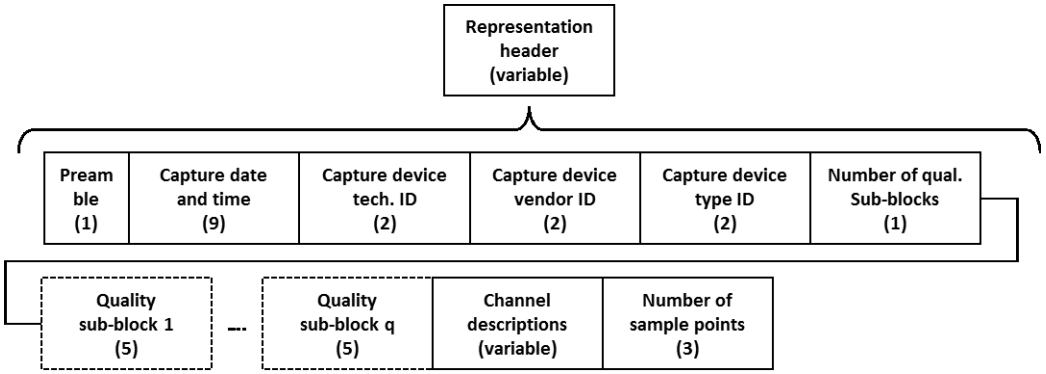


Figure 3-17 BDB Representation Header of 19794-7.2WD2 Full Format

The Channel Descriptions data (see Figure 3-18) begins with the "Channel Inclusion Field" indicating the presence or absence of particular channels. X and Y channels are mandatory as well as the T channel, DT channel, or uniform sampling, where the channel used must be indicated. Following the indication of the channels present, will be a single channel description field for each channel, this is shown as "Channel Inclusion Field" (see Figure 3-18). These channel descriptions contain information such as: scaling value, minimum and maximum possible channel values, mean value and standard deviation of the channel values, whether the channel value is constant or not and if the linear component of the regression line for this channel has been removed.

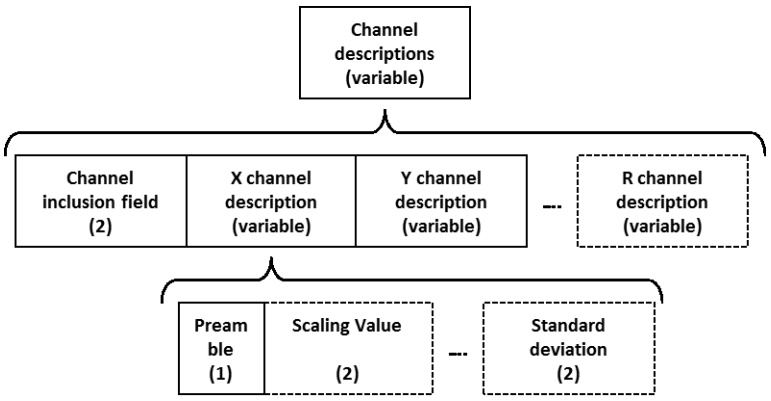


Figure 3-18 Channel descriptions of 19794-7.2WD2 Full Format

The number of sample points included in the BDB Representation Body is indicated in its corresponding field “Number of Sample Points”.

Within the second generation of 19794-7, in the BDB Representation Body a sequence of sample points can be found and also, if indicated at the preamble, the extended data (see Figure 3-19). The structure of the optional extended data is not defined within this format.

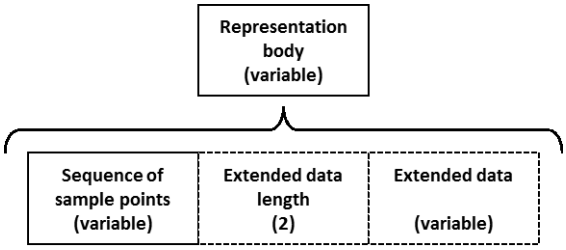


Figure 3-19 BDB Representation Body of 19794-7.2WD2 Full Format

The sequence of sample points (see Figure 3-20) remains the same as that defined in the first generation. Here each sample point contains the values of the channels indicated in the “Channel Inclusion Field”, and are stored in 2 bytes, with the exception of the S Channel, which is stored in 1 byte.

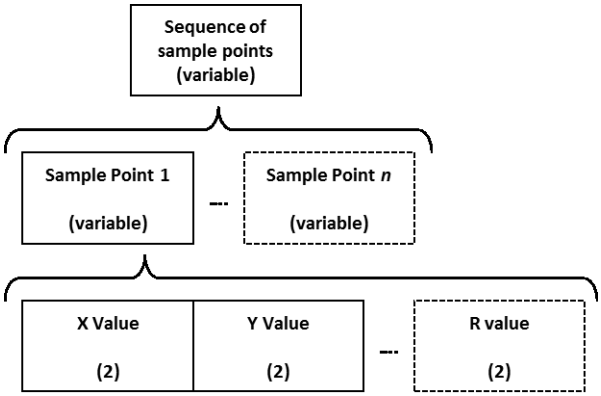


Figure 3-20 Sequence of Sample Points of 19794-7.2WD2 Full Format

3.6.1.2 COMPACT FORMAT

The Compact Format remains the same as that defined in the first generation. This Compact format is defined for use with smart cards and other tokens requiring a smaller representation size. The Compact Format Representation Body is made up of a sequence of sample points, where each sample point value is stored in just 1 byte (see Figure 3-21), also this particular Compact Format does not contain a header.

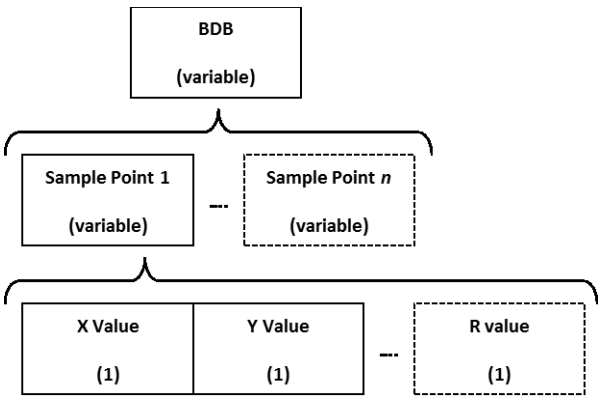


Figure 3-21 BDB of 19794-7.2WD2 Compact Format

As in the first generation, the Compact Format does not allow multiple signature representations. Information regarding the structure and contents of the data block (BDB) are contained within separate matching algorithm parameter data objects, defined in [118].

3.6.2 2ND COMMITTEE DRAFT 19794-11, BIOMETRIC DATA INTERCHANGE FORMATS – PART 11: SIGNATURE/SIGN PROCESSED DYNAMIC DATA

The Part 11 of the 19794 [119] project is currently under development, where its second Committee Draft (CD2) version was released in 2009 (this will be referenced as 19794-11CD2 hereinafter). The work carried out in this Thesis has been started, based on this second Committee Draft, in order to improve it. Due to the late start of this part, it has been directly aligned with the second generation of the 19794 series of standards.

This part is self-described as a compression of part 7, and is sufficiently compact to be stored in smart cards and other tokens. This compression is based on the segmentation of the signature into components of pen-strokes and pressure-strokes. Instead of storing all the sample values of every stroke, an account of both pen-strokes and pressure-strokes is stored. The information content of the Strokes is provided using several values (initial value, end value, min, max, mean, etc.) of the x and y axes, velocity, acceleration, pressure and time.

The 19794-11CD2 defines two different segmentations, one for x and y channels (pen-strokes), and another for the pressure channel (pressure-strokes).

A BDB conformance with 19794-11CD2 is made up of a BDB Header and a BDB Body.

The BDB Header, see Figure 3-22, stores the information received on the data interchange format including its version as well as the length, in bytes, of the BDB, the number of representations (number of views) stored within this BDB, and also information regarding the capture device.

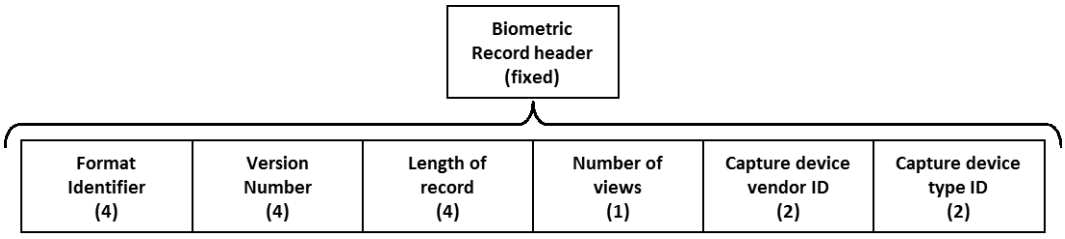


Figure 3-22 BDB Header of 19794-11CD2

The BDB Body, see Figure 3-23, stores information on the capture device (such as scaling values and sample resolution), it also allows the inclusion of Extended data, again without defining its structure.

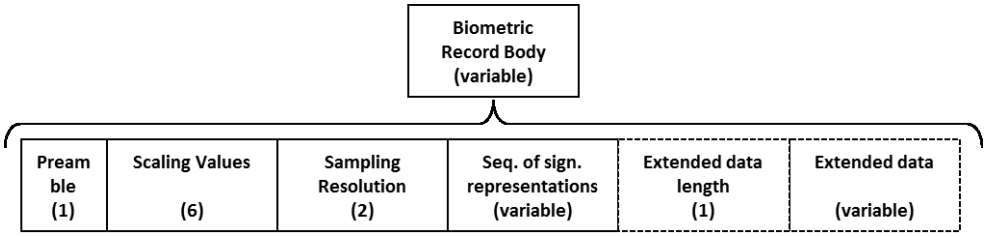


Figure 3-23 BDB Body of 19794-11CD2

The BDB Body also includes the representations of the signatures, see Figure 3-24. These representations are made up by a sequence of pen and pressure strokes together with several other overall features.

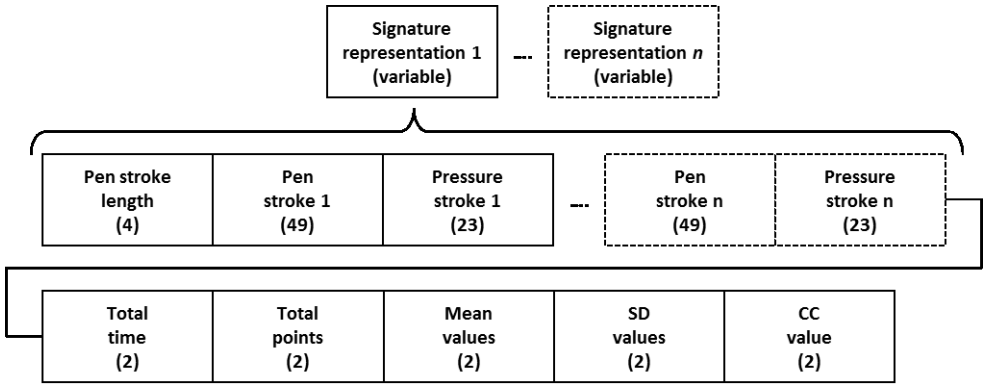


Figure 3-24 Signature Representation of 19794-11CD2

Pen-Strokes

The X and Y segmentation is based on pen-strokes. A pen-stroke is defined as the "movement of a pen between two singular points" [119]. These singular points can be a pen-down, a pen-up or a turning point. A turning point is defined as "a sample point where either x, y or both axes values change from increasing to decreasing" [119]. As a result, 4 types of strokes may be defined:

- Pen-down to turning point,
- Turning point to turning point,
- Turning point to pen-up,
- Pen-down to pen-up.

For each pen-stroke, see Figure 3-25, its starting point and end point (x-plane, y-plane and t values) will be stored, and also attributes such as the velocity (maximum, minimum and mean of v_x and v_y values), acceleration (maximum, minimum and mean of ax and ay values)

and pressure (maximum, minimum and mean of pressure values) during the pen-stroke. The length of the Pen-stroke and vector direction is also recorded.

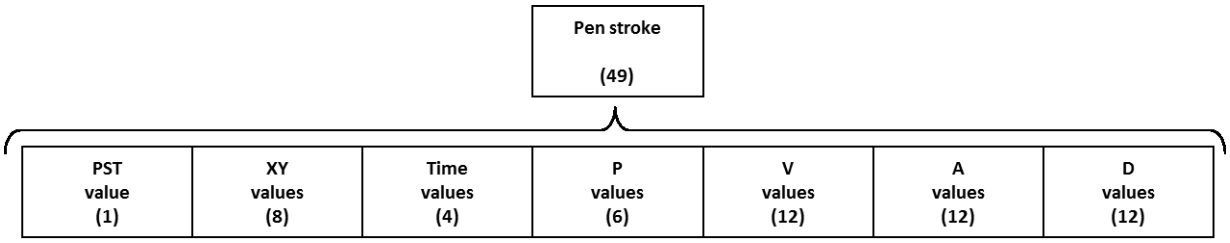


Figure 3-25 Pen-Stroke defined in 19794-11CD2

Pressure-Strokes

Analogous to the pen-strokes, the pressure segmentation is based on pressure-strokes, which are defined as the movement of a pen between two singular points. These singular points can be a pen-down, pen-up and also as a pressure turning point. A pressure turning point is defined as a sample point where the pressure value changes from increasing to decreasing or vice versa. Again, 4 types of pressure-strokes may be defined:

- Pen-down to turning point,
- Turning point to turning point,
- Turning point to pen-up,
- Pen-down to pen-up.

For each pressure-stroke, see Figure 3-26, the start and end of the pen movement (x-plane, y-plane and t values) will be stored, along with the pressure data (end, start, maximum, minimum and mean values).

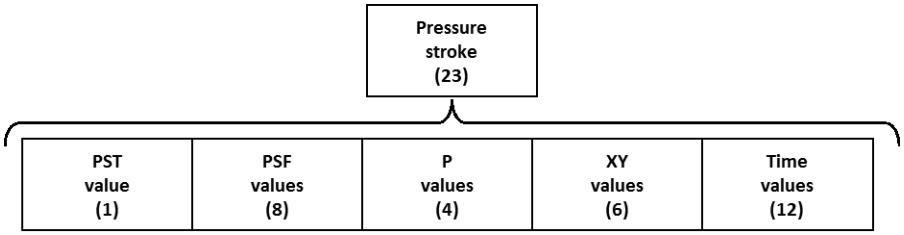


Figure 3-26 Pressure-Stroke defined in 19794-11CD2

Overall Data

Together with the pen-stroke and pressure-stroke data, additional data is recorded that represents the overall features of the signature representation, see Figure 3-27. These data are:

- Total time
- Total number of sample points
- Mean values of x, y and p channels.
- Standard deviation of x, y and p channels.
- Correlation coefficient between x and y channel.

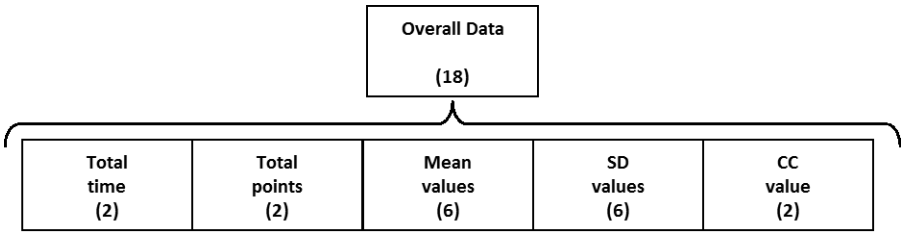


Figure 3-27 Overall Data defined in 19794-11CD2

3.7 CONCLUSIONS

In this chapter an introduction to automatic signature verification has been presented. The differences between online and offline signature verification systems were defined. The work presented in this Thesis focuses on online systems. It has been indicated that such behavioural biometric modality systems are susceptible to fraud as they allow the possibility of impersonation attempts by users who not only imitate the shape of the signature but also the way of performing the signature. Considering this, different levels of signature forgeries have been described.

The main techniques used for comparison algorithms have been introduced along with the databases used for their evaluation. Following this, details were provided on the two public signature evaluation campaigns, where the results obtained have been explained in detail.

Once the automatic signature verification systems were introduced, the necessity of signature data format standards arose as a result of the wide range of capture devices available on the market. The two ISO/IEC standards which deal with signature (part 7 and part 11 of ISO/IEC 19794 project) have been explained, where details on the data format structure were presented.

In the following chapters, the research done for this Thesis on signature verification algorithms and the development of the two ISO/IEC signature standards will be presented.

Chapter 4 IMPROVEMENT IN AUTOMATIC SIGNATURE VERIFICATION

4.1 INTRODUCTION

The wide range of applications suited to Automatic Signature Verification Systems (ASVS) has been discussed in the previous chapter. This is particularly true for banking transactions, paperless processes and, authorisation documents etc. In automatic biometric verification systems the key issues to be addressed are generally the error rate levels, specifically in signature verification for both random and skilled forgeries, and implementation feasibilities, i.e. computational load and storage needs. Implementation feasibilities are significant for systems based on biometric tokens, where the aforementioned aspects are their main constraints.

This chapter will deal with both error rates and implementation feasibility with respect to the storage needs and computational load. As a starting point, two well-known signature verification algorithms will be analysed, these are the Gaussian Mixture Models (GMM) and the Dynamic Time Warping (DTW). Reduction of the computational load and storage requirements are confronted using feature selection.

It would seem natural that improving the systems performance is achieved by increasing the number of features. However, this is not always true and may in fact create several drawbacks. This is especially the case for computation load, storage space required and data exchange.

Feature selection methods reduce the size of the user model while maintaining, or even improving, the error rate levels. Also, reducing the number of features leads to reduced computational requirements.

The following section in this chapter will introduce briefly the GMM and DTW algorithms, providing details on the features analysed. In section 4.3, different feature selection techniques (Fisher Ratio (FR), Principal Component Analysis (PCA) and the Hellinger Distance (HD)) are explained. Following this, the results from applying these techniques to both algorithms will be presented, indicating the selected reduced feature vectors and their performance in terms of verification error rates (section 4.4 for GMM and section 4.5 for DTW). This chapter concludes with a review of the influence of different parameters on both algorithms (GMM and DTW), and will show how the verification error rates are affected by their values.

4.2 SIGNATURE VERIFICATION ALGORITHMS

Among all the different techniques proposed in relevant references for signature verification systems [38-41], the Gaussian Mixture Models (GMM) and Dynamic Time Warping (DTW) have been selected for this work as good candidates for ASVS algorithms for several reasons which are presented now.

The GMM has been successfully used in other biometric modalities [110] [120], but its use in ASVS is not very extensive, where there are only a small number of published works with results [80, 83-85]. The GMM will be explained briefly in section 4.2.1, followed by the set of features which will be analysed in section 4.2.1.3.

The DTW is currently regarded as the most successful technique for SAVS. This was found to be the most suited algorithm for this application at two signature verification competitions: SVC'2004 [22] and BSEC'2009 [70]. Section 4.2.2 will briefly explain the DTW technique. In section 4.2.2.3 the derived temporal signals used to calculate the features will be introduced.

4.2.1 GAUSSIANS MIXTURE MODELS

The Gaussians Mixture Models (GMM) are well known and a highly referenced technique for pattern recognition. The appearance of the Expectation-Maximization algorithm [121] resulting from its training has demonstrated that this technique is an appropriate alternative for pattern recognition tasks. The GMM has been widely and successfully used in an extensive range of applications. Its first and most well-known application in the biometrics field was demonstrated by Reynolds [110], where it was used for voice recognition. Another biometric modality that has successfully used this technique is hand geometry [120].

Although the GMM has been applied to On-line Signature Verification [80, 83-85], the literature is not as extensive as with other techniques, i.e. the HMM and the DTW. Table 4-1 shows the most relevant published results from GMM-based systems, detailing the year of publication, the database used on the evaluation and the equal error rates (EER) obtained for skilled forgeries (SF) and random forgeries (RF) if available (N/R, Not Reported, indicates that the error rates were not given).

Table 4-1 On-Line Signature Verification Systems based on GMM

Authors	Year	Database	EER (SF)	EER (RF)
J. Richiardi A. Drygajlo *[83]*	2003	MCyT	3.4%	N/R
W. Liang *et al. [84]*	2005	Not Public	6.7%	N/R
Bao Ly, Van Garcia-Salicetti, S. Dorizzi, B. *[80]*	2007	MCyT	6.7%	N/R
A. Ahrary *et al. [85]*	2009	SVC2004	14%	11%

In general terms, the underlining concept of the GMM is to represent the user's biometric characteristics, not as a set of features but as a weighted sum of probability Gaussian functions. One of the powerful attributes of the GMM is its ability to obtain smooth approximations for arbitrarily-shaped densities by means of a linear combination of probabilistic functions.

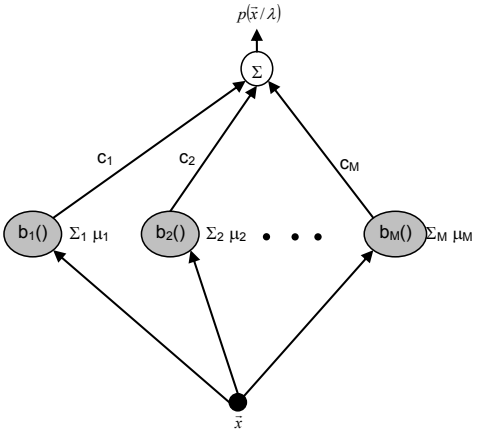


Figure 4-1 Gaussians Mixture Model Representation

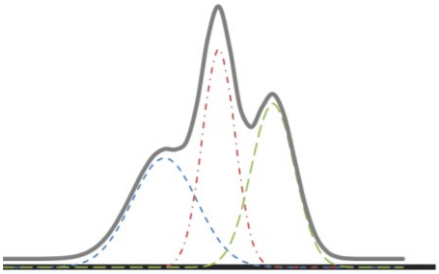


Figure 4-2 Mixture of Gaussian Probabilistic functions

The probability density function for a feature vector "$\vec{x}$" and a user model "λ", is defined as a weighted sum of the M probabilistic Gaussian functions, as follows:

$$p(\vec{x}/\lambda) = \sum_{i=1}^{M} c_i \cdot b_i(\vec{x}) \qquad (1)$$

where "c_i" is the mixture weights and "M" is the number of mixed Gaussian probability functions (3). The term "c_i" must satisfy the constraint (2) in order to obtain a legitimate probability density function.

$$\sum_{i=1}^{M} c_i = 1 \tag{2}$$

Each Gaussian probabilistic function is defined as:

$$b_i(\vec{x}) = \frac{exp\left(-\frac{1}{2}(\vec{x}-\overrightarrow{\mu_i})^T \Sigma_i^{-1}(\vec{x}-\overrightarrow{\mu_i})\right)}{(2\pi)^{L/2}\sqrt{\|\Sigma_i\|}} \tag{3}$$

Where "$\overrightarrow{\mu_i}$" and "Σ_i" are the mean vectors and covariance matrixes respectively for each Gaussian function i= 1 … M. The term "$\vec{x}$" is the input feature vector and "L" is the number of features within the feature vector.

Therefore, the complete Gaussian Mixture Model (4) for a user, "λ", is represented by a set of three different parameters:

- mean vectors "μ_i",
- covariance matrixes "Σ_i",
- weight factors "c_i", all of them for i=1...M.

Thus, it may be said that each user will have his/her own GMM model, "λ_s", which can be represented by the following notation:

$$\lambda_s = \{\overrightarrow{\mu_i}, \Sigma_i, c_i\}_{s=1\ldots S}^{i=1\ldots M} \tag{4}$$

Here "S" is the number of users, and "M" the number of probability Gaussian functions.

In these general GMM models several simplifications, based on previous works, can be performed [110] [122]. One of the most important simplifications relates to the covariance matrix of the Gaussian function. This matrix can be considered full or diagonal. Other theoretical studies have confirmed that the results obtained using diagonal matrices are equal and in some cases better than those containing a full matrix. This is also true for cases where the features are not statistically independent [110] [122] [83]. Moreover, different types of covariance matrices can be considered: a) one for each node, designated nodal covariance; b) one for each user, designated total covariance; c) one for the complete system, designated global covariance. Due to the improved results obtained, observed from a literature review, [110] [122] [83] [123], the nodal covariance matrix has been employed.

In the following sections of this chapter the methods used to obtain the user model and the techniques to verify a signature sample will be explained.

4.2.1.1 EXPECTATION-MAXIMIZATION ALGORITHM, TRAINING GMM

The Expectation-Maximization (EM) algorithm was first presented, and named, in a well-known publication that dates back to 1977 [121]. The EM algorithm has been used in the training process which, from a set of genuine signature samples, produces user models. The EM algorithm detailed in [120] has been applied to obtain the user models.

Given the data training, $x = \{\vec{x}_1, \vec{x}_2, \dots, \vec{x}_T\}$ - "T" feature vectors from genuine signature samples are obtained. The EM algorithm attempts to determine the user model λ which maximizes the likelihood of the GMM, this can be expressed as:

$$p = \prod_{t=1}^{T} p(\overrightarrow{x_t}/\lambda) \tag{5}$$

The EM algorithm estimates the model iteratively. From an initial model λ, the EM produces a new model $\hat{\lambda}$, which has a greater likelihood than the previous one. The new model becomes an initial model for a following iterative training step. This iterative process is repeated until a specified threshold is achieved.

The EM training process is based on two different steps referred to as *expectation* and *maximization*:

i) **Expectation Step**: computes *a posteriori* the expectation of the log-likelihood evaluated using the current model, for each sample signature from the training data. The *a posteriori* probability is given by:

$$p(i/\vec{x}, \lambda) = \frac{c_i \cdot b_i(\overrightarrow{x_t})}{\sum_{k=1}^{M} c_k \cdot b_k(\overrightarrow{x_t})}, \; 1 \le i \le M, \; 1 \le t \le T \tag{6}$$

ii) **Maximization Step**: computes new parameters, from the initial model, that maximize the expected log-likelihood found during the expectation step. The following estimation equations are used to calculate the new weight coefficients "$\widehat{c_i}$", mean vectors "$\widehat{\overrightarrow{\mu_i}}$" and covariance matrixes "$\widehat{\sigma^2}_i$" respectively:

$$\widehat{c_i} = \frac{1}{T}\sum_{t=1}^{T} p(i/\overrightarrow{x_t}, \lambda) \tag{7}$$

$$\widehat{\overrightarrow{\mu_i}} = \frac{\sum_{t=1}^{T} p(i/\overrightarrow{x_t}, \lambda) \cdot \overrightarrow{x_t}}{\sum_{t=1}^{T} p(i/\overrightarrow{x_t}, \lambda)} \tag{8}$$

$$\widehat{\sigma^2}_i = \frac{\sum_{t=1}^{T} p(i/\overrightarrow{x_t}, \lambda) \cdot \left(\overrightarrow{x_t} - \widehat{\overrightarrow{\mu_i}}\right) \cdot \left(\overrightarrow{x_t} - \widehat{\overrightarrow{\mu_i}}\right)'}{\sum_{t=1}^{T} p(i/\overrightarrow{x_t}, \lambda)} \tag{9}$$

The EM steps are repeated until a convergence threshold is reached. The threshold used was the minimum difference between two consecutive mean vectors [120]. When that difference is below the threshold, the EM algorithm is assumed to have converged.

Before the EM algorithm is applied, an initial user model, λ_1, has to be defined. There are no adequate theoretical methods or guidelines on the initialization of the users model, thus the initial model was defined from a literature review as [83, 110] [122] [123]:

- **Number of Gaussians "M"**: The M value is set after heuristic tests. Large M values generally lead to improved performance, however they incur greater computational loads.
- **Weight Coefficients"c_i":** Weight coefficients must only satisfy (2), these are typically initialized as 1/M.
- **Covariance matrixes Σ_i:** The covariance matrixes are initialized as identity matrixes. It is also worth highlighting that a minimum value is predefined to avoid a "0" and small values which imply $\|\Sigma_i\|$ is equal to "0" in (3), although depending on the number of features. A high number of features imply that the product of all the elements from the diagonal is below the computer resolution threshold, which entails a greater set minimum value.
- **Mean Vectors "$\overrightarrow{\mu_i}$":** There are different methods for initializing the mean vectors, these are: K-means clustering [83], random mean selection vector from the input data [120], and a combination of both [110]. Previous work has found no significant difference between these different initialization techniques [110]. To reduce the computational load, random mean selection has been chosen.

4.2.1.2 Signature Scores Computation

Once the user model has been obtained by the EM algorithm, $\lambda_s = \{\overrightarrow{\mu_i}, \Sigma_i, c_i\}^{i=1\ldots M}$, a signature score is calculated using the equations (1) and (3).

4.2.1.3 Extracted Features

More than 140 features, based on those selected from a literature review have been analysed [40, 41 , 105, 124-126]. These features are calculated from the 5 original signals acquired by the input device (i.e. x and y position, pressure, azimuth and elevation) and their derived signal velocity and acceleration.

4.2.1.3.1 Preprocessing

The raw signals captured in the MCyT need to be preprocessed to remove any noise and irrelevant information. The following preprocessing steps are used:

i) ***Smoothing*** of the five temporal signals (x and y axis pressure, azimuth and inclination) using a low pass filter to eliminate the noise introduced during the capture process. This filter is a 5 point moving average.

ii) ***Location normalization***: both the x-axis and y-axis temporal functions are normalized by centring the signature at its mass centre:

$$\hat{x}(t) = x(t) - x_m \quad (10)$$

$$\hat{y}(t) = y(t) - y_m \quad (11)$$

iii) ***Size normalizing***: both the x-axis and y-axis are normalized using the norm of their 2 dimension vector [x,y]:

$$\hat{x}(t) = {x(t)}/{norm([x,y])} \quad (12)$$

$$\hat{y}(t) = {y(t)}/{norm([x,y])} \quad (13)$$

where the norm used is defined as:

$$norm([x,y]) = \sqrt{\int_0^1 \|[x(t),y(t)]\|^2 dt} \quad (14)$$

The Pressure, azimuth and inclination are normalized using their maximum values. These values are given by the capture device and are detailed in [21]). The following equations apply:

$$\hat{p}(t) = {p(t)}/{1023} \quad (15)$$

$$\widehat{az}(t) = {az(t)}/{360} \quad (16)$$

$$\widehat{in}(t) = {in(t)}/{90} \quad (17)$$

4.2.1.3.2 Derived Signals: Velocity and Acceleration

The velocity of the coordinates x-axis, y-axis, pressure, azimuth and inclination are calculated as the first derivative of their signals:

$$v_s(t) = \frac{s(t+1) - s(t)}{(t+1) - t} \quad (18)$$

The acceleration of these coordinates is calculated taking the first derivative of their velocities:

$$a_s(t) = \frac{v(t+1) - v(t)}{(t+1) - t} \quad (19)$$

Both temporal functions (velocity and acceleration for the x and y coordinates) are also normalized by using their norm. Where the definition of the norm is the same as (14), but replacing the 2 dimension vector [x,y] with a one dimension vector (velocity or acceleration).

4.2.1.3.3 FEATURE EXTRACTION

A set of 143 features is obtained from the temporal signals acquired in the MCyT database (x and y position "*x*" "*y*", pen pressure "*p*", pen azimuth "*az*" and pen inclination "*in*"). Also, as previously mentioned, the velocity and acceleration from these signals have also been considered (v_x, v_y, v_p, v_{az}, v_{in}, a_x, a_y, a_p, a_{az}, a_{in}).

More precisely, for the signals captured directly using the tablet input device, the following features are extracted, detailed in Table 4-2, where "*s*" (signal) means any of them (i.e. *x*, *y*, *p*, *az* or *in*), and "s" is a vector of n elements, $s = \{s_1, s_2, \dots, s_n\}$:

Table 4-2 Features analysed from the signals captured directly using the tablet input device, GMM algorithm

	Name	Description/Mathematical Expression
1	$s_{\max\_end}$	$\max(s) - s_1$
2	s_{ini_min}	$s_1 - min(s)$
3	$s_{\text{end\_min}}$	$s_n - min(s)$
4	s_{mean}	$\frac{1}{n}\sum_{i=1}^{n} s_i$
5	s_{rms}	$\sqrt{\frac{1}{n}\sum_{i=1}^{n} {s_i}^2}$
6	$s_{\text{max\_mean}}$	$\max(s) - \frac{1}{n}\sum_{i=1}^{n} s_i$
7	$s_{\text{mean\_min}}$	$\frac{1}{n}\sum_{i=1}^{n} s_i - s_n$
8	$s_{\text{max\_rms}}$	$\max(s) - \sqrt{\frac{1}{n}\sum_{i=1}^{n} {s_i}^2}$
9	s_{rms_min}	$\sqrt{\frac{1}{n}\sum_{i=1}^{n} {s_i}^2} - min(s)$
10	s_{std}	$\sqrt{\frac{1}{n}\sum_{i=1}^{n} (s_i - \mu)^2}$
11	$s_{zero\_crossing}$	Number of points where the sign of a function changes (e.g. from positive to negative), represented by a crossing of the axis (zero value) on the graph of the function

As there are five signals captured directly using the tablet (*x*, *y*, *p*, *az* and *in*), this means that there are 55 features in a single signature.

From the velocity "v" and accelerations "a" of the signals captured directly by the tablet (v_x, v_y, v_p, v_{az}, v_{in}, a_x, a_y, a_p, a_{az}, a_{in}), the features detailed in Table 4-3 are calculated, "s" again means any of them, being a vector of n elements, $s = \{s_1, s_2, \ldots, s_n\}$:

Table 4-3 Features analysed for the velocity and accelerations, GMM algorithm

	Name	Description/Mathematical Expression
1	$s_{\text{time\_over0}}$	% of time the signals is positive
2	$s_{zero\_crossing}$	Number of times where the sign of a function changes (e.g. from positive to negative), represented by a crossing of the axis (zero value) in the graph of the function
3	s_{rms}	$\sqrt{\frac{1}{n}\sum_{i=1}^{n} {s_i}^2}$
4	$s_{mean\_\max\_ratio}$	$\frac{1}{n}\sum_{i=1}^{n} s_i \Big/ \max(s)$
5	$s_{\text{max\_mean}}$	$\max(s) - \frac{1}{n}\sum_{i=1}^{n} s_i$
6	$s_{\text{max\_min}}$	$\max(s) - \min(s)$
7	$s_{\text{mean\_over0}}$	$\frac{1}{m}\sum_{i}^{m} s_i \in s > 0$
8	$s_{\text{mean\_below0}}$	$\frac{1}{m}\sum_{i}^{m} s_i \in s < 0$

Where "m" is the number of elements which satisfy the condition, i.e. s<0 or s>0.

These eight features per derived signal imply 80 features extracted from 5 velocities and 5 accelerations.

Moreover, the following global features related to the completed signature sample have also been analysed:

Table 4-4 Global features analysed, GMM algorithm

	Name	Description/Mathematical Expression
1	N_Strokes	Number of pen down events
2	Time	Total time of a signature
3	T_writting	Ratio between the writing time (where the tip of the pen is touching the surface of the input device) and the total time of a signature
4	Width_Height	$(\max(y) - \min(y))/(\max(x) - \min(x))$
5	Heigth_Width	$(\max(x) - \min(x))/(\max(y) - \min(y))$
6	Area	$(\max(x) - \min(x)) \cdot (\max(y) - \min(y))$
7	Length	$\sum_{i=1}^{n-1} \sqrt{(x_i - x_{i+1})^2 + (y_i - y_{i+1})^2}$
8	Length_Area	$\frac{\sum_{i=1}^{n-1} \sqrt{(x_i - x_{i+1})^2 + (y_i - y_{i+1})^2}}{(\max(x) - \min(x)) \cdot (\max(y) - \min(y))}$

Summarizing the previous paragraphs and tables of this section, 55 features are calculated from the five temporal signals captured by the input devices. Another 80 features are calculated from their velocities and accelerations and, finally, 8 global features are added. In total, 143 features will be analysed using the feature selection technique, where different approaches have been used to identify the reduced feature subsets. These subsets can offer smaller sized feature vectors and user models and also a reduced computational load.

4.2.2 DYNAMIC TIME WARPING

The Dynamic Time Warping (DTW) algorithm originates from the field of speech recognition [127]. One of the first successfully attempts of the DTW for handwritten signature verification came from Sato and Kogure [49]. It was also successfully used in the two signatures competitions: SVC'2004 [22] and BSEC'2009 [70].

In Table 4-5 the most relevant published results achieved using DTW-based systems, specifying the equal error rates (EER) achieved for both skilled forgeries (SF) and random forgeries (RF) when available (N/R, Not Reported, indicates that the error rates were not given):

Table 4-5 On-Line Signature Verification Systems based on DTW

Authors	Year	Dataset	EER (SF)	EER (RF)
A. Kholmatov K. Yanikoglu *{Kholmatov, 2005 #159}*	2003	No Public	1.41%	N/R
Alisher Kholmatov Berrin Yanikoglu *[128]*	2004	SVC2004	2.84%	2.79%
G.K. Gupta R.C. Joy *[129]*	2007	No Public	4.8%	2.25%
L. Nanni A. Lumini *[130]*	2006	MCyT	7.6%	2.3%
Jain, A.l K. Griess, F. D. Connell, S. D. *[60]*	2002	No Public	3.3 %	3.2 %
Pascual-Gaspar, J. Cardeñoso-Payo, V. Vivaracho-Pascual, C. *[58]*	2009	MCyT	1.06%4	0.38/%[4]
BSEC 2009 *[70]*	2009	MCyT	3.91%	1.20%
Pascual-Gaspar, J. Cardeñoso-Payo, V. Vivaracho-Pascual, C. *[58]*	2009	BioSecure	2.20%	0.97%

The DTW deals with the non-linear temporal axis alignment between two signatures, one of them referred to as the reference (black line in Figure 4-3) and the other as sample (grey line in Figure 4-3).

[4] User dependent threshold

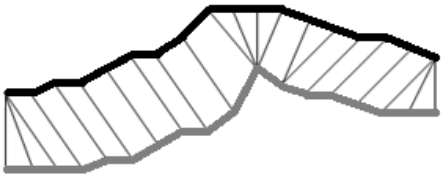


Figure 4-3 Time Alignment between two sequences

In Figure 4-3, the time axis is warped so that each data point in the grey sequence is optimally aligned to a point in the black sequence.

In this work, the signatures are defined as temporal signals formed by a 2 dimensional vector [49], created by joining the x and y position:

$$\begin{matrix} r_j(t_j) = [x_r(t_j), y_r(t_j)] & j = 1 \dots T_r \\ s_i(t_i) = [x_s(t_i), y_s(t_i)] & i = 1 \dots T_s \end{matrix} \quad (20)$$

In order to perform the DTW, a measure of the distance (Euclidean distance in a 2D space) between two points is defined:

$$d(i,j) = \|s_i - r_j\| \quad (21)$$

The DTW algorithm, using Dynamic Programming, iteratively fills a distance matrix between the pattern and sample points. From the different options proposed in the work of Sakoe and Chiva [127], to fill the distance matrix, a symmetric DP-Algorithm has been used which has a slope constraint condition equal to 1, this was implemented as:

$$g(i,j) = min \begin{cases} g(i-1,j-2) + 2 \cdot d(i,j-1) + d(i,j) \\ g(i-1,j-1) + 2 \cdot d(i,j) \\ g(i-2,j-1) + 2 \cdot d(i-1,j) + d(i,j) \end{cases} \quad (22)$$

After the distance matrix is filled, the optimum way of aligning the sample and the reference is found using backtracking to find the best (minimal distance) time alignment. This is known as the warping path (black dots in Figure 4-4).

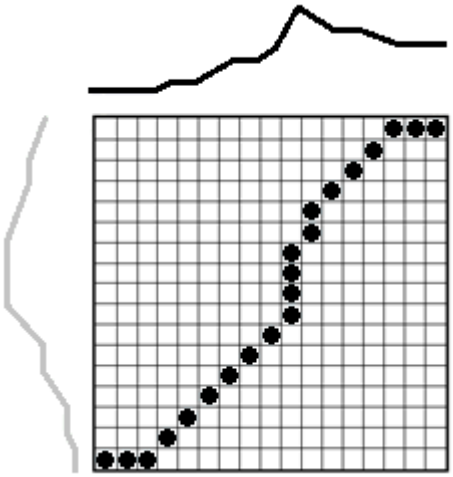


Figure 4-4 Warping Path aligning the pattern and the sample

When there is no time difference between the two samples, this warping function coincides with the diagonal line ($r_i = s_j$).

Several restrictions on the matrix distances can be made to minimize the number of calculated points. These may include, amongst others, adjustment window requirements, slope constraints, etc. [127].

4.2.2.1 DTW USER MODEL

The DTW is used to construct the user model. This model serves as a reference for alignment of signature samples used for verification before extracting the features.

This user model is built by taking one signature from the training data set (normally between 3 and 10 samples), and aligning the rest of the training signatures with it, using X and Y position signals. Once all signatures have been aligned, the average X and Y signals are calculated.

These X and Y averaged signals are inversely aligned temporally to create the DWT user model. To perform this inverse alignment, a distance matrix calculated as the average of the distance matrices from the training data set alignments is used. This technique was proposed by Sato and Kogure [49].

Hence, the user model is composed of at least the X and Y position signals and the time evolution (if the sampling rate is fixed, the time evolution may be omitted). The inclusion of pressure and inclination data depends on the features that will be used by the algorithms. The number of points for each signal is defined during the preprocessing stage (see 4.2.2.3.1).

4.2.2.2 ALIGNING SIGNATURE SAMPLES TO THE USER MODEL

The first step is to align the signature samples used for verification with the DTW user model belonging to the identity claimed. This alignment will identify the warping path which minimizes the distance between the sample and the user model by only using the information included in the X and Y position signals.

Once this warping path is found, it is used to align the 5 signals captured by the input device (X and Y position, pressure, azimuth and elevation). Following the development of this alignment, the feature extraction process takes place, as will be explained in the next section.

4.2.2.3 EXTRACTED FEATURES

4.2.2.3.1 PREPROCESSING

As in the GMM algorithm, the signals captured by the input device are preprocessed to reduce noise and irrelevant information. The same preprocessing steps are used (**¡Error! No se encuentra el origen de la referencia.**) along with a further step, based on linear Interpolation, to reduce the number of sample points. This transforms the original temporal signals to equi-spaced 256-point temporal signals (the influence of the number of points used in this step is analyzed in section 4.5.3).

After these steps, the 5 captured signals are time aligned with the user model from the identity claimed (see 4.2.2.2).

4.2.2.3.2 FEATURE EXTRACTION

Once the signals acquired by the digital tablet have been preprocessed and aligned, the feature extraction process begins by calculating 25 derived signals, as detailed in [80]. These derived signals have been successfully used by B. Ly Van, Sonia Garcia-Salicetti and B. Dorizzi whom have used the Viterbi Path and HMM's likelihood technique [80].

To calculate the derivative signal, a regression formula with 'O' order [131] is used:

$$reg\ (z(t), O) = \frac{\sum_{k=1}^{O} k \cdot \big(z(t+k) - z(t-k)\big)}{2 \cdot \sum_{k=1}^{O} k^2} \qquad (23)$$

As in [80], an order value of 2 has been used to obtain an approximation of the derivative signal, providing softened waveforms, removing slight noise variations.

The 25 signals used are presented in Table 4-6, where "reg" represents the regression formula as expressed in (23):

Table 4-6 Derived signals used for DTW algorithm

	Description	Name	Mathematical Expression
1	Velocity in x	$v_x(t)$	$reg(x(t),2)$
2	Velocity in y	$v_y(t)$	$reg(y(t),2)$
3	Absolute velocity	v(t)	$\sqrt{v_x^2(t)+v_y^2(t)}$
4	Acceleration in x	$a_x(t)$	$reg(v_x(t),2)$
5	Acceleration in y	$a_y(t)$	$reg\big(v_y(t),2\big)$
6	Absolute acceleration	a(t)	$\sqrt{a_x^2(t)+a_y^2(t)}$
7	Tangential acceleration	$a_t(t)$	$reg(\|\|v(t)\|\|,2)$
8	Angle α	$\alpha(t)$	$asin\left(\frac{v_y(t)}{\|\|v(t)\|\|}\right)$
9	Cosine of angle α	$cos_\alpha(t)$	$\frac{v_x(t)}{\|\|v(t)\|\|}$
10	Sin of angle α	$sin_\alpha(t)$	$\frac{v_y(t)}{\|\|v(t)\|\|}$
11	Angle Φ	$\phi(t)$	$reg(\alpha(t),2)$
12	Cosine of angle Φ	$cos_\phi(t)$	$cos\big(\phi(t)\big)$
13	Sin of angle Φ	$sin_\phi(t)$	$sin\big(\phi(t)\big)$
14	Pressure	p(t)	Captured by the input device
15	Velocity of p	$v_p(t)$	$reg(p(t),2)$
16	Azimuth angle	az(t)	Captured by the input device
17	Inclination angle	$in(t)$	Captured by the input device
18	Velocity of azimuth angle	$v_{az}(t)$	$reg(az(t),2)$
19	Velocity of inclination angle	$v_{in}(t)$	$v_{in}(t)=reg(in(t),2)$

	Description	Name	Mathematical Expression
20	Curvature radius	$r(t)$	$\frac{a_t(t)}{reg(\phi(t),2)}$
21	The length to width ratio for windows of size 5, centred on the current point	$lw_5(t)$	$\frac{\sum_{i=t-2}^{t+1}\sqrt{(x_i-x_{i+1})^2+(y_i-y_{i+1})^2}}{\lvert max\{y_i \dots y_{t+2}\} - min\{y_i \dots y_{t+2}\}\rvert}$
22	The length to width ratio for windows of size 7, centred on the current point	$lw_7(t)$	$\frac{\sum_{i=t-3}^{t+2}\sqrt{(x_i-x_{i+1})^2+(y_i-y_{i+1})^2}}{\lvert max\{y_i \dots y_{t+3}\} - min\{y_i \dots y_{t+3}\}\rvert}$
23	The ratio of the minimum over the maximum velocity for a window of 5 points, centred on the current point	$\mathrm{minmaxv}_5(t)$	$\left\lvert\frac{min\{v_{t-2} \dots v_{t+2}\}}{max\{v_{t-2} \dots v_{t+s}\}}\right\rvert$
24	Coordinate x	x(t)	Captured by the input device
25	Coordinate y	y(t)	Captured by the input device

All these signals, "s", are normalized using their maximum and minimum in order to obtain comparable differences between them:

$$\hat{s} = \frac{s - min\{s\}}{max\{s\} - min\{s\}} \quad (24)$$

After all 25 signals have been generated and normalized, their distances from the reference signals obtained in the enrolment process are calculated. To calculate this distance, the definition provided by Kogure and Sato in [49] called pseudo-distance (pd) is used. In this Thesis, the distance was calculated for all 25 signals, as the mean absolute differences between the acquired and derived signals from the input signature (s_{sample}) and the signals and derived signals from the user pattern (s_{Model}) as follows:

$$pd_s = \frac{1}{N} \cdot \sum_{i=1}^{N} \left| s_{sample}(i) - s_{Model}(i) \right| \quad (25)$$

Here pd_s is the pseudo-distance for 's' signals and 'N' is the number of points, fixed by the equi-spaced preprocess.

Additionally, the features used by Kogure and Sato, described in [49], have been used:

Table 4-7 Kogure and Sato features used for the DTW algorithm

	Description	Name	Mathematical Expression
26	Complex Trace Coordinate	z	$x + y \cdot i$
27	Pseudo Distance	pd	As described in [49]
28	Pseudo Shape	ps	As described in [49]
29	Pseudo Pressure	pp	As described in [49]

The complex trace coordinate pseudo distance is calculated in the same way as mentioned before, normalizing the temporal signal using (24) and calculating the distance using (25).

4.2.2.4 Comparison Distance

To avoid increasing the computational load, a straightforward technique based on the Euclidean distance was used. The pseudo-distance vectors are normalized using the mean pseudo-distances vector obtained from the training set, according to the following relation:

$$\overrightarrow{pd}' = \overrightarrow{pd} - \overrightarrow{\overrightarrow{pd}} \tag{26}$$

Here $\overrightarrow{pd}$ is the pseudo-distances vector obtained from the sample signatures, and $\overrightarrow{\overrightarrow{pd}}$ is the mean pseudo-distance vector obtained from the training signature samples set.

These elements in the new pseudo-distance vector, $\overrightarrow{pd}'$, with negative values , are set to 0, so that only the elements from the pseudo-distance vector that are bigger than their respective mean values are considered.

$$\overrightarrow{pd}'' = \begin{cases} \overrightarrow{pd}', & pd' \geq 0 \\ 0, & pd0 < 0 \end{cases} \tag{27}$$

The Euclidean distance for this new pseudo-distances vector is:

$$d = \sqrt{\sum_{f=1}^{F} \left(pd''_f\right)^2} \tag{28}$$

4.3 FEATURES SELECTION TECHNIQUES FOR DIMENSIONALITY REDUCTION

4.3.1 INTRODUCTION

There has been much study carried out on the discriminative power of features and an in-depth analysis of the most suitable feature set for handwritten signatures [105, 125-126, 132]. For example, in [132] and [126] an initial set of 46 global and 39 local features were analyzed using a modified Fisher Ratio, where signatures from the MCyT database were tested using the GMM algorithm. Their results show equal error rates of 4.5% for skilled forgeries, using two subsets of 12 global and 12 local features. In [125] 49 features where analysed using a private database, obtaining 2.5% equal error rate with only 15 individualized parameter feature subsets selected from the 49-feature set. The feature selection technique used was based on the distance between users of the features mean values normalized to their standard deviation (using only genuine samples).

In this Section of Chapter 4 different feature selection techniques will be analysed. These techniques are: Fisher Ratio (FR), Principal Component Analysis (PCA) and the Hellinger Distance (HD), and will be applied to two different algorithms: DTW and GMM. The main aim is to reduce the number of features used by the algorithms to obtain a reduced user model size and to lower the computational load. Also it will be shown here how different feature selection techniques perform for the different algorithms considered.

The Fisher Ratio has been successfully used in many different studies [132] [126] [120]. The PCA is proposed in this Thesis as a feature selection technique in the signature verification field. In other works, the PCA has been used for verification algorithms [133], in this work it will be used as a feature selection technique as described in [134]. The Hellinger Distance (HD) [135] is proposed as a novel feature selection technique. The HD is used to quantify the similarity between feature probability distributions, comparing genuine and forgery feature distributions amongst users and selecting those features which maximize the difference between genuine and skilled forgery signatures.

The MCyT Signature Subcorpus Database [21] was used to carry out these studies. The MCyT database has been split into two datasets: its first 50 users have been used for feature selection, this dataset is named $MCyT_S$. The analysis outcomes are used to determine the most reliable reduced subsets of features, which will be evaluated, in term of verification error rates, using the rest of the users, whom are part of the MCyT Test dataset, named $MCyT_T$.

The first two techniques, FR and PCA, are performed using only genuine signatures. Therefore, since there are no skilled forgery signatures the feature selection will be based on identification error rates only. The feature subsets which obtain the best error rates and

feature vector lengths are chosen. However, for the Hellinger Distance, forgery samples (provided in the MCyT database) have been taken into consideration. It will be shown whether the use of forgery samples during feature selection improves the error rate levels achieved, or, if the information obtained from genuine samples only is sufficient for reliable feature selection.

The following subsections will explain in-detail the different feature selection techniques used.

4.3.2 FISHER RATIO

The Fisher Ratio [136] is used to compare the discriminative power among features. The Fisher Ratio has been successfully used in different studies and modalities [126] [120]. This technique is an intuitive method used to measure the discriminative power of features. It expresses the idea that the variability within-classes (values of the features from the same user) should be small, whilst the variability between-classes (values of the features amongst all the users) should be large. In order to calculate this, the variability of the feature within its class (user) has been measured as the variability of the mean of the feature for each class (user). The variability within-classes should be small; this means that this feature is stable for each genuine signature samples from the same user. However, the variability between-classes (between users) is calculated as the variability of the mean of the features amongst all users, a large value shows that this feature is sufficiently distinctive for each user (class). The higher the Fisher ratio (FR), the higher the discriminative power of the feature.

The Fisher Ratio is calculated as follow:

$$FR = \frac{\frac{1}{N}\cdot\sum_{j=1}^{N}\left(m_j-\bar{m}\right)^2}{\frac{1}{N\cdot p}\sum_{j=1}^{N}\sum_{i=1}^{p}\left(x_{ji}-m_j\right)^2} \tag{29}$$

where:

$$\bar{m} = \frac{1}{N}\sum_{j=1}^{N} m_j \tag{30}$$

$$m_j = \frac{1}{p}\sum_{i=1}^{p} x_{ji} \tag{31}$$

Where "$\bar{m}$" is the user mean, "m_j" is the mean of the feature "x" for the class "j", "x_{ji}" is the sample "i" of the feature "x" for the class "j", "N" is the number of users and "p" is the number of samples per user.

Since the variance is affected by the feature range, data normalization is of special concern when calculating the Fisher Ratios. To avoid this effect, all features are normalized within the range [0,1].

Once the Fisher Ratio has been calculated for all the features, they are then sorted using their discriminative power (FR).

4.3.3 PRINCIPAL COMPONENT ANALYSIS FOR FEATURE SELECTION

The Principal Component Analysis (PCA) is used in pattern recognition to perform feature selection [134]. The PCA can be defined as a process to obtain linear combinations which can be used to summarize data, where the least amount of information possible is lost. This method seeks to reduce the sum of the mean squared error. The solution of this problem involves the so-called *scatter* matrix S. Given "p" samples and "L" dimensional data (where "L" means the number of features), X= $\{\overrightarrow{x_1}, \overrightarrow{x_2}, \dots, \overrightarrow{x_p}\}$, the scatter matrix is defined as:

$$S = \sum_{i=1}^{p} (\overrightarrow{x_i} - \overrightarrow{m}) \cdot (\overrightarrow{x_i} - \overrightarrow{m})' \quad (32)$$

Where "m" is the sample mean:

$$\overrightarrow{m} = \frac{1}{p} \sum_{i=1}^{p} \overrightarrow{x_i} \quad (33)$$

The eigenvectors of the scatter matrix S are the principal axes that minimize the sum of the mean squared error. One of the properties of the PCA is that Eigen-values represent the proportion of total variation defined by the Eigen-vectors. The eigenvectors corresponding to the smallest Eigen-values should be selected and reject the feature with the largest coefficients (absolute value). Then the next smallest Eigen-values considered. This principle is consistent with the notion that we regard a component with small Eigen-values as of less importance and, consequently, the variable which dominates it should be of less importance. This process has been carried out to identify the most important features. This process continues obtaining again a sorted list of features by means of their discriminative power measured by their Eigen-values. Continuing this process selects the most significant features suitable for the verification process.

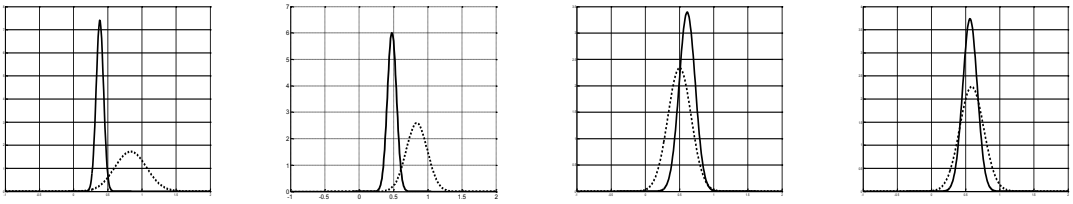
4.3.4 FISHER RATIO COMBINED WITH PRINCIPAL COMPONENT ANALYSIS

As an alternative option, the combination of the Fisher Ratio and the PCA has been analysed. The Fisher Ratio is used as an initial filter, selecting those features which

demonstrate a sufficient discriminative ratio. This is then followed by the PCA analysis to sort the subset selected by the Fisher Ratio.

4.3.5 HELLINGER DISTANCE

To analyze the discriminate power of the features, the distributions generated by each feature are calculated for each user. This is done by comparing the distributions from genuine signature samples with those obtained from forgeries. In contrast to the dimension reduction performed using the Fisher Ratio and the Principal Component Analysis (where only genuine signature samples were used) both genuine and forgery signature samples from the first 50 users of the MCyT have been used.



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

Figure 4-5 Example of feature distribution. Dark line shows the genuine distribution, dot line shows the forgery distribution

In the work presented in this Thesis, it has been hypothesized that features that have less overlap between genuine and forgery distributions (Figure 4-5 (a) and (b)) will achieve improved results for posterior comparison. By removing those features which share more area (Figure 4-5 (c) and (d)) the performance of the ASVS should be improved.

In order to quantify the overlap between the genuine and forgery distributions, the squared Hellinger Distance [135] between two normal distributions has been used. The Hellinger Distance, defined as the distance between probability measures, is calculated using:

$$H^2(P,Q) = \sqrt{\frac{2\cdot\sigma_1\cdot\sigma_2}{\sigma_1^2+\sigma_2^2}}\cdot exp\left(-\frac{1}{4}\cdot\frac{(\mu_1-\mu_2)^2}{\sigma_1^2+\sigma_2^2}\right) \quad (34)$$

Where P and Q are two normal distributions:

$$\begin{aligned} P &\sim N(\mu_1,\sigma_1^2) \\ Q &\sim N(\mu_2,\sigma_2^2) \end{aligned} \quad (35)$$

Here (μ_1,σ_1^2) are the mean and standard deviation from genuine sample features, respectively, while (μ_2,σ_2^2) are the mean and standard deviation from forgery sample features, respectively.

4.4 Feature Selection Applied to The GMM

The 143 features extracted from a signature (detailed in 4.2.1.3.3) have been analyzed using four different feature selection techniques detailed in the previous section, and are: The Fisher Ratio (FR), Principal Component Analysis (PCA), a combination of the FR and PCA (FRPCA) and the Hellinger Distance (HD).

The first three techniques only make use of genuine signatures to perform their analysis. From these three different analyses, the 143 features have been sorted based on their discriminative power, calculated using the three different techniques. Using these three sorted lists, as the outputs for the feature analysis techniques, the identification error rates have been used to select the best possible feature subsets, where the loss in information is as little as possible. Also, several extremely reduced feature subsets have been selected to fit the reduced resource characteristics of tokens.

It has been seen that the Hellinger Distance feature selection technique uses both genuine and forgery signature samples. In this technique the mean overlap of each feature has been calculated and different feature subsets have been chosen by excluding those features which present the most amount of overlap.

4.4.1 Based on Identification Error Rates

In Figure 4-6 the identification error rates are presented, using the $MCyT_S$ dataset, for different lengths of the features vector sorted using the Fisher Ratio (FR), Principal Component Analysis (PCA) and the combination of the Fisher Ratio and Principal Component Analysis (FRPCA).

For the feature vector sorted using the Fisher Ratio (FR), it can be seen that up until a length of approximately 80 features, the error rate level maintains constant. From 80 features lengths to 40 features, the error rates increase slowly, and when the length is lower than 25, there is a great loss in information and the error rates are observed to be high.

The PCA results demonstrate improved performance. The identification process slightly improves in performance when the number of features decreases from 143 to 60. From 60 features to 30, the error rates maintain a level of 2.5%. When the length is below 30 features, the error rates begin to increase, maintaining reasonable levels as far down as 13 features, at this point the error rates begin to increase rapidly.

It may be concluded from these results, that the PCA feature selection achieves better results than the Fisher Ratio, demonstrating an improved performance when considering the discriminative power of individual features.

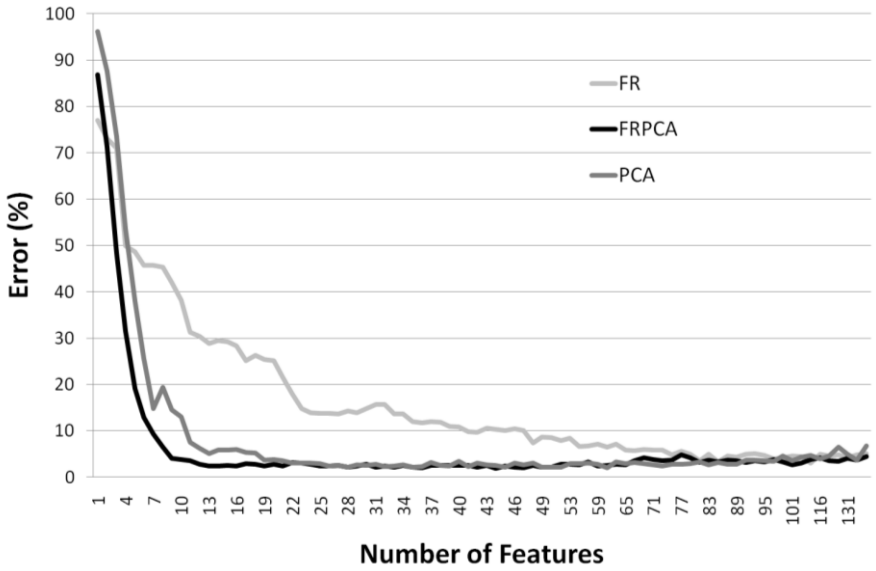


Figure 4-6 Identification Error Rates vs. Number of Features

The third option tested makes use of a combination of both methods. First, the features with a low Fisher Ratio were removed from the vector, excluding 54 features. With the remaining features, the PCA analysis was carried out. This option, FRPCA, has improved on the results obtained from each of the other two techniques used individually, as it can be seen in Figure 4-6.

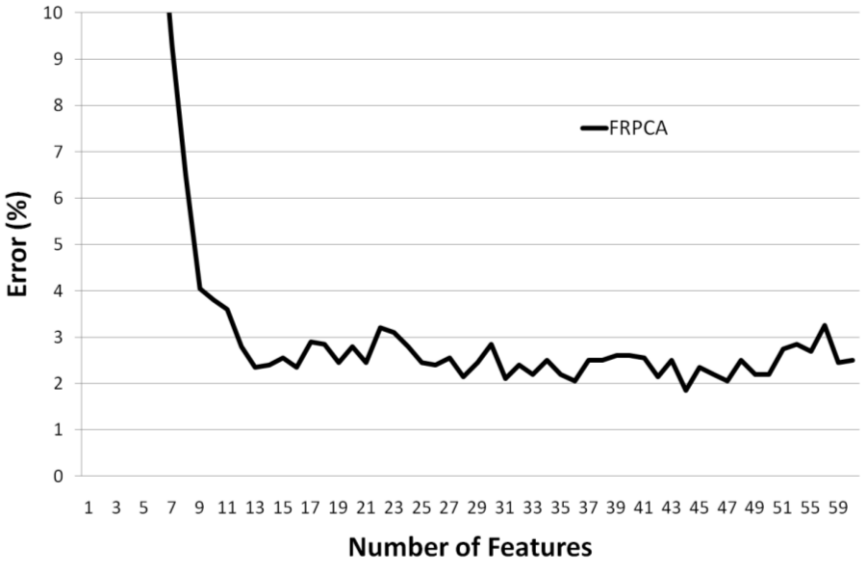


Figure 4-7 Identification Error Rates vs. Number of Features, Fisher Ratio combined with PCA Analysis, Detail

It may be seen in both Figure 4-6 and Figure 4-7 that using the FRPCA the identification error rates have a similar behaviour than using only the PCA, improving as far down as 40 features, maintaining constant down to 30 features and slightly growing as far down as 13, at

this point the error rates increase dramatically. But, especially for low numbers of features, the FRPCA improves on the identification error rates achieved by the PCA.

The combination of both the Fisher Ratio and the PCA has demonstrated improved results, where more information is kept for feature vector lengths of less than 30 where a reduced subset of only 13 features is shown to maintain a low identification error.

Based on these results, the combination of the Fisher Ratio and PCA demonstrates the best performance. From this analysis, three different features subsets have been selected. The first, attempts to obtain a reduced feature vector for low level resource systems in both computational power and storage, as it only requires a vector length of 13 features. The second subset is made up of 28 features, where this is a compromise between obtaining a small feature vector length and loss in as little information as possible. The third and final subset is composed of 44 features; this is the feature vector that is observed to obtain the lowest error rate (1.8%).

The 13 Feature Vector Subset, named "FRPCA 13", contains the following features:

Table 4-8 GMM 13 Features Vector Subset selected using FRPCA

FRPCA 13		
$Length\_Area$	$y_{\text{end\_min}}$	el_{mean}
$\boldsymbol{x_{\text{max\_end}}}$	$y_{zero\_crossing}$	$vx_{\text{max\_min}}$
$\boldsymbol{x_{\text{ini\_min}}}$	p_{rms}	$vx_{\text{mean\_over0}}$
$\boldsymbol{x_{std}}$	p_{std}	
$\boldsymbol{y_{\text{ini\_min}}}$	$p_{zero\_crossing}$	

The second subset, with 28 features, "FRPCA 28", contains all the previous selected features plus the following ones:

Table 4-9 GMM 28 Features Vector Subset selected using FRPCA

FRPCA 28		
$T\_writting$	$vp_{\text{mean\_over0}}$	$vin_{\text{time\_over0}}$
$\boldsymbol{x_{\text{end\_min}}}$	$vp_{zero\_crossing}$	$ax_{\text{time\_over0}}$
$\boldsymbol{vx_{zero\_crossing}}$	$vaz_{\text{time\_over0}}$	$ax_{mean\_\text{max}\_ratio}$
$\boldsymbol{vx_{zero\_crossing}}$	$vaz_{zero\_crossing}$	$ax_{\text{max\_mean}}$
$\boldsymbol{p_{\text{time\_over0}}}$	vaz_{rms}	$ap_{\text{mean\_over0}}$

For the third subset, "FRPCA 44", with 44 features the following features have been added:

Table 4-10 GMM 44 Features Vector Subset selected using FRPCA

FRPCA 44		
N_Strokes	$vp_{\text{max\_mean}}$	aaz
$\boldsymbol{y_{mean}}$	$vaz_{\text{time\_over0}}$	$ain_{\text{time\_over0}}$
$\boldsymbol{y_{\text{rms}}}$	$ax_{\text{mean\_over0}}$	$ain_{\text{max\_mean}}$
$\boldsymbol{vx_{mean\_max\_ratio}}$	$vay_{\text{time\_over0}}$	$ain_{\text{mean\_below0}}$
$\boldsymbol{vy_{\text{time\_over0}}}$	$ap_{zero\_crossing}$	
$\boldsymbol{vy_{mean\_max\_ratio}}$	$ap_{mean\_max\_ratio}$	

4.4.2 Features Selection Based on Hellinger Distance

In Figure 4-8 the overlap between the genuine and forgery distributions measured for each of the 143 features analysed using the Hellinger Distance (see 4.3.5) are presented. The values on the vertical axis are the averages of the Hellinger Distance among all users, the horizontal axis represents the feature (described in 4.2.1.3.3). It may be observed that the range in overlap lies between 40% and 90%.

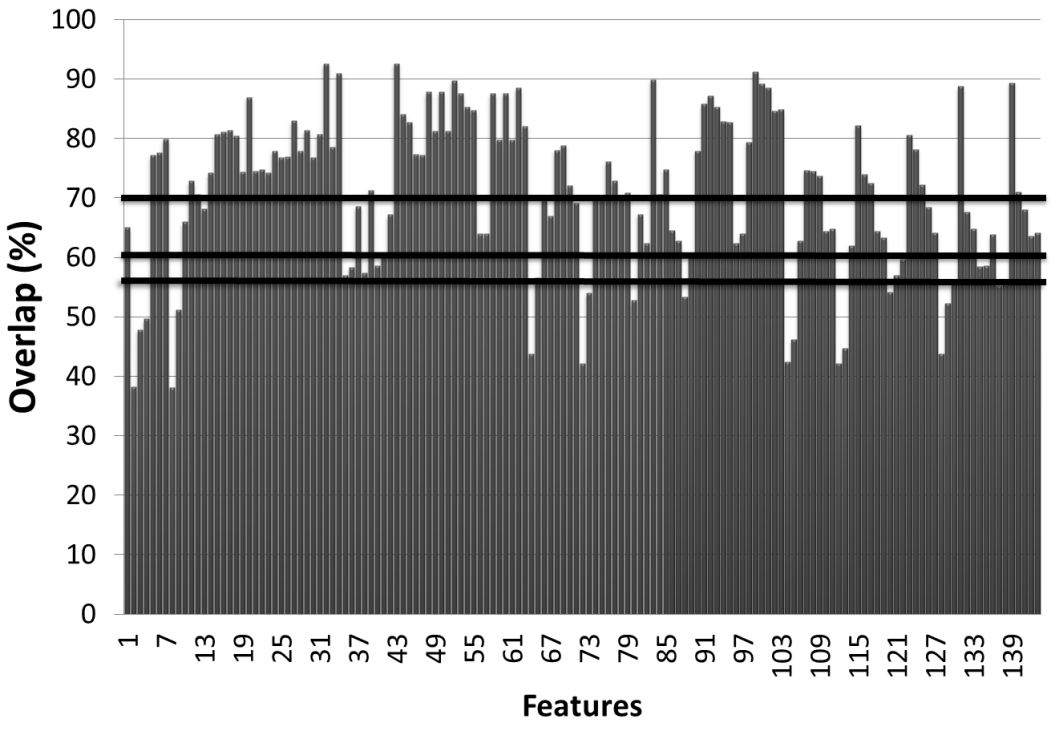


Figure 4-8 Overlap for GMM features analysed

Three different subsets have been considered from this outcome and are represented on Figure 4-8 by 3 black horizontal lines. The first subset, named “HD 16”, consists of features which have an overlap of less than 55%, below the lowest black line. The subsets named “HD 28” and “HD 60” contain features with an overlap of less than 60% and 70%, respectively.

The first mask, HD 16, is composed of the following 16 features:

Table 4-11 GMM 16 Features Vector Subset selected using HD

HD 16		
Time	$vy_{zero\_crossing}$	$ay_{zero\_crossing}$
T_writting	$vp_{\mathrm{time\_over0}}$	$ap_{\mathrm{time\_over0}}$
Length	$vaz_{\mathrm{time\_over0}}$	$vx_{\mathrm{time\_over0}}$
Length_Area	$ax_{\mathrm{time\_over0}}$	$ael_{\mathrm{mean\_below0}}$
$\boldsymbol{vy_{\mathrm{time\_over0}}}$	$ax_{zero\_crossing}$	
$\boldsymbol{vy_{\mathrm{time\_over0}}}$	$ay_{\mathrm{time\_over0}}$	

The HD 28, is formed from the same 17 features detailed above as well as the following 12:

Table 4-12 GMM 28 Features Vector Subset selected using HD

HD 28		
$\boldsymbol{p_{mean}}$	$vx_{zero\_crossing}$	$aaz_{\mathrm{mean\_over0}}$
$\boldsymbol{p_{\mathrm{rms}}}$	$ap_{zero\_crossing}$	$aaz_{\mathrm{mean\_below0}}$
$\boldsymbol{p_{\mathrm{mean\_min}}}$	ap_{rms}	$ael_{zero\_crossing}$
$\boldsymbol{p_{\mathrm{rms\_min}}}$	aaz_{rms}	ael_{rms}

Finally, the HD 60 is made up of all the previous features as well as the 32 features listed below:

Table 4-13 GMM 60 Features Vector Subset selected using HD

HD 60		
N_Strokes	vp_{rms}	$ap_{\text{mean\_over0}}$
$\boldsymbol{x_{\text{max\_}end}}$	$vp_{\text{mean\_over0}}$	$ap_{\text{mean\_below0}}$
$\boldsymbol{x_{mean}}$	$vp_{\text{mean\_below0}}$	$az_{\text{max\_min}}$
$\boldsymbol{p_{\text{max\_rms}}}$	$vaz_{zero\_crossing}$	$az_{\text{max\_min}}$
$\boldsymbol{p_{std}}$	$vel_{\text{mean\_over0}}$	$az_{\text{max\_min}}$
$\boldsymbol{p_{zero\_crossing}}$	ax_{rms}	$el_{\text{max\_min}}$
$\boldsymbol{el_{mean}}$	$ax_{\text{mean\_over0}}$	$el_{\text{max\_min}}$
$\boldsymbol{el_{rms}}$	$vax_{\text{mean\_below0}}$	$el_{\text{mean\_below0}}$
$\boldsymbol{vx_{rms}}$	$vax_{\text{mean\_below0}}$	
$\boldsymbol{vx_{mean\_\text{max}\_ratio}}$	ay_{rms}	
$\boldsymbol{vx_{\text{mean\_below0}}}$	$ay_{\text{mean\_over0}}$	
$\boldsymbol{vp_{zero\_crossing}}$	$ay_{\text{mean\_below0}}$	

4.4.3 VERIFICATION ERROR RATES FOR THE SELECTED FEATURES SUBSETS

Once the different features vector subsets have been identified and selected, their performance was tested using the GMM-based signature verification system described in the previous section (4.2.1).

The tests were carried out using the last 50 users from the MCyT database ($MCyT_T$) [21]. Five random genuine signature samples were selected from each user to obtain the GMM user models. The 20 remaining genuine and 25 skilled forgery signatures were also used, making a total test set of 1000 authentic comparisons and 1250 forgery comparisons. Also, all the genuine signature samples from other users were compared against each user model to obtain the error rates for random forgeries, this has resulted in 61250 comparisons. This process was repeated 10 times to obtain sufficient results for a statistic analysis.

In Figure 4-9 the skilled forgery ROC-graphs for different feature vector lengths is presented, the light colored lines represent the subset selected using the Hellinger Distance (HD) and the black lines are the subsets selected using the combined Fisher Ratio and Principal Component Analysis (FRPCA).

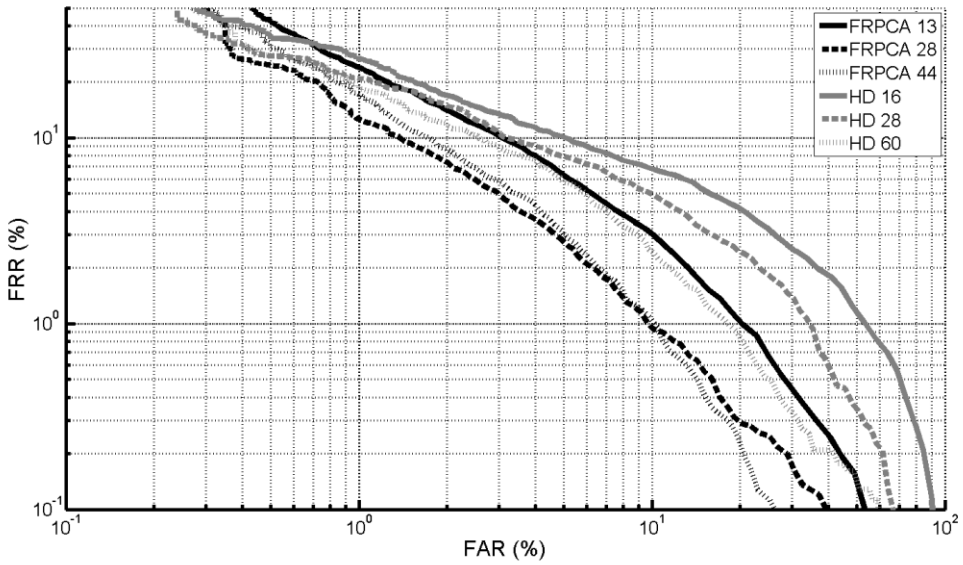


Figure 4-9 GMM Signature Verification System Results for Skilled Forgeries, length analysis of the feature vector

The Equal Error Rates (ERR) obtained for skilled forgeries are detailed in Table 4-14. It can be seen from both Figure 4-9 and Table 4-14 that the subsets selected using the combined FR and PCA obtain improved results when compared to those selected using the HD. This result is of particular interest as the HD subsets were selected to maximize the distribution distances between genuine and forgery samples and the subsets selected using the combined FR and PCA maximize the discriminative power between users (identification) where no forgery samples were used. These results demonstrate how genuine samples and feature selection based on identification error rates is a useful technique for achieving high-quality verification error rates

Table 4-14 Equal Error rates for skilled forgeries and different features subsets

	FRPCA_13	FRPCA_28	FRPCA_44	HD_16	HD_28	HD_60
ERR Skilled Forgeries	5.6%	3.8%	4.0%	7.7%	6.7%	5.4%

Focusing on the FRPCA subsets alone shows that the 13 feature subset performs worse than the other 2 subsets selected using the FRPCA, where an EER of close to 5.4% is observed. However, the performance is still considered as good when compared to published results for GMM Signature Verification systems (Table 4-1) and also taking into account the small number of features that the vector contains. This small number of features also implies a much reduced size for both sample and template references (user model) and reduced computational requirements.

There are no significant differences between the 28 and 44 feature vector performances, achieving an EER of 4.0% for 44 features and 3.8% for the 28 feature vector, this is in agreement with the state of art presented in Table 4-1 [83]. The size is lower for both sample and user model which implies reduced computational requirements when compared with the results shown in Table 4-1. In [83] a GMM with 64 for Gaussians mixture components is used, while in the system presented here only 4 are required.

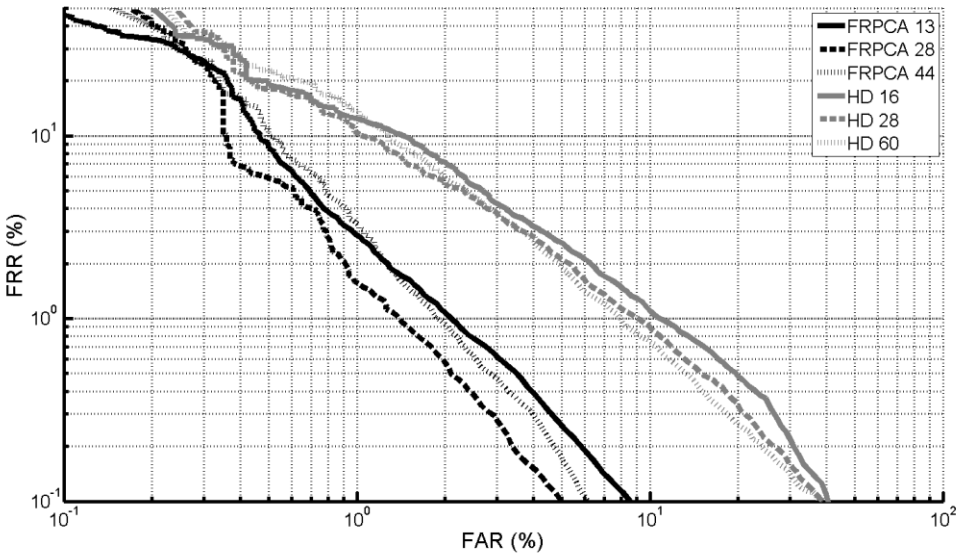


Figure 4-10 GMM Signature Verification System Results for Random Forgeries, length analysis of the feature vector

In Figure 4-10 and Table 4-15, the results for the random forgery are presented. In this case, the superior performance is clearly seen for the subsets selected using the FRPCA when compared to those obtained using the HD. This may be explained by the maximization of the discriminative power between users owed to the FRPCA, whereas the HD only considers genuine and forgery feature distributions for each individual user and attempts to obtain improved results for skilled forgeries.

Table 4-15 Equal Error rates for random forgeries and different feature subsets

	FRPCA_13	FRPCA_28	FRPCA_44	HD_16	HD_28	HD_60
EER Random Forgeries	1.6%	1.3%	1.5%	3.5%	3.3%	3.3%

The Equal Error Rate achieved for the subsets selected using the FRPCA for random forgeries is close to 1.5%, and maintains the same level for the three different subsets, regardless of the feature vector length. The same EER results are obtained for the feature subsets selected using the HD, however, in this case with an EER close to 3.5%.

4.5 Feature Selection Applied to DTW

The dimensionality reduction study and the performance evaluation of the DTW on-line signature verification system proposed have been carried out using the same methodology described in 4.4 for the GMM.

4.5.1 Based on Identification Error Rates

The identification error rates for different feature vector lengths sorted using the Fisher Ratio (FR), Principal Component Analysis (PCA) and the combination of the Fisher Ratio and Principal Component Analysis (FRPCA) are shown in Figure 4-11.

The identification error rates obtained using the DTW algorithm demonstrates a significant improvement in performance when compared with results from the GMM algorithm. However, when considering the comparison among the three feature selection techniques, the results obtained are similar to those obtained with the GMM algorithm (shown in 4.4.1).

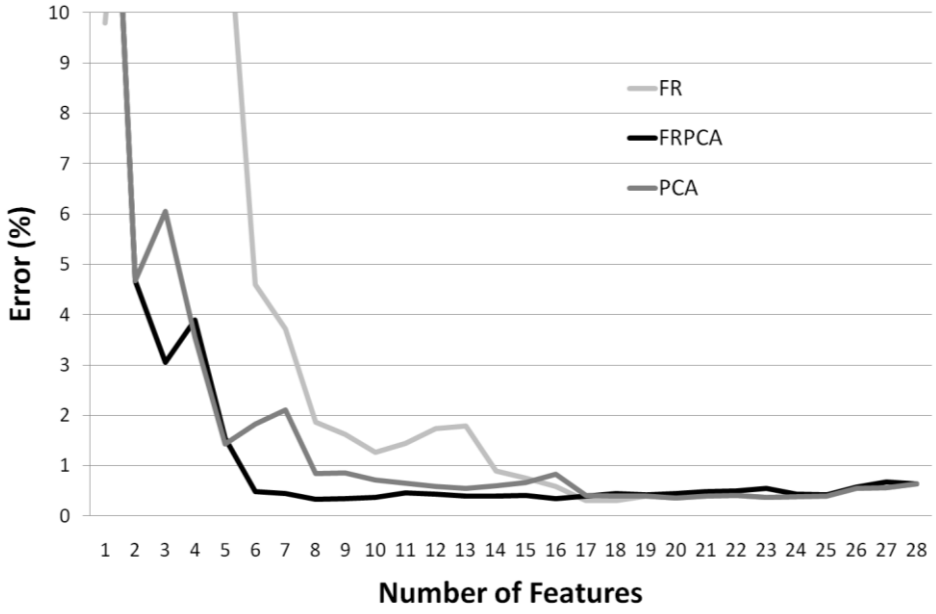


Figure 4-11 Identification Error Rates vs. Number of Features for the DTW Algorithm

The Fisher Ratio (FR) technique demonstrates the worst performance, maintaining reasonable identification error rates only as far down as a feature subset of 15. The Principal Component Analysis (PCA) improves on the FR and maintains low error rates down to a subset composed of 8 features.

The best results are obtained by the FRPCA combination. This technique removes the features with low Fisher Ratios (in this case 10 features are removed), and the PCA is applied to the remaining set of features. The FRPCA maintains good error rates for only 6 features, obtaining its minimum error rate for a subset of 8 features.

Based on these results, the combined Fisher Ratio and PCA were used to select 2 feature subsets. The first has been composed of only 6 features; this is in accordance with the aim of minimizing the size of the user template. The second subset is composed of 8 features, as it has demonstrated the best error rate results.

The 6 Feature Vector Subset, named "FRPCA 6", contains the following features:

Table 4-16 DTW 6 Features Vector Subset selected using FRPCA

FRPCA 6		
Absolute velocity	Sine of angle α	Sin of angle Φ
Cosine of angle α	Angle Φ	Coordinate x

The second subset, with 8 features, "FRPCA 8", contains all the previous selected features in addition to the following:

Table 4-17 DTW 8 Features Vector Subset selected using FRPCA

FRPCA 8	
Cosine of angle Φ	Pseudo Distance

4.5.2 FEATURES SELECTION BASED ON HELLINGER DISTANCE

The feature distributions were calculated for every user in the $MCyT_S$, using all genuine and forged signatures and based on results from 10 simulations. The results have been averaged amongst users. In each simulation, the DTW user models created were based on 5 randomly chosen genuine signature samples.

In Figure 4-12 the overlap between the genuine and forgery distributions measured for all the pseudo-distances are presented. The values on the vertical axis are the average of the Hellinger Distance among all users, the horizontal axis represents the features described in 4.2.2.3.2. It can be seen that most of the features demonstrate little overlap, this means that they are good candidates for discrimination between forgeries and genuine signatures. This is true for all sets where the common area is less than 15%.

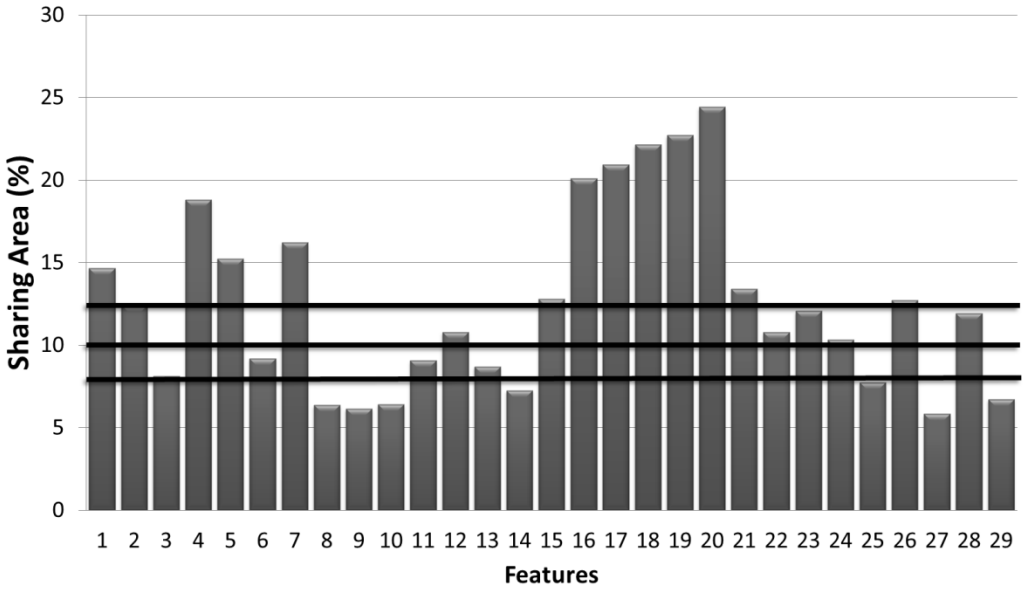


Figure 4-12 Common Distribution Area for Pseudo-distances

The pseudo-distances are seen to have a much lower overlap between forgery and genuine signatures than the features analysed using the GMM algorithm.

In Figure 4-12 it may be observed that several features, numbered from 16 to 20, stand out negatively:

(16) Azimuth angle, $\mathbf{az}(\mathbf{t})$ (18) Velocity of azimuth angle, $\boldsymbol{v_{az}}(\mathbf{t})$ (20) Curvature radius, $\boldsymbol{r}(\boldsymbol{t})$

(17) Inclination angle, $\boldsymbol{in}(\mathbf{t})$ (19) Velocity of inclination angle, $\boldsymbol{v_{in}}(\boldsymbol{t})$

This particular conclusion is in agreement with results from other published works [52] [58] [137], where it was indicated that the tilts captured by the input devices do not improve the performance of signature verification systems.

Three different feature subsets were composed from the results presented in Figure 4-12, where three different thresholds were set, these are represented as black lines in the figure.

The first, is for those features where the overlap is lower than 7.5% (lowest black line), named “HD 6”, and contains the 6 following features:

Table 4-18 DTW 6 Features Vector Subset selected using HD

HD 6		
Angle α	Sine of angle α	Pseudo Distance
Cosine of angle α	Pressure	Pseudo Pressure

The second feature subset is composed of those features where the overlap is less than 10% (mid blank line), named "HD 11", and contains the same 6 features from the previous subset in addition to the following:

Table 4-19 DTW 11 Features Vector Subset selected using HD

HD 11		
Absolute velocity	Sin of angle Φ	The ratio of the minimum over the maximum velocity for a window of 5 points, centred on the current point
Absolute acceleration	Angle Φ	

The last subset includes features where the overlap is less than 12.5% (highest black line), named "HD 17", where the following 6 features are added:

Table 4-20 DTW 17 Features Vector Subset selected using HD

HD 17		
Velocity in y	Coordinate y	The length to width ratio for windows of size 5, centred on the current point
Cosine of angle Φ	Pseudo Shape	The length to width ratio for windows of size 7, centred on the current point

4.5.3 VERIFICATION ERROR RATES FOR THE SELECTED FEATURES SUBSETS

Using the 5 different feature vector subsets selected by the FRPCA and the HD, the DTW algorithm verification performance was tested using the $MCyT_T$ and by following the same methodology explained in 4.4.3: 5 randomly chosen signatures are used to train the user model, 20 genuine and 25 forgery signatures for each user to obtain the skilled error rates, and 25 genuine signatures from each of the other users to obtain the random error rates. All were repeated 10 times to obtain sufficient information for a statistical analysis.

In Figure 4-13 the results obtained for the different features subsets selected are presented. As opposed to the GMM results, for the DTW algorithm the feature vectors selected using the Hellinger Distance (HD) technique have obtained much improved error rates than those selected using the combined Fisher Ratio and Principal Component Analysis (FRPCA).

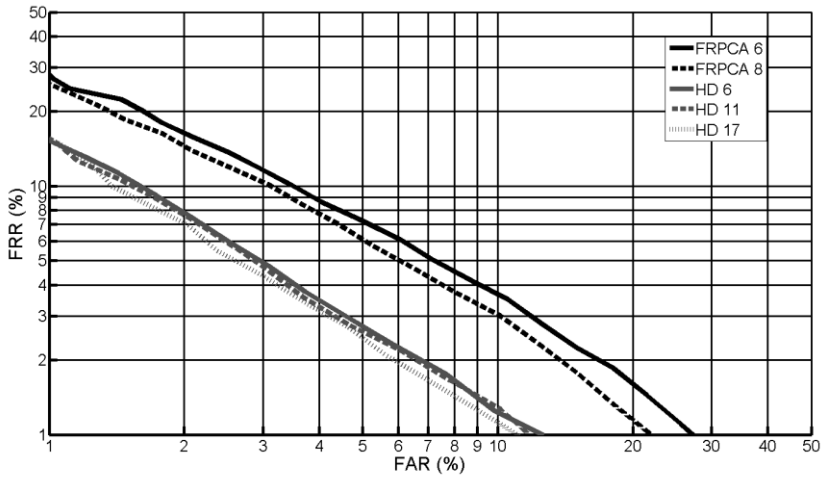


Figure 4-13 DTW Signature Verification Results for Skilled Forgeries, different feature vector Analysis

The equal error rates are detailed in Table 4-21. These results show that the HD feature vectors obtain close to 3.7% EER for all of the subsets.

Table 4-21 DTW Equal Error rates for skilled forgeries and different features subsets

	FRPCA 6	FRPCA 8	HD 6	HD 11	HD 17
ERR Skilled Forgeries	6,0%	5,5%	3,7%	3,6%	3,6%

In the Random Forgeries case, the results are presented in Figure 4-14 and Table 4-22. Again the HD subsets demonstrate improved performance when compared to the FRPCA subsets. The three HD subsets obtain a remarkable 0.5% EER for random forgeries, this result is in agreement with results observed from the state-of-art.

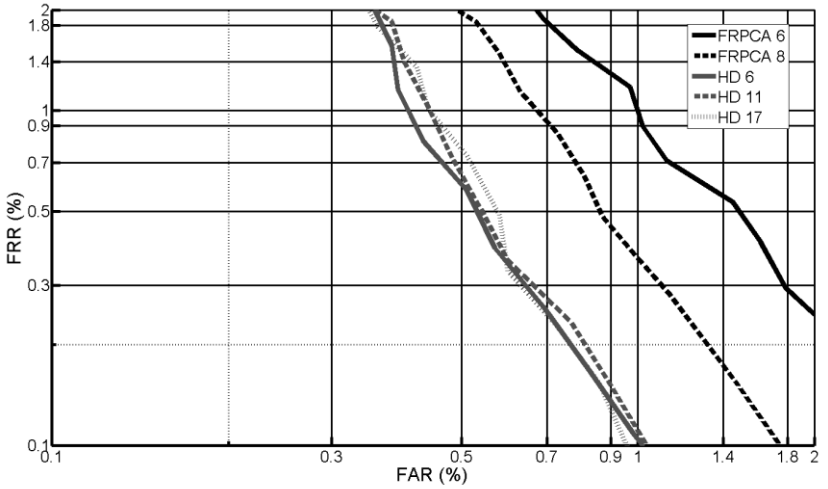


Figure 4-14 DTW Signature Verification Results for Random Forgeries, feature vector Analysis

Table 4-22 DTW Equal Error rates for Random forgeries and different features subsets

	FRPCA 6	FRPCA 8	HD 6	HD 11	HD 17
EER Random Forgeries	1.0%	0,8%	0,5%	0,5%	0,5%

It is worth highlighting the high-quality of the results achieved from the smallest feature subset that is formed by only 6 features, these are related to velocity (angle α) and pressure. These results are in agreement with the conclusions obtained from other studies [58] [124] and indicate that the tilt information is not relevant for signature verification tasks.

These results are also in agreement with the state of the art, in particular for random forgeries where 0.5% EER has been obtained. The results released by the two signature competitions SVC2004 [22] and BIOSEC2009 [70] based also on the DTW demonstrate better performance than the system proposed for skilled forgeries, but it is also worth highlighting that a slightly bigger error level has been obtained than that published by Van Bao Ly et all [80] (3.4% for skilled forgeries) but using less than 4 times the amount of signals (6 instead 25) and reducing the User Model to only 4 vectors (coordinates x and y, pressure and time) of 256 sample points, which results in a reduction for computational load and storage resources.

4.6 Analysis of the GMM Algorithm Parameters

This section will analyze two parameters of the GMM algorithm. The first, is the number of Gaussian probabilistic functions used to create the GMM model. This parameter is important due to the impact of the number of Gaussian on the computational load and in the size of the user model for the GMM algorithm.

The second parameter analyzed is the number of signature samples taken for the enrolment process. This number has not been homogenous throughout the literature, where numbers between 1 and 10 are used. Generally, it is agreed that 3 signature samples for the enrolment represents the minimum sample set size required and takes into account the intrinsic variability of the user signature. In both signature evaluation campaigns, SVC'2004 and BSEC'2009, 5 signature samples were used for the enrolment process. It is generally accepted that 5 signatures are the best option as they provide sufficient statistical information on the intrinsic user variability and also presents a user friendly enrolment [138]. More than 10 signature samples make the enrolment process too tedious for users, therefore 9 signature samples will be the maximum value analyzed. Summarizing, values from 3 to 9 signature samples have been used and the performance results using the GMM algorithm will be presented.

4.6.1 Number of Gaussians in the GMM model

The number of Gaussians (M) to use in the GMM System has been analysed, where 4 to 32 Gaussians have been tested.

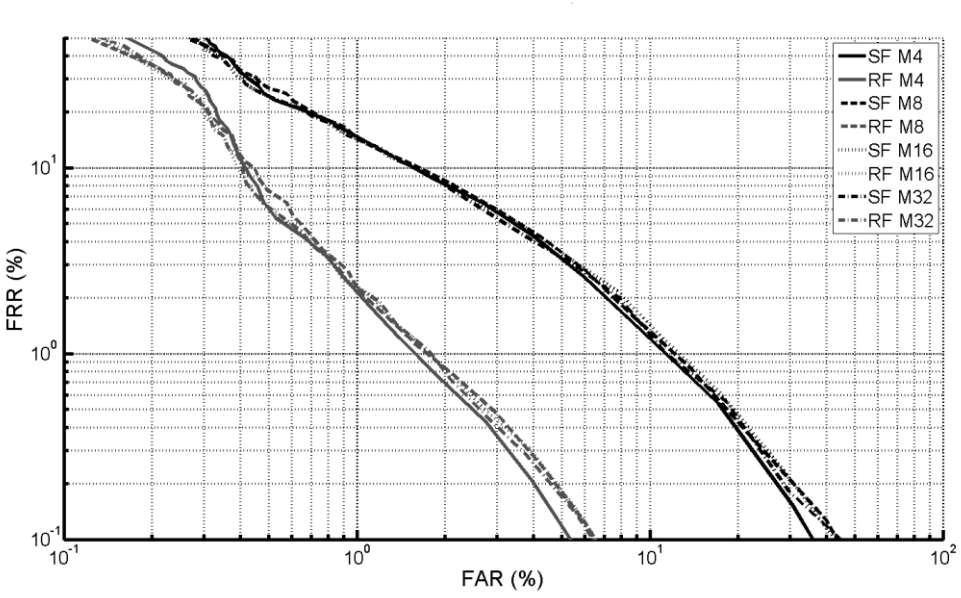


Figure 4-15 GMM Signature Verification System Results, number of Gaussians Analysis

From Figure 4-15 and Table 4-23 it can be concluded that there are no significant differences amongst all the tests performed, where it can be seen that 4 Gaussians are sufficient to model the systems probabilistic space. Increasing the number of Gaussians does not lead to improved performance of the system described above, and would imply a much increased computational load.

Table 4-23 GMM equal error rates, number of Gaussians analysis

NUMBER OF GAUSSIANS	4	8	16	32
ERR Skilled Forgeries	4.1%	4.2%	4.2%	4.1%
ERR Random Forgeries	1.4%	1.4%	1.4%	1.4%

4.6.2 NUMBER OF SIGNATURES SAMPLES TAKEN FOR THE ENROLMENT PROCESS

The number of sample signatures taken for the enrolment process is an important issue in all biometric systems. Generally, the more signature samples requested the better the performance, however this implies greater efforts on behalf of the user for enrolment.

The typical number of samples captured for enrolment is between 3 and 10. Less than 3 typically implies an inadequate estimation of the intrinsic variability of the user signature, while more than 10 results in a tedious process.

In Figure 4-16 and Table 4-24 the performance improvement for 3 signatures samples to 9 may be observed. In the case of only 3 signature samples, the model maintains sufficient performance with an EER of 5%. This result is in agreement with the state of the art for GMM systems.

Table 4-24 GMM equal error rates, training samples analysis

Enrolment Samples	3	5	7	9
ERR Skilled Forgeries	5.0%	3.9%	3.3%	3.1%
ERR Random Forgeries	1.9%	1.3%	1.0%	1.0%

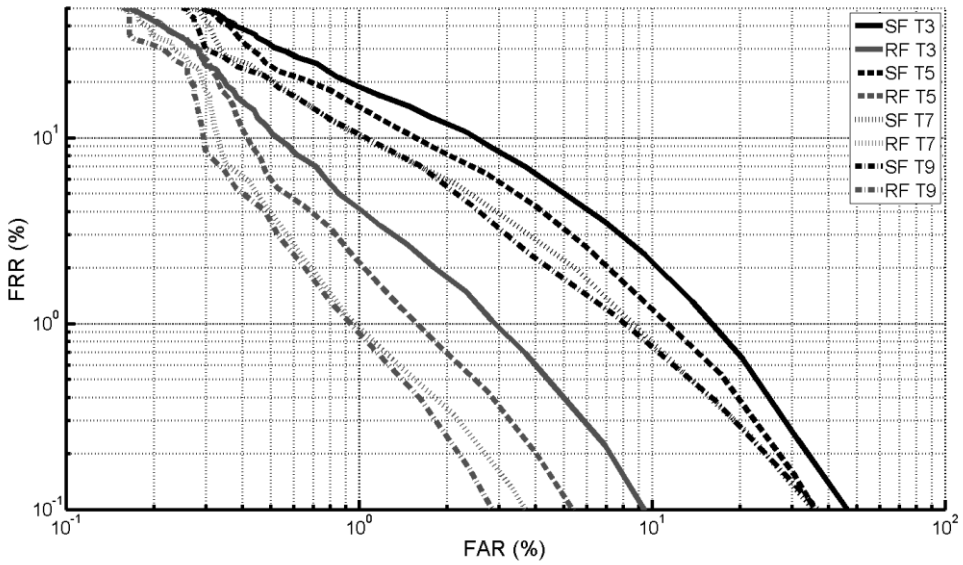


Figure 4-16 GMM Signature Verification System Results, Training Samples Analysis

The greater the number of signatures taken during the enrolment process, the better the performance. For Skilled Forgery error rates, the results show an EER of 3.9 % for 5 samples, 3.3% for 7 samples and 3.1% for 9 samples. For Random Forgery error rates the EER ranges from 1.9% to 0.9%.

This clearly demonstrates that the greater the number of signatures taken during the enrolment phase the better the system performs. However, user acceptance must be considered when designing a real system. A tradeoff between both of these factors, performance vs. usability, should be made, while at the same time considering the level of security required.

4.6.3 User Model Size

The number of elements in the feature vector has an impact on the size of the user model. The model size for the GMM algorithm, as it was detailed in section 4.2.1, is represented by a set of three different parameters:

- mean vectors "μ_i",
- diagonal covariance matrixes "Σ_i",
- weight factors "c_i".

All of the above for i=1...M, where "M" is the number of Gaussians functions. The size of the user models, storing these elements as floating point values within 4 bytes is:

- $size(\mu) = M \cdot L \cdot 4bytes$,
- $size(\Sigma) = M \cdot L \cdot 4bytes$,
- $size(c) = M \cdot 4bytes$.

Where L is the number of elements of the feature vector.

Taking into consideration the 3 different feature vector subsets selected by the FRPCA, which have demonstrated the best performances, the user model size for the 3 feature subset and different number of Gaussians Functions is presented in the following table:

Table 4-25 User model size for different feature subsets and different numbers of Gaussian functions

Feature Vector Length / Number of Gaussians Functions	FRPCA 13 (KBytes)	FRPCA 28 (KBytes)	FRPCA 44 (KBytes)
4	0.44	0.91	1.41
8	0.88	1.81	2.82
16	1.75	3.63	5.63
32	3.50	7.26	11.26
64	7.01	14.51	22.52

It may be observed from Table 4-25, that the user model size varies greatly, depending on the GMM configuration (number of Gaussian functions) and the elements of the feature vector. It can be as little as 0.44 Kbytes for a vector of 13 elements and only 4 Gaussian functions, and up to 22 Kbytes for 44 features and 64 Gaussian functions.

An interesting advantage of the GMM algorithm proposed is that it achieved the same error rates for different numbers of Gaussian functions, where it has been seen that as few as 4 functions may be used. With this configuration, the user model sizes vary from 0.44 KBytes for 13 features to 1.41 KBytes for 44. This reduced size is ideal for systems which are limited in size.

4.6.4 COMPUTATIONAL LOAD

The computational load of the GMM algorithm proposed is based on two factors. Firstly, the feature calculation from the raw data acquired by the input device. Secondly, the Gaussian function calculation required to determine the likelihood of the sample signature with the user model.

The impact of the feature vector length and the number of Gaussian functions on the computational load is shown in Table 4-26. The times taken for comparisons are shown in

milliseconds. This time has been calculated by implementing the GMM in Matlab®, running on a computer with the following characteristics: Intel® Core® 2 Duo E670@2.66GHz, 3 GBytes RAM, Windows® 7 32 bits. The time represents the average time after calculating 1000 comparison scores

Table 4-26 Comparison time for different feature subsets and different number of Gaussian functions

Feature Vector Length / Number of Gaussians Functions	FRPCA 13 (ms)	FRPCA 28 (ms)	FRPCA 44 (ms)
4	4.1	4.2	4.5
8	4.2	4.6	5.1
16	4.5	5.2	5.9
32	5.1	6.2	7.9
64	6	8.3	11.8

In the case of the impact on the computational load, the length of the feature vector does not imply a great change on the time taken to perform a comparison when the number of Gaussian functions is low. When the number of Gaussian functions goes beyond 32, the Gaussian calculations begin to prevail in the comparison time. In the case of 64 Gaussian functions, the number of features has a relevant impact, demonstrating close to double the time required (from 6 ms to 11.8 ms).

Again, the GMM algorithm proposed has the advantage of obtaining state-of-the art EERs for a configuration containing only 4 Gaussian functions, where the time per comparison remains low and stable for different feature vector sizes.

4.7 ANALYSIS OF THE DTW ALGORITHM PARAMETERS

As previously carried out for the GMM algorithm, an analysis of the impact of two different factors on the performance of the DTW algorithm is presented in this section. These two parameters are the number of equi-spaced points and the number of signature samples used for the enrolment process.

The first is related to the computational load and the user model size of the DTW algorithm. The impact of this factor is analyzed in the following section 4.7.1.

The number of signature samples taken for the enrolment process was discussed in Section 4.6. The impact on the performance of the DTW algorithm will also be analyzed using signature sample values between 3 and 9 in section 4.7.2.

4.7.1 NUMBER OF EQUI-SPACED POINTS USED

The DTW preprocessing steps include the transformation of the original temporal signals acquired at a fixed sampling rate of 100Hz, into equi-spaced 256-point temporal signals by means of Linear Interpolation. This step normally reduces the number of sample points within the temporal series channels acquired by the input devices (x and y axes, pressure and tilts), except for the case of very short signatures.

In this section the impact of the number of equi-spaced points has been analyzed, evaluating different configurations with 64, 128, 256, 512 and 1024 equi-spaced points. The feature subset used has been HD 6, and is composed of the 6 features detailed in 4.5.2.

In Figure 4-17 and Table 4-27 the results obtained (ROC graphs and EER) for skilled and random forgeries are presented. It may be seen that both random and skilled error rates improve as the number of equi-spaced points increase. For random forgeries the error rate decreases from 1.2% (64 points) to 0.5% (1024 points) and for skilled forgeries from 6.2% to 3.4% for 64 and 1024 points, respectively.

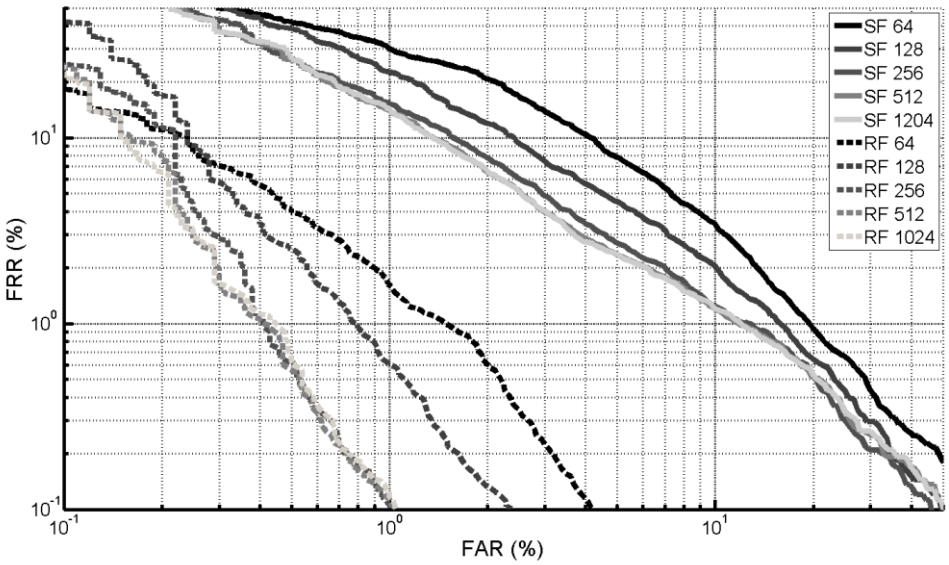


Figure 4-17 DTW error rates, number of Equi-Spaced points analysis

Table 4-27 DTW equal error rates, number of Equi-Spaced points analysis

Equi-Spaced Points	64	128	256	512	1024
ERR Skilled Forgeries	6.2%	4.7%	3.6%	3.4%	3.4%
ERR Random Forgeries	1.2%	0.8%	0.5%	0.5%	0.5%

Two important conclusions may be derived from Figure 4-17 and Table 4-27. First, there are no differences observed between 512 points and 1024 points for both random and skilled error rates, therefore it has been shown that 512 equi-spaced points are sufficient to maintain the relevant information for performing signature verification. This conclusion leads to a significant size reduction in the user template, especially when considering that new capture devices will increase their sampling rate to 200Hz.

Regarding a lower limit for the number of equi-spaced points, the EER obtained for 64 points indicates that relevant information is lost. On the other hand 128 points maintains a good EER. This amount of points is considered as a suitable option for systems with limited storage and/or computational resources, as the error rates only increase moderately, obtaining 4.7% for skilled forgeries and 0.8% for random forgeries.

4.7.2 NUMBER OF SIGNATURES SAMPLES TAKEN FOR THE ENROLMENT PROCESS

The improvement on the error rates from the different number of signatures taken during the enrolment process was analysed using the DTW. In Figure 4-18 and Table 4-28 are the results obtained (ROC graphs and EER) for both skilled and random forgeries.

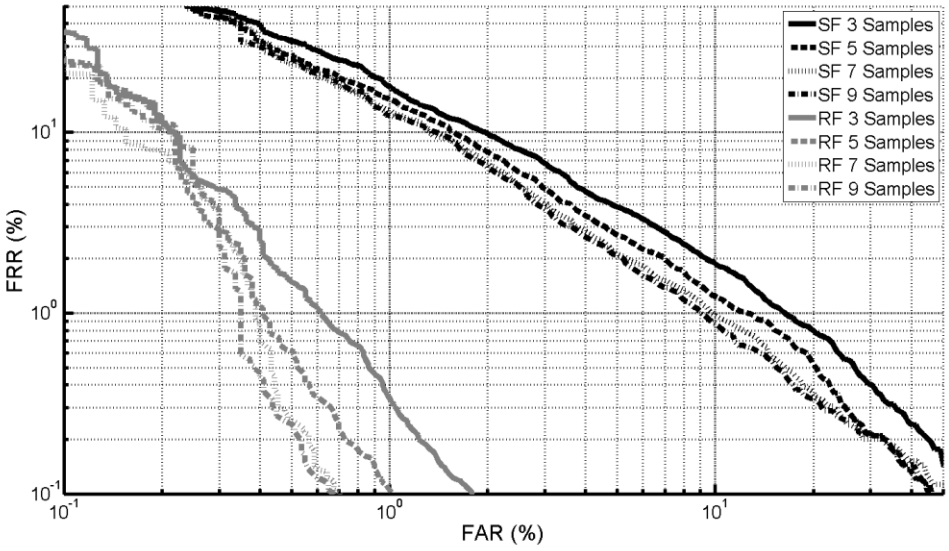


Figure 4-18 DTW Equal Error rates, training samples analysis

As expected, the higher the number of signature training samples used, the lower the error rate obtained. However, unlike the GMM algorithm, the difference between using only 3 signatures and the more common 5 signatures are observed to have little impact on the error rates. This is particularly true for skilled forgery error rates, as the EERs only decrease from 4.3% to 3.6%, indicating that 3 signatures are sufficient in the DTW algorithm for reasonable equal error rates.

Table 4-28 DTW Equal Error rates, training samples analysis

Enrolment Samples	3	5	7	9
ERR Skilled Forgeries	4.3%	3.6%	3.4%	3.2%
ERR Random Forgeries	0.7%	0.5%	0.4%	0.4%

It may also be seen in Figure 4-18 that by increasing the number of signatures beyond 5 does not have a significant effect on the error rates, where a 3.4% ERR was obtained for 7 training samples, decreasing to 3.2% for 9 training samples.

In the case of random forgeries the differences are observed to be even smaller than those shown for the skilled forgeries.

4.7.3 USER MODEL SIZE

In the DTW algorithm proposed, the number of elements within the feature vector does not impact the user model size. This issue has been discussed in section 4.5 where it was shown that the 3 feature vectors imply the storage of the time series of the x and y axes, pressure and time, whereas the inclinations time series (elevation and azimuth) are not required.

The user model size for the DTW algorithm is a function of the number of equi-spaced points used in the preprocessing steps. The bigger the number of equi-spaced points used, the bigger the user model. This 4 time series (x and y axes, pressure and time) is stored using integer values within 2 bytes.

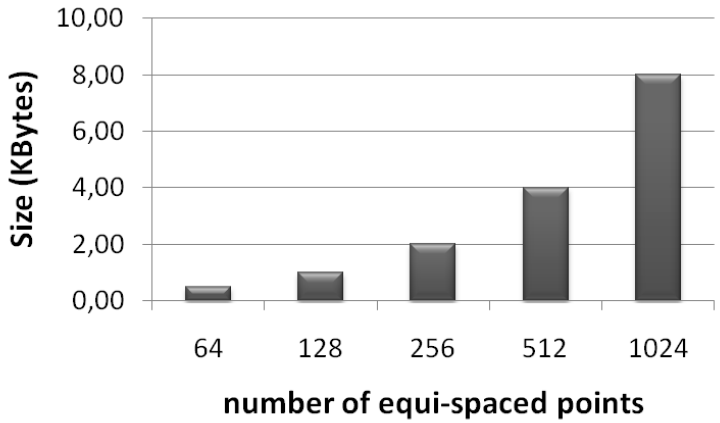


Figure 4-19 User's model size for DTW algorithm and different number of equi-spaced points

The user model sizes for the number of equi-spaced points was analyzed in the previous section 4.7.1, where in Figure 4-19 the points from 64 to 1024 are represented. In the case of 64 equi-spaced points the resultant EER was too high and not considered to be an option in section 4.7.1. However, 128 equi-spaced points demonstrates sufficient EERs, and considered a good option for systems with limited storage requirements. In Figure 4-19 it can be seen that the user model size for this option is only 1 Kbyte. The user model size increases from this 1 Kbyte to 8 Kbytes for 1024 equi-spaced points. The size values are acceptable for most systems, and, as it will be shown in next section, the computational load for the DTW algorithm with 1024 equi-spaced points is much higher than that of 128 equi-spaced points.

One advantage of the proposed DTW algorithm is that the number of stored points in the time series is known and fixed by designers, avoiding storage issues in cases of exceptionally long signatures. The MCyT database has an average of 350 points for each signature and a maximum of 1163 points. This number of points can be increased with newer signature input devices that have higher sampling rates than 100 Hz.

4.7.4 COMPUTATIONAL LOAD

The computational load of the DTW algorithm proposed is greatly influenced by the number of equi-spaced points used in the preprocessing step. The dynamic programming algorithm has to fill two matrices with distance as described in [127] and the calculation of these two matrices are very computational demanding processes. The derived signal calculations, as well as the feature calculation, have a very low impact on the overall comparison process.

The time taken for the comparisons is presented in milliseconds. This time has been calculated using the DTW implemented in Matlab® 2008. The preprocessing step and feature calculation were implemented in Matlab Code while the DTW algorithm itself has been optimized using Mex-C functions[5]. The algorithm was executed on a computer with the following characteristics: Intel® Core® 2 Duo E670@2.66GHz, 3 GB of RAM and Windows® 7 32 bits. The time represents the average time after calculating 1000 comparison scores.

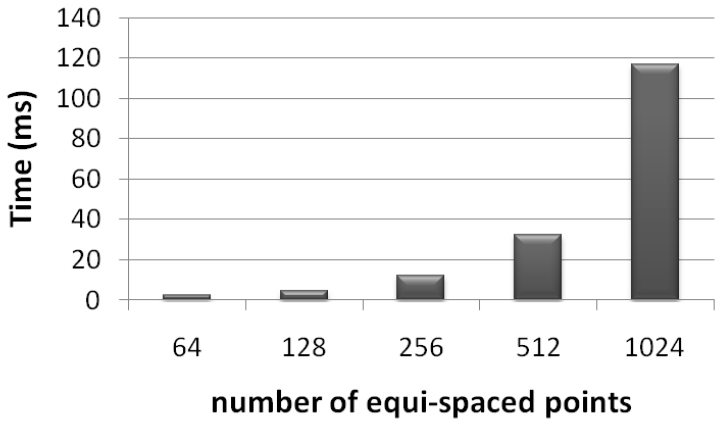


Figure 4-20 Comparison time for DTW algorithm and different number of equi-spaced points

From Figure 4-20 the impact the number of equi-spaced points has on the computational load of the DTW algorithm can clearly be seen. The comparison time for 256 equi-spaced point signals is 12.4 ms while for 1024 equi-spaced points it increases to 116.8

[5] Due to the fact of having used Mex-C functions in order to perform the DTW algorithm, these values are not comparable with the comparison times obtained with the GMM algorithm.

ms, approximately 10 times more. It was shown in the previous section 4.7.1 that larger numbers of equi-spaced points do not imply any improvement in the algorithm's performance in term of the EERs.

To reiterate, the fact of having fixed numbers of equi-spaced points avoids comparison time issues for longer signatures and prevents high comparison times when using new input devices with higher sampling frequencies. Such systems will be introduced in the near future.

4.8 CONCLUSIONS

In this chapter 4 different feature selection techniques have been used to select reduced feature vectors for two different algorithms. The aim of this analysis was to obtain reduced subsets to minimize the size of the user model to be stored and the algorithms computational requirements.

The different feature selection techniques used are based on the Fisher Ratio, Principal Component Analysis and Hellinger distance. The Principal Component Analysis was used as a feature selection technique for signature verification systems, as it demonstrated better performance when compared to the more extensively used Fisher Ratio technique. Furthermore, the combination of both techniques, using the Fisher Ratio as a first filter to remove the features with low discriminative power followed by the PCA to sort the remaining set of features, obtains the best performance when compared to each technique individually.

The Hellinger Distance is a novel feature selection technique which, to date, has not been used in signature applications. In this chapter it has been demonstrated that this method improves on the results obtained from the combination of the Fisher Ratio and the PCA for algorithm features with low levels of overlap between the genuine and forgery distributions, such as the DTW algorithm proposed.

The reduced feature vector subsets have been selected for both algorithms, GMM and DTW, reducing the user model size and the computational load, while at the same time, maintaining an equal error rate which is in agreement with the state of the art for such algorithms. In particular, considering the GMM algorithm, a feature vector composed of up to 28 features has been selected, and has demonstrated a state-of-the art equal error rate level. Furthermore, an extremely small feature vector composed only of 13 subsets has also been proposed, as it has obtained reasonable error rates. For the DTW algorithm, a reduction in the number of derived signals used on the verification process from 25 to 6 has been carried out and has obtained the same error rate levels presented and published in [80].

It is worthwhile highlighting that the DTW algorithm improves on the performance of the GMMs in both random and skilled forgery error rates. The GMM presents better characteristics when considering the user model size and comparison time.

Moreover, the influence of several parameters pertaining to these two algorithms has been analysed. For the GMM-based algorithm proposed, it has been shown how the number of mixed Gaussian distributions can be as little as 4. The number of signatures taken during the enrolment process has also been analysed. As expected, the more signatures taken during the enrolment process, the better the equal error rate obtained.

For the DTW-based algorithm proposed, the number of sample points of the signature acquired is extremely critical from the point of view of the computational loads and user

model size. It has been demonstrated how this number of sample points can be reduced by linear interpolation to 512 points while maintaining the same verification error rate levels obtained from 1024 sample points. This demonstrates that 512 points are sufficient to maintain all the relevant information for the DTW algorithm. Also, the performance of 256 sample points has been demonstrated to be adequate, although there is a reduction in the performance (from an EER of 3.4% to 3.6%). For systems in which storage and computational requirements are an important issue, 128 sample points may still be considered as a possible alternative. Regarding the number of signatures used for the enrolment process, the DTW-based algorithm is more robust than the proposed GMM-based algorithm, where a reasonable equal error rate with only 3 signature sample sets for enrolment is observed.

Chapter 5 VIABILITY ANALYSIS OF SIGNATURE STANDARD DATA FORMATS

5.1 INTRODUCTION

In section 3.6 the three different international standard signature data formats have been explained, where two of them claim to be a compressed data format for signatures, i.e. part 7 compact data format and part 11.

There are no previous studies on the affect that these specific data formats have on the Biometrics Data Interchange Record (BDIR) sizes, also, there is no information regarding the compression ratio achieved for these formats which claim to be compact formats. This chapter will attempt to discover the answer to these questions.

Following this analysis on the size and the compression ratio presented for the 19794-7.2WD2 Compact Format and the 19794-11CD2, section 5.3 will examine whether these compression formats have any impact on the performance of the algorithm.

5.2 BDIR SIZE ANALYSIS

The previously mentioned lack of information has motivated part of the research work presented in this Thesis. All three signature data formats that exist as part of the preliminary version of the signature standards [115] [119] have been implemented and used to store signature samples from three different public databases: MCyT [21], SVC2004 [22] and MyIDea [101]. Only genuine signatures have been used, as forgery signatures do not represent real examples.

These databases contain, in total, 4614 original samples from 210 users where the users are from different countries:

- MCyT: Spanish users,
- MyIDea: French and English users,
- SVC2004: Chinese users signing in English or in Chinese character sets.

The use of these three databases also allows exploration of the differences that may arise regarding the compression ratios of signatures from different countries and also the effect of using different characters.

The BDIR average sizes, compression ratio, number of samples, total signing time, number of pen-strokes and pressure-strokes have been calculated for the three databases. These 210 users have been split into five different subsets which depend on the origin and character set used for each user. This has been done to analyse the effects of these parameters on the compression ratios. The five different subsets are defined in the following table:

Table 5-1 Subsets used for BDIR Size Study

Subset ID	Country	Character Set Used	N# of Users	N# of Signatures per User
MCyT	Spain	Latin	100	25
MyIDea_FR	France	Latin	46	18
MyIDea_En	England	Latin	27	18
SVC2004_OCC	China	Latin	24	20
SVC2004_ORI	China	Chinese	16	20

In the following sections, details on how the three standard data formats have been implemented will be given. This is then followed by results obtained from the five datasets.

5.2.1 IMPLEMENTATION DETAILS

5.2.1.1 19794-7.2WD2 FULL FORMAT

All part 7 Full Format instances share the same BDB General Header (7 bytes) and BDB Representation Header, which include all the mandatory fields (19 bytes) plus the channel descriptions (50 bytes) which details the channels that are included (time, x and y position, switch and pressure), their scaling values and maximum and minimum values.

Although all datasets include tilt information (azimuth and elevation), these channels have not been included due to the fact that they cannot be stored within the Part 11 data format. Therefore, in order to perform a comparative analysis of the compression ratios achieved using the 19794-7.2WD2 Compact Format and 19794-11CD2, it has been decided to use only the common channels between all the standard data formats.

5.2.1.2 19794-7.2WD2 COMPACT FORMAT

The Compact Format instances do not have any header, they contain the number of samples (3 bytes) and their values for channels: x and y position, switch and pressure.

The time channel has not been included in this format, this is because the values are to be stored in only 1 byte and there is a lack of resolution in 1 byte (values range from 0 to 255) for signatures made up by more than 255 sample points.

For signature samples from the MCyT and MyIDea datasets the time channel exclusion does not imply any loss in information, this is because the sample points are captured at an equidistant time difference, 10 milliseconds, and both pen-up and pen-down movements are recorded.

For samples coming from the SVC2004 datasets the time channel exclusion will lead to a loss in information. Although the sample points are also captured at an equidistant time difference of 10 milliseconds, only pen-down movements are recorded. This fact implies that the time difference between a pen-up event and the next pen-down event will be lost.

To store the X and Y position channels in only 1 byte, the values from the different datasets have been converted from 2 bytes (ranging from -32768 to 32767) to 1 byte (ranging from -128 to 127). This conversion has been made using a linear interpolation between the maximum and minimum values for each single record to the range of -128 and 127, where the result is rounded off to its closest integer.

A further implementation detail which does not conform with the data format defined in 19794-7.2WD2 has been carried out for the signature samples of the SVC2004 datasets. This implementation deals with the switch channel. According to 19794 part 7, the switch channel is defined as “for recording whether the pen tip touches the writing plane or not. The data values shall be 0 in case of non-touching and 1 in case of touching”. If this definition were followed, and taking into account that no information regarding time will be stored for the 19794-7.2WD2 Compact Data Format, it would not be a trivial task to ascertain when a new

stroke (indicated by the presence of a pen-down event) has started. This information is provided within the SVC2004 datasets in the switch channel with the presence of a 0 for a pen-down event sample point, and 1 for the rest of the pen-down sample points. This concept has been used in the SVC2004 Compact Format BDIRs as opposed to that provided in the part 7, 19794-7.2WD2.

This issue has been discussed by SC37 WG3 signature experts as a result of comments made by the Spanish National Body, sent by the author of this Thesis. As a result of this discussion, a new definition for the switch channel was agreed upon following the implementation explained above.

5.2.1.3 19794-11CD2

All instances stored using the data format defined in 19794-11CD2 share the same Biometric Record Header (17 bytes). Within the BDB Body, channel scaling values and sampling resolution have been filled, and the absence of extra data has been indicated in the preamble field.

In order to obtain the data format defined in 19794-11CD2 from the data stored in the 19794-7.2WD2 Full Format, several solutions have been taken into consideration.

These are:

1. Velocity and Acceleration have been calculated from the X and Y channels as:

$$v_{Xi} = \frac{x_{i+1} - x_i}{t_{i+1} - t_i} \tag{36}$$

with $v_{X0} = 0$

$$a_{Xi} = \frac{v_{i+1} - v_i}{t_{i+1} - t_i} \tag{37}$$

with $a_{X0,1} = 0$

(Same for the y channel)

2. Pen-strokes and pressure-strokes between turning points are forced to be longer than 2 points. This simplification avoids strokes being recorded that are a result of noise or small vibrations during the signing process. Here the dotted lines indicate pen-up movements, square markers indicate pen-down events, diamond markers indicate pen-up events and round markers indicate turning points.

In order to calculate the singular points and the pressure turning points, the definition provided by the standard has been followed: “a sample point where either x, y or both axes values change from increasing to decreasing”.

5.2.2 RESULTS OBTAINED

As previously mentioned, for all of the different data formats analysed only genuine signatures from all datasets have been used. These have been used to determine the average BDB size and compression ratio for the 19794-7.2WD2 Compact Format and 19794-11CD2 data format, along with other key characteristics such as the mean number of points and pen and pressure strokes. In the following graphs the different data formats will be named “FF” for Full Format defined in part 7, “CF” for compact format defined in part 7, and “P11” for part 11.

In Figure 5-1 the Full Format, Compact Format and part 11 BDIR sizes are shown. To begin with, it is worth highlighting the differences observed between the five datasets regarding the size of the BDIR. It may be seen that the MCyT dataset has the greatest size of all the datasets, especially when compared with the MyIDea datasets. The MCyT dataset contains Spanish users, whereas MyIDea datasets contain French and English users, where all of them use the Latin character set. Spanish users are known to have more pictorial strokes than English and French users, where this explains why the size of the Spanish BDIR signatures is greater than the MyIDea users. This can also be observed in Figure 5-3 to Figure 5-5, where the MCyT dataset has the greatest average number of points, total time and strokes (except when compared to the SVC2004 oriental users dataset). These facts indicate that, in general, Spanish signatures are of a more complex nature than French and English ones.

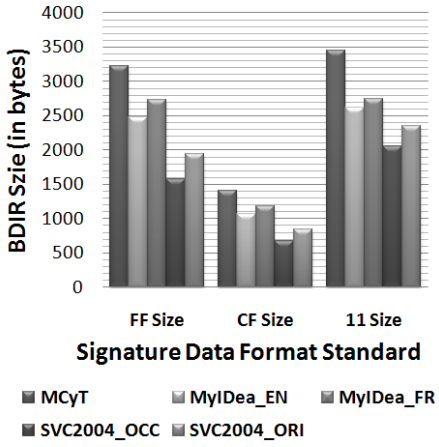


Figure 5-1 BDIR Sizes for different Signatures Data format Standards

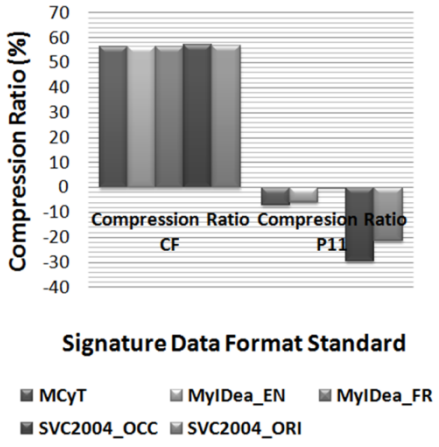


Figure 5-2 Compression Ratios for different Signatures Data format Standards

The data extracted from the SVC2004 dataset which contains Chinese users, signing in either Latin or Chinese character sets, shows that the average number of points in the signature is clearly lower than the other datasets. This is because the SVC2004 database does not store the pen-up movements, as opposed to the MCyT and MyIDea databases. The

differences that arise between occidental users from the MCyT and MyIDea datasets and SVC2004 databases, in both number of sample points and signature total time, may also be a result of the fact that the SVC2004 database has no real user signatures. Users had to create new signatures, where they may have chosen easy and straightforward ones in order to get used to doing them quickly. This fact can make the resulting signatures less stable than real ones, and as a result may lead to non-satisfactory algorithm verification performances.

In Figure 5-2 the compression ratios achieved by the 19794-7.2WD2 compact format and 19794-11CD2 compared with the 19794-7.2WD2 Full Format are shown. As it was expected, the Compact Format achieves a compression ratio close to 56%. This value comes from the conversion of the channel point values from 2 bytes to 1 byte. This conversion affects position channels X and Y, and also pressure (the channel time, as described in section 5.2.1.2, is not stored in the Compact Format). The channel switch maintains the same size, as it is defined as a 1 byte length in Full Format. Also, further compression is obtained as a result of the header absence in the Compact Format.

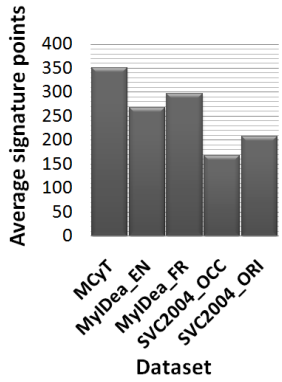


Figure 5-3 Average Signature Points for Different Datasets

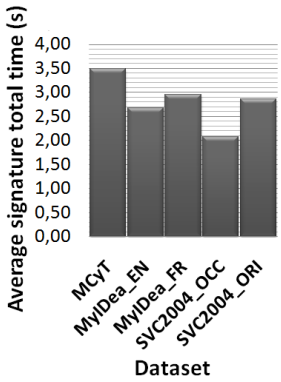


Figure 5-4 Average Signature Total Time for Different Datasets

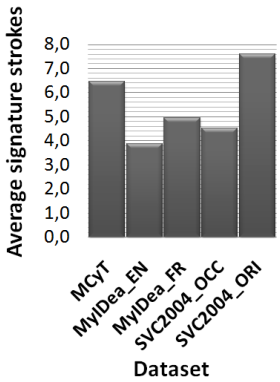


Figure 5-5 Average Signature Strokes for Different Datasets

The Part 11 demonstrates poor compression performance, see Figure 5-2. 19794-11CD2 does not compress the part 7 Full Format as expected, and has achieved a greater average BDIR size for all the datasets tested. The reasons behind this revolve around the fact that there is too much information recorded for each singular point (6 velocities, 6 accelerations, etc.). Some of the information is also duplicated for consecutive strokes, i.e. each stroke records initial and end point information. Since consecutive strokes share information on their starting and ending points, this duplicated information can be removed. The data format is also affected by channel noise, small amounts of trembling at the X, Y and Pressure channels, which may arise from the capturing device and/or from user signing noise. This noise leads to spurious singular points, thus increasing the number of pressure and pen strokes, hence increasing the size of the BDIR defined in Part 11. This noise and its effect on the size of the BDIR can be avoided using a low pass filter before the velocity is calculated and by defining turning points as a zero-crossing velocity event. The use of this low pass filter

would mean a lower number of singular points, and therefore, a lower BDB size. Including a filter can be avoided by using a more complex velocity calculation formula that is not affected by small variations between sample points. However, this would lead to changes in the description of turning points for a velocity zero-crossing based definition.

The results from the pen and pressure turning points are presented in Figure 5-6. The difference between the MCyT dataset and the rest of datasets is especially significant. Again, these results may indicate that the MCyT signatures are more complex in terms of x and y axis variations as well as pressure variations. The large difference between occidental users from the MCyT-MyIDea and SVC2004 databases may be caused by the fact that the SVC2004 database does not contain real signatures.

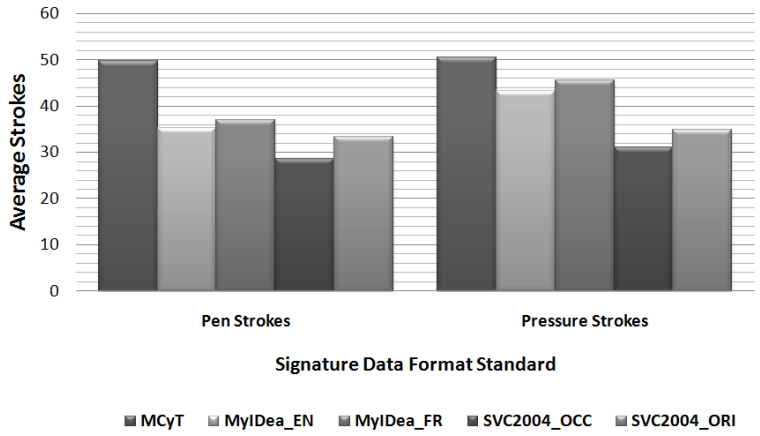


Figure 5-6 Average Pen and Pressure Strokes for different Signatures Data Format Standards

5.3 ALGORITHM PERFORMANCE ACHIEVED WITH CURRENT DATA FORMATS

The 19794-7.2WD2 Full Format describes how the raw data coming from the signature input devices is stored, where no preprocessing has been carried out, thus implying that no information is lost. However, the other two signature data formats claim to be compact formats, where the raw information is preprocessed to achieve a certain level of compression, but this preprocessing leads to a loss in information.

In the 19794-7.2WD2 Compact Format a transformation from a 2 byte to 1 byte resolution has been completed in order to achieve a compression ratio of close to 56%. This reduced resolution can imply inferior algorithm performances.

In the 19794-11CD2 the compression is obtained by means of signal segmentation through the definition of singular points. In the previous section 5.2 it has been shown that the way this data format is defined leads to no compression ratio, where for some cases the size of a sample stored following this data format is greater than samples from the 19794-7.2WD2 Full Format. Also, the segmentation that is required implies information loss and as a result only information on the singular points are saved and the information during pen-up movements is completely missed.

This section analyses whether the information lost on both ISO/IEC compact formats (19794-7.2WD2 Compact Format and 19794-11CD2) has an impact on the performance of signature verification algorithms.

In order to use the standard 19794-11CD2 Compact Format on the algorithms used to perform the test, no additional preprocessing has been required.

Unlike the previous case, the 19794-11CD2 requires the addition of a further step to obtain the complete temporal signals used as inputs for both algorithms. From the information stored in each singular point (x and y position, pressure, time and initial vector direction), the corresponding four temporal signals have been interpolated. This interpolation has been carried out using MatLab's Interpolation Toolbox [139-140]. Cubic spline interpolation has been used on X and Y channels whereas Piecewise Cubic Hermite interpolation has been used for the Pressure Channel.

In order to evaluate the impact on the performance of the compact formats being tested, the signature verification algorithms introduced in section 4.2 have been used. The first of them is a Gaussian Mixture Model, which is based on the features extracted from the signature signals. The second is Dynamic Time Warping (DTW) which is based on temporal alignment between the biometric sample and the biometric reference template.

Using the $MCyT_T$ dataset and following the same methodology explained in chapter 4, five random signatures have been used to train the user models, therefore 20 genuine and

25 forgery signatures for each user were used to obtain the skilled error rates. This test has been repeated 10 times in order to reduce the impact of the selected signature for training.

Figure 5-7 and Figure 5-8 show the results obtained for both algorithms, GMM and DTW, respectively.

In Figure 5-7 it may be observed, that when using the GMM signature verification algorithm, the 19794-7.2WD2 Full Format achieves better results than the other two data formats. In terms of EERs, the 19794-7.2WD2 Full format achieves 3.9% while the Compact Format increases the EER to 4.3%, the 19794-11CD2 achieves 5.1%. In relative terms, the Compact Format increases the EER by a factor of 10%, while the part 11 increases it by a factor of 30%.

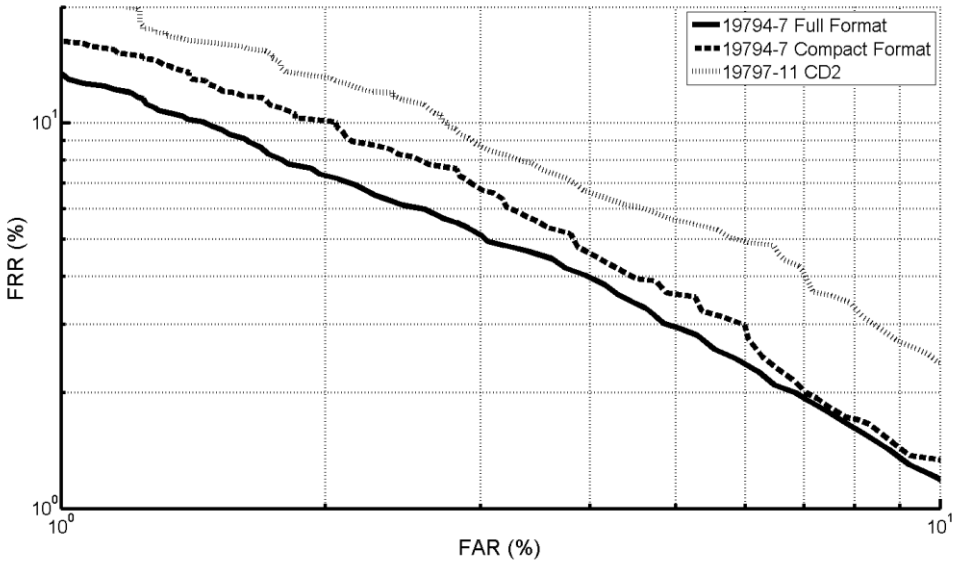


Figure 5-7 Error Rates for GMM and Different Data Formats

Results from the DTW are presented in Figure 5-8 where it can be seen that this particular algorithm is more robust against the 2 to 1 byte transformation introduced by the 19794-7.2WD2 Compact Format, achieving in both 19794-7.2WD2 formats an EER of close to 4.1%. However, the 19794-11CD2 data format again demonstrates inferior algorithm performance, with an EER of 5.4%, again close to 30% less effective than the EER obtained using the 19794-7.2WD2 data formats.

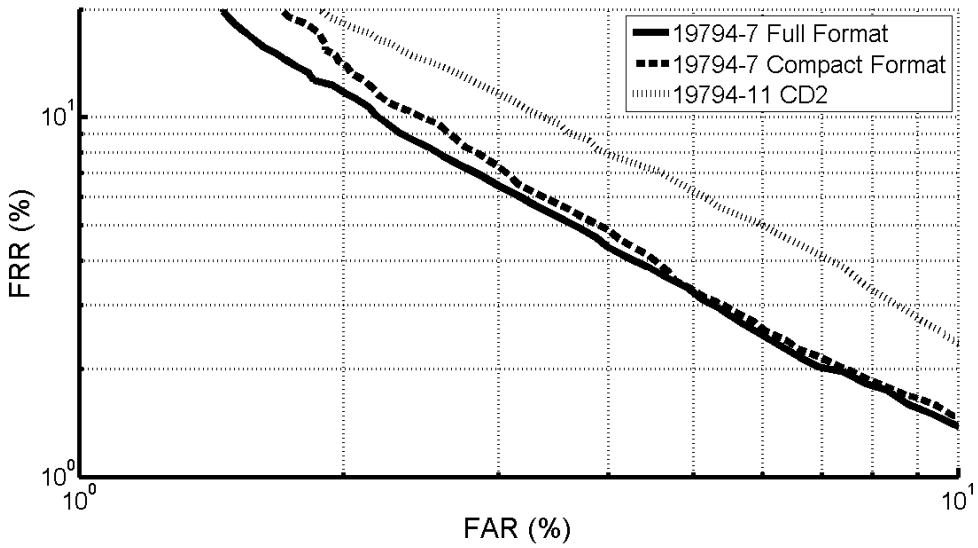


Figure 5-8 Error Rates for DTW and Different Data Formats

5.4 CONCLUSIONS

Due to the growth in Signature Automatic Verification Systems applications, and the increase in the number of different types of capture devices, it has become necessary to have standardized data formats to guarantee interoperability.

In this chapter, three signature data formats defined by ISO/IEC have been analysed.

As a result of this analysis, it has been shown that the 19794-7.2WD2 Compact Format achieves a 56% compression ratio, whereas the 19794-11CD2 does not imply any compression, increasing the BDB data size. This greater size is due to the amount of information stored within each pen and pressure stroke. The use of a low pass filter or more complex velocity calculations are possible options to improve the compression ratio of part 11.

An analysis of different kinds of users (Spanish, French, English and Chinese) has also been carried out. Nationality is seen to affect several signature characteristics such as: number of strokes, total writing time and complexity. These characteristics have an impact on the size of the BDB as well as on the complexity of the signatures. The occidental signatures collected in the SVC2004 database are the most straightforward. This can be explained by considering the fact that this database is composed of false signatures that have been created only for the purpose of collecting information for the database. Signatures collected in the MCyT have been shown to be more complex, in terms of the quantity of sample points, turning points and pen-up and pen-down events.

Due to the data information lost within the compact formats, it has been proved that the reduction from 2 bytes to 1 byte specified in the 19794-7.2WD2 Compact Format may imply an impact on the performance of the signature verification algorithm.

Using the 19794-7.2WD2 Compact Format with the GMM algorithm obtains an EER 10% inferior to that when compared to the raw data storage in the 19794-7.2WD2 Full Format. The DTW algorithm is observed to be a more robust algorithm regarding the range of values used for storing the signature data and is not affected by the reduction from 2 bytes to 1 byte.

In the case of the 19794-11CD2 data format, it has been shown that by using the information stored within the strokes (x and y position, pressure and time), it is possible to recreate the original signals captured by the input devices. However, the performance of the signature algorithms tested is observed to be lower, where the EER is close to 30% worse for both the GMM and DTW algorithms. The inferior performance is a result of the information lost when transforming the signature data from the 19794-7.2WD2 Full Format to the 19794-11CD2, and is also due to the recreation of the 19794-7.2WD2 Full Format signature data from the 19794-11CD2 using interpolation.

In the following chapters 6 and 7, the problems detected from this analysis have been faced, where solutions are proposed for each of them. Chapter 6 will solve the lower performance obtained using the 19794-7.2WD2 Compact Format, by proposing a new signature data format based on lossless compression methods, thus attempting to obtain satisfactory compression ratios which are lossless, or near lossless.

The drawbacks related to the 19794-11CD2 are covered in chapter 7, where attempts to solve the lack of compression and minimizing the loss of information are presented. In order to solve these problems, a new data format structure is proposed which avoids the storage of duplicated data, reducing the data stored and by developing the definition of the singular point to reduce the information lost.

Chapter 6 INTEROPERABILITY OF SIGNATURE BIOMETRICS AT SIGNAL LEVEL

6.1 INTRODUCTION

In the previous Chapter 5, a *Viability Analysis Of Signature Standard Data Formats*, 19794-7.2WD2 [115] was implemented and analysed using three different public datasets [21-22, 101]. From the results obtained, it was shown that the Compact Format achieves a compression rate of 56% when compared with to the Full Format, this is accomplished by transforming the sample point channel values from 2 bytes to 1 byte.

This compression implies some loss of information. Values are stored in only 1 byte, therefore there is a loss in precision of 16 bits to 8 bits. Furthermore, the channel time cannot be stored as there is not sufficient precision in just 1 byte. Alternatively, a DT (time difference) channel or uniform sampling should be used.

It was also shown how this loss in information may affect the performance of several of the signature verification algorithms, i.e. the GMM algorithm increases its equal error rate from 3.9% to 4.3%, whereas, the DTW was shown to be more robust against bit depth.

This chapter will introduce an alternative data format that is based on a lossless compression data algorithm where the objectives laid out have been to achieve reasonable and sufficient compression ratio, without any loss in information, and therefore, not affecting the performance of the signature verification algorithms. It explores different strategies of re-organizing the sample point values from on-line signature data to obtain

higher performance of the compression data algorithm and consequently a lower biometric data record size.

A new near-lossless strategy of re-organizing the sample point values is also explored by introducing an error difference control level between consecutive sample mechanisms.

Different data compression algorithms are tested to ascertain their suitability for on-line signature data characteristics, where these have been verified using the same three databases discussed in chapter 5. These databases will also allow identification of any demographic differences between them for the compression data formats proposed.

6.2 COMPRESSED DATA FORMAT

The new compressed data format, presented in this Thesis, is based on organizing sample point values to improve the functioning of data compression algorithms. In order to find the best way to re-organize the sample point values, different versions have been tested using the aforementioned mentioned databases.

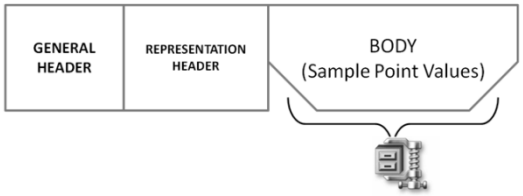


Figure 6-1 BDB Body Data Compression in Compressed Data Format

The first Version (V1) (based on the Signature/Sign Time Series Full Format 2nd Generation, 19794-7.2WD2) has been used as a reference to provide baseline results for other methods, allowing the performance of the data compression algorithms to be assessed. Version 2 (V2) and Version 3 (V3) have been implemented to explore two further methods of reorganizing sample point values to improve the performance of the lossless compression algorithm, finally, Versions 4 to 6 (V4-6) use a novel near-lossless strategy to arrange sample point values which entail a controlled loss of information. These last three have been used to test and identify the level of information loss.

Following is a brief description of each version of the compressed data format tested.

Version 1 (V1)

This version compresses the sequence of sample points as defined in 19794-7.2WD2 Full Format BDB Representation Body. It is also used to provide baseline results for other methods as a means of assessing the performance of the data compression algorithms. Compression algorithms are applied to the sequence of sample points as defined in the standard (19794-7.2WD2 Full Format), i.e. as a sequence of points, where each point contains the values from all the channels included.

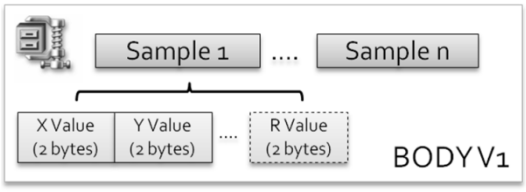


Figure 6-2 Sample Point Values order for Compression Data Format Version 1

Version 2 (V2)

Instead of storing the sequence of samples as defined in the 19794-7.2WD2 Full Format, it has been considered that a further option may be to store each channel separately, followed by linking all channels together before finally compressing the resulting data structure. This has been implemented to improve the compression algorithms performance as a result of similarities in the consecutive samples from the same channel.

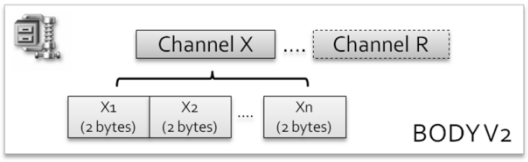


Figure 6-3 Sample Point Values order for Compression Data Format Version 2

Version 3 (V3)

In this version it is proposed that the values of each channel are not stored separately but rather the difference between consecutive samples of the same channel. In this way, the first value of a channel is stored as its actual value, and the following values are the difference with the previous one. These difference values are stored as a 2 byte signed integer instead of a 2 byte unsigned integer. After calculating the differences, the results from each channel are linked together and then compressed. This strategy is envisaged to improve the performance of the compression algorithms compared with Version 2, as lower values are required to code each sample, achieving more repetitive values, hence its suitability for data compression algorithms.

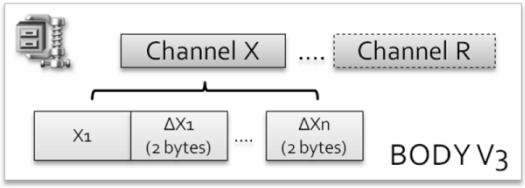


Figure 6-4 Sample Point Values order for Compression Data Format Version 3

Version 4 (V4)

Version 4 is a novel technique proposed to store sample point values. As in the case of Version 3, it does not store the sample point values but rather the difference between consecutive sample point channel values, however in this case, just 1 byte is used as a signed integer. Again, storing the difference values within just 1 byte will lead to more repetitive values, which improves the performance of the compression data algorithm. Furthermore, due to storing the differences within only 1 byte, this leads to lower body sizes, which after compression, will achieve even smaller BDB sizes for this data format.

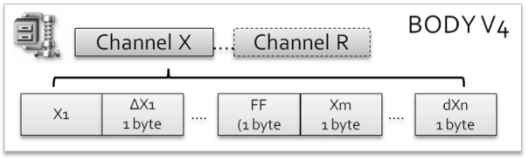


Figure 6-5 Sample Point Values order for Compression Data Format Version 4

In order to store difference values as a signed integer within just 1 byte, a mechanism to control whether the absolute difference value is greater than 127 (limit for a 1 byte signed integer) has to be implemented. The mechanism used is described as follows:

1. Every channel will start with its initial value V_1 (stored in 2 bytes).
2. The initial value, V_1, is followed by a sequence of differences between consecutive samples, $\Delta V_2 ... \Delta V_n$ stored in 1 byte, where:

$$\Delta V_i = V_{i+1} - V_i \tag{38}$$

3. When the difference between consecutive samples is greater than 126, a control character FF_{HEX} is stored, followed by the original channel value for that sample (stored in 2 bytes).
4. The control character FF_{HEX} and the original channel value, are again followed by a sequence of differences between consecutive samples, $\Delta X_2 ... \Delta X_n$ that are stored in 1 byte.

The following block diagram, Figure 6-6, explains the methodology used to obtain the values that are stored within just 1 byte:

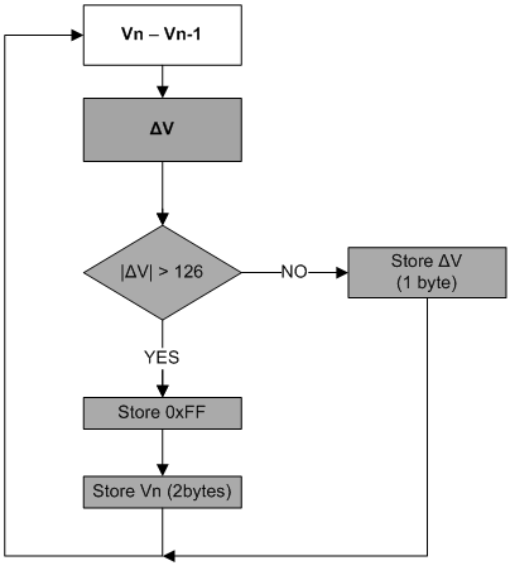


Figure 6-6 Block Diagram for Compressed Data Format Version 4, without any error.

Version 5 (V5)

Version 5 brings the previously described Version 4 a step further. Here the differences between consecutive sample channel values are also stored within 1 byte as an unsigned integer. But this version allows a specified error in the difference stored, called *Error_Difference_Level,* which will allow 1 byte absolute differences bigger than 126 to be stored.

The following block diagram, Figure 6-7, explains the mechanism used to obtain the values stored within just 1 byte, where a specified error level is allowed:

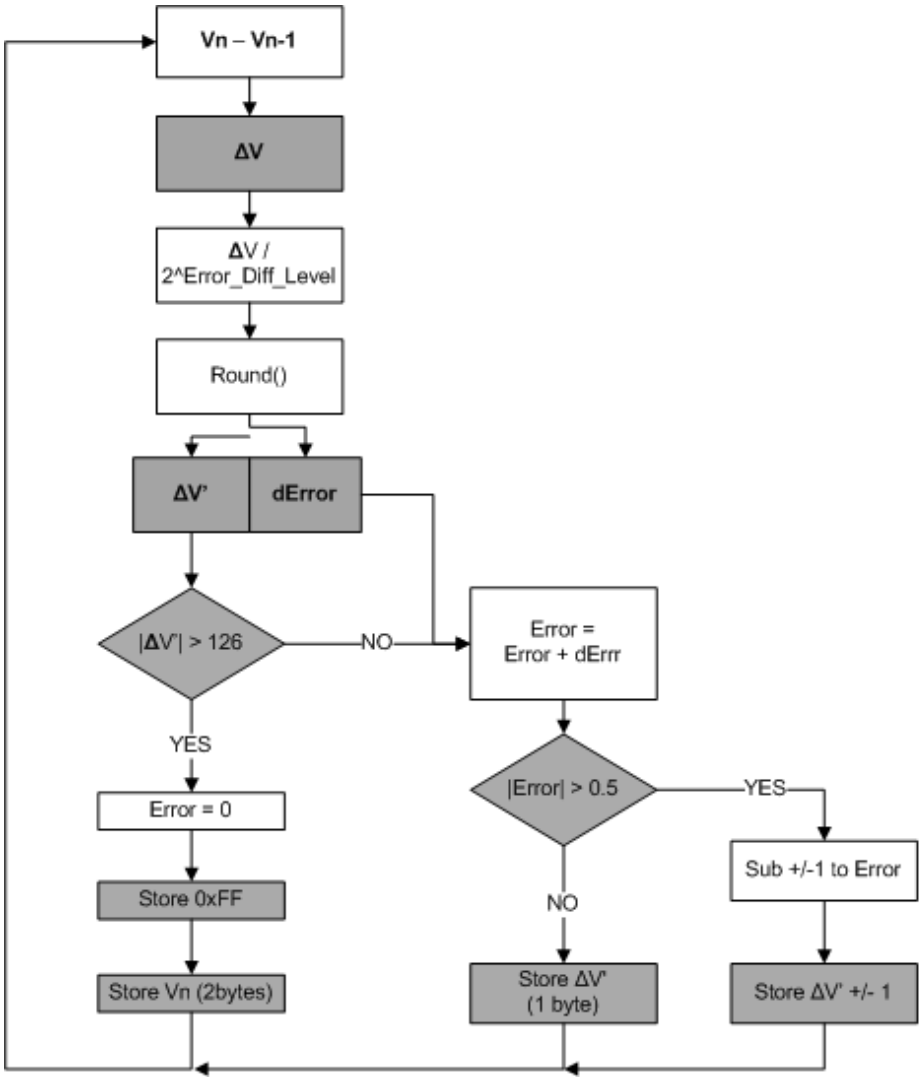


Figure 6-7 Block Diagram for Compressed Data Format Version 4, with error.

The methodology used is described as follows:

1. Every channel starts with an initial value V_1 (stored in 2 bytes).
2. The initial value, V_1, is followed by a sequence of differences between consecutive samples, $\Delta V'_2 ... \Delta V'_n$ stored in 1 byte, where:

$$\Delta V'_i = \text{round}\left(\frac{V_{i+1} - V_i}{2^{\text{Error Difference Level}}}\right) \quad (39)$$

3. When the difference between, $\Delta V'_i$, consecutive samples is greater than 126, a control character FF_{HEX} is stored, followed by the original channel value for that sample (stored in 2 bytes), updating the accumulative error to 0.

    3.1. The control character FF_{HEX} and the original channel value, are again followed by a sequence of differences between consecutive samples, $\Delta V'_i$, stored in 1 byte, according to the definition given in (2) .

4. When the difference between consecutive samples, $\Delta V'_i$, is lower than 126, an accumulative error is updated.
    4.1. If the accumulative error is lower than 0.5, then the difference is stored within 1 byte.
    4.2. If the accumulative error is bigger than 0.5, then the accumulative error is corrected with +/- 1 as well as the difference between consecutive values, ΔV_i, and then the corrected difference, $\Delta V'_i$, is stored within 1 byte. This method is such that the error is within the limit defined by:

$$\text{Maximum Error} = 2^{\text{Error\_Differen\_Level} - 1} \quad (40)$$

In this version, V5, an *Error_Difference_Level* of 1 has been tested, a value of 1 for channels X, Y, T, S and P, which will entail a maximum difference between the original data and the compressed data of 1.

Version 6 (V6)

Implemented as described in Version 5, but an *Error_Difference_Level* value of 2 will be used for channels X, Y, T, S and P, which involves a maximum difference between the original data and the compressed data of 2. Here it will be seen how allowing a bigger error effects the compression ratios achieved.

The same block diagram for Version 5 (see Figure 6-7) is also valid for this version.

6.3 COMPRESSION ALGORITHMS TESTED

In order to determine the compression algorithm which best suits the characteristics of the signature data, seven different lossless compression algorithms have been tested. The majority of these are included in the 7-Zip Command Line Version application [141]. Also GZip [142], LZW [143] and BZip2 [144] compression algorithms have been tested.

Table 6-1 summarizes the compression algorithms used and their source:

Table 6-1 Compression Algorithms Tested

Identifier	Compression Algorithm
1	Zip (7 – zip) [141]
2	LZMA (7-zip) [141]
3	PPMd (7-zip) [141]
4	Deflate (7-zip) [141]
5	GZip [142]
6	LZW [143]
7	Bzip2 [144]

6.3.1 7 ZIP

The algorithm 7-Zip [141] is a free open source piece of software which provides a high compression ratio. The supported formats for packing and unpacking are: 7z, ZIP, GZIP, BZIP2 and TAR. The version used for the experiments presented in this Thesis has been 4.65 for Windows released on 03-02-2009. Within the 7z format, several methods have been tested: Zip, LZMA (Improved and optimized version of the LZ77 algorithm), PPMd (Dmitry Shkarin's PPMdH with small changes) and Deflate (Standard LZ77-based algorithm). Further details can be found in [141].

6.3.2 GZIP

GZip [142] is a free cross-platform software application for file compression. It claims to have, as a main advantage over other compression software, a much improved level of compression and freedom from patented algorithms. It has been adopted by the GNU project and is now commonly used on the Internet. The Gzip file format was standardized as RFC 1952 [145]. The Version 1.2.4 for the Windows platform has been used. Further details can be found in [142].

6.3.3 BZIP2

BZip2 [144] is a free and open source lossless compression algorithm developed by Julian Seward. It also claims to be patent free. The main advantages of this algorithm are its good compression ratio and fast compression and decompression times. The version used in this work has been 1.0.5, which was released on the 17th of March 2008. Further details can be found in [144].

6.3.4 LZW06

The LZW06 lossless compression algorithm is an implementation of the Lempel-Ziv-Welch encoding /decoding algorithm by Michael Dipperstein [143]. Version 0.6 has been used which was released on the 21st of December 2009. Further details can be found in [143].

6.4 BDIR SIZE ANALYSIS

In this section, results will be presented for the data storage of signatures included in the MCyT, MyIDea and SVC2004 databases using the different versions for the new compression data format described in section 6.2 and using the compression data algorithms detailed in Section 6.3. These results will show the compression ratio obtained for each of the different versions along with the performance of the different compression algorithms tested. Also, results obtained will be compared to the compression ratio achieved using the 19794-7.2WD2 Compact Format.

6.4.1 RESULTS OBTAINED

Genuine signatures included in the datasets described before have been stored according to the 19794-7.2WD2 specification in both Full Format and Compact Format and following the implementation described in section 5.2.1.1 and 5.2.1.2.

For each Compression Format Version proposed the same BDB General Header and BDB Representation header have been used, however, the BDB Body has been formed by organizing the data as explained in Section 6.2 and then compressed with the different compression algorithms detailed in Section 6.3.

In Figure 6-8 the average data size in Kilobytes for the different Data Formats and the different datasets are presented. The average has been obtained for each of the compression algorithms used.

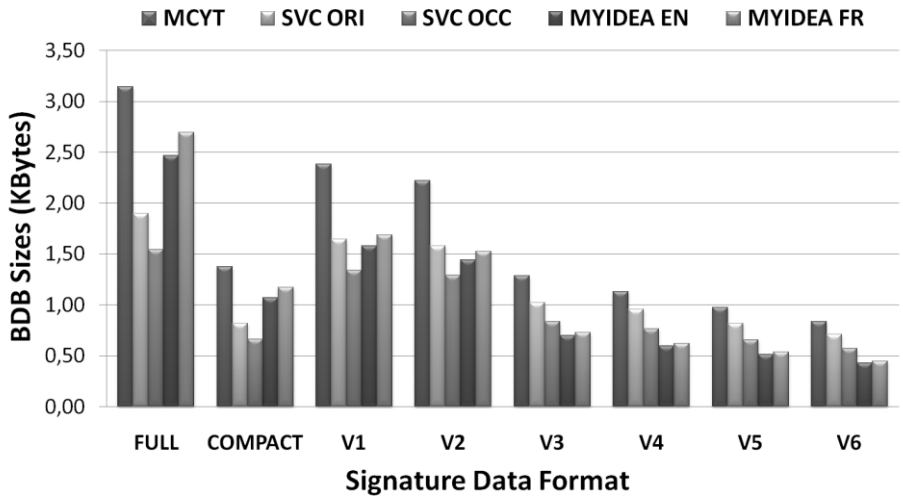


Figure 6-8 Average BDIR Sizes for Different Data Formats and datasets

The Full Format average sizes clearly show the difference between different types of users and datasets. Again, as mentioned in the previous chapter, it is worth highlighting the small sizes for the SVC2004, this is due to the non-recording of the pen-up movement, whereas MCyT and MyIDea contain this data. Different nationalities have different average data sizes, with the largest data size group formed from contributors to the MCyT dataset (Spanish users).

In Figure 6-9 the average compression ratio achieved from the different Data Formats and datasets are depicted. The Compact Format obtains a compression ratio of close to 56%, this is due to the storage of values for the position of both channels X and Y and the Pressure is stored in 1 byte instead of 2 bytes as is the case for the Full Format. The Compact Format does not contain a header and the channel time has not been included, this is because the values obtained do not have sufficient resolution within the range of 0-255.

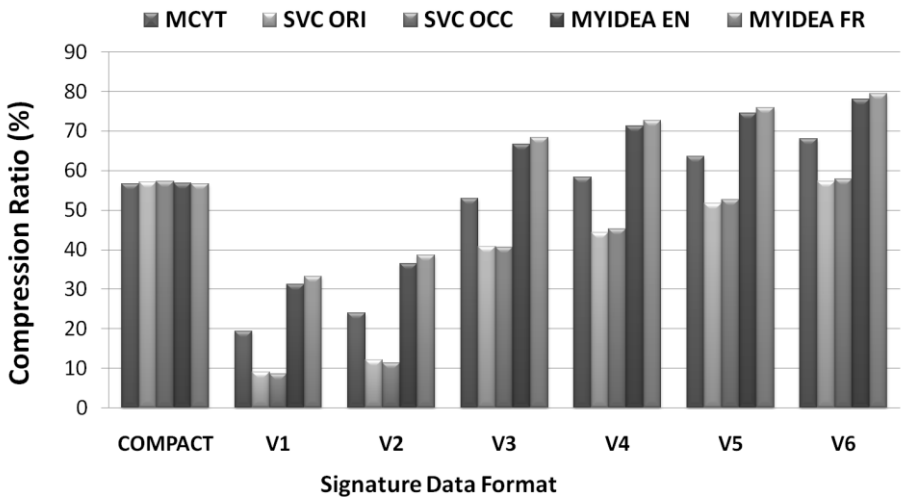


Figure 6-9 Compression Ratios for Different Data Formats and Datasets

An improved performance level for the compression ratio is achieved with the Compression Data Format Versions 4 to 6 (V4-6), obtaining close to 60% for all datasets in the case of Version 6. The larger the *Error_Difference_Level* value, the greater the compression ratio achieved.

The MyIDea datasets demonstrate greater compression ratios than the other datasets. This is due to the channels within the dataset which have the smallest data ranges, hence the consecutive differences between values are seen to be more repetitive leading to improve in the compression algorithm performance.

In Figure 6-10 the performance is shown for the algorithms according to the Compression Format Versions proposed. These results show the average between all datasets. The average compression ratio shown is the compression ratio achieved for the BDB Body only (i.e. sample points values). The best performance is achieved using Version 3

(V3), which stores the difference between channels in 2 bytes, therefore there are a large number of bytes with 0 values that represent a difference of less than 127. Alternatively, taking into account the final BDIR data, for Version 4 to 6 (V4-6), the compression algorithm ratios are lower than those achieved using Version 3. However, the final data size (BDB size, which contains the General Header + Representation Header + BDB Body Compressed) achieved is smaller due to the fact that the data that is compressed (the differences between consecutive sample point values stored in only 1 byte instead stored in 2 byte as in Version 3) also has a much smaller size.

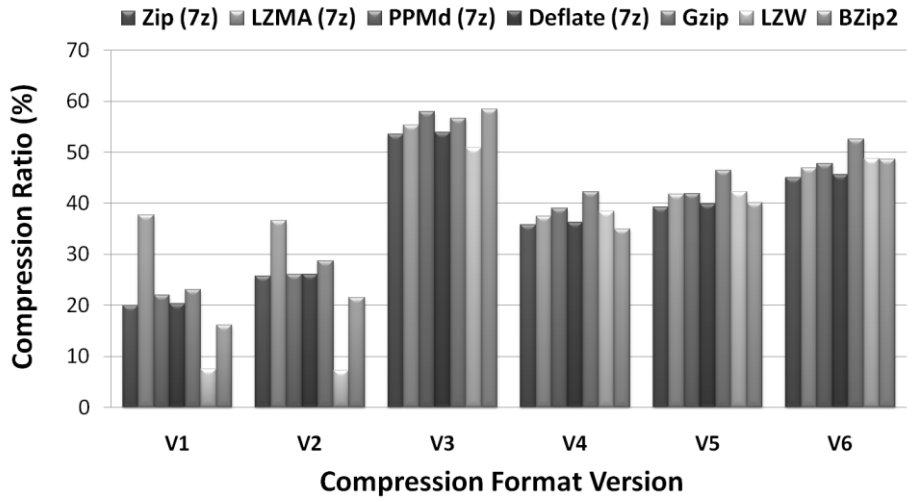


Figure 6-10 Compression Ratio for Different Compression Format Versions and Compression Algorithms

Regarding the performance of the different compression algorithms, it is worth pointing out the improved results achieved by GZip for Versions 4 and 5, although, all of the algorithms demonstrate good performance for Versions 3 to 5. For the versions that store the sample point values (Versions 1 and 2) instead of the difference between consecutive sample point values, the compression algorithm presents a significantly different level of performance, where the LZMA obtains the best results and LZW presents the lowest performance for both Version 1 and 2.

An additional analysis of the compression ratio, shown in Figure 6-11, presents the average compression ratios achieved for different datasets and Compression Algorithms, again, only for the compression ratio achieved within the BDB Body. The average compression ratios have been calculated across all the Compression Format Versions tested.

There are important differences between datasets. The MyIDea dataset obtains the best compression ratios, this is due to the reduced channel resolution when capturing the signatures, resulting in the difference between consecutive values being more repetitive. Within the SVC2004 dataset, there is not a large difference between Occidental and Oriental users, indicating that the compression algorithms obtain the same performance regardless of the type of signature. Another important result obtained from this analysis is that no one

Compression algorithm achieves much improved results over any other algorithm, however, it has been found that the LZW obtains the poorest compression ratios.

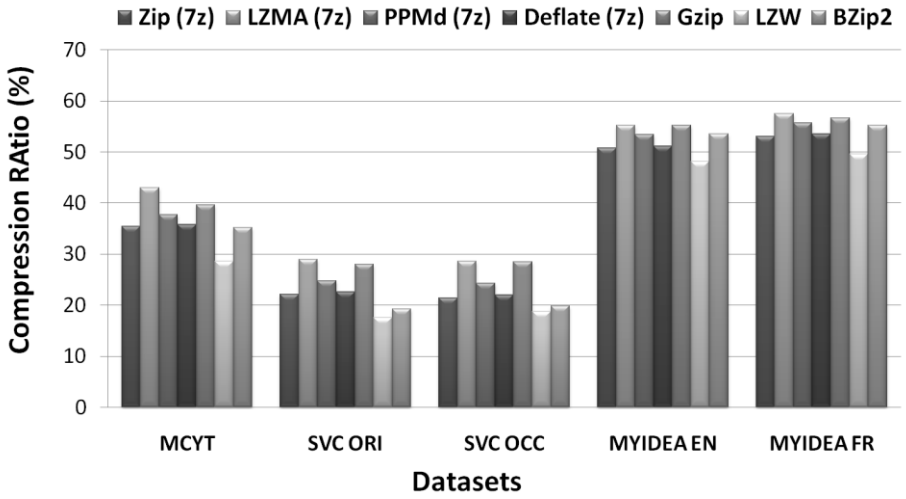


Figure 6-11 Compression Ratio for different Datasets and Compression Algorithms amongst all versions tested

6.5 ALGORITHM PERFORMANCE RESULTS

This section investigates the impact of the new and near-lossless compression format (version 5 and 6 defined in section 6.2), which imply a level of error or information lost, on the performance of the signature verification algorithm.

To demonstrate this, the same methodology presented in section 5.3 has been followed, using GMM and DTW algorithms that have been introduced in Chapter 4 along with the $MCyT_T$ dataset.

In Figure 6-12 the GMM algorithm performances for different data formats is presented: 19794-7.2WD2 Full Format and Compact Format, and the new near lossless data format, Versions 5 with an Error Level of 1, and Version 6 with an error level of 2. Again, as shown in the previous chapter 5, the Compact Format achieves poorer error rates than the Full Format, while the new near lossless data formats maintain the same level of error rates.

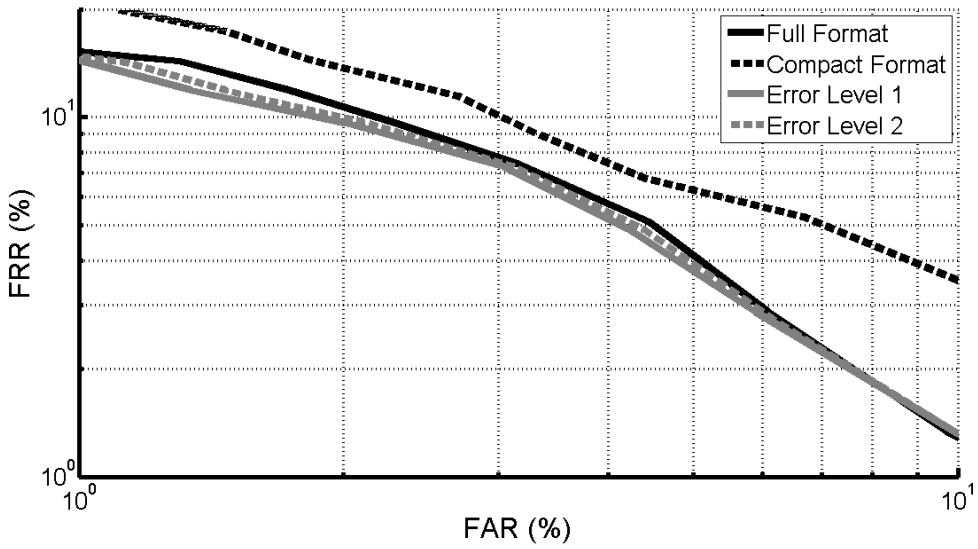


Figure 6-12 Error Rates for GMM and Different Data Formats

For the DTW algorithm case, in Figure 6-13 it is also shown that Version 5 (Error Level 1) and Version 6 (Error Level 2) have no impact on the performance of the verification algorithm.

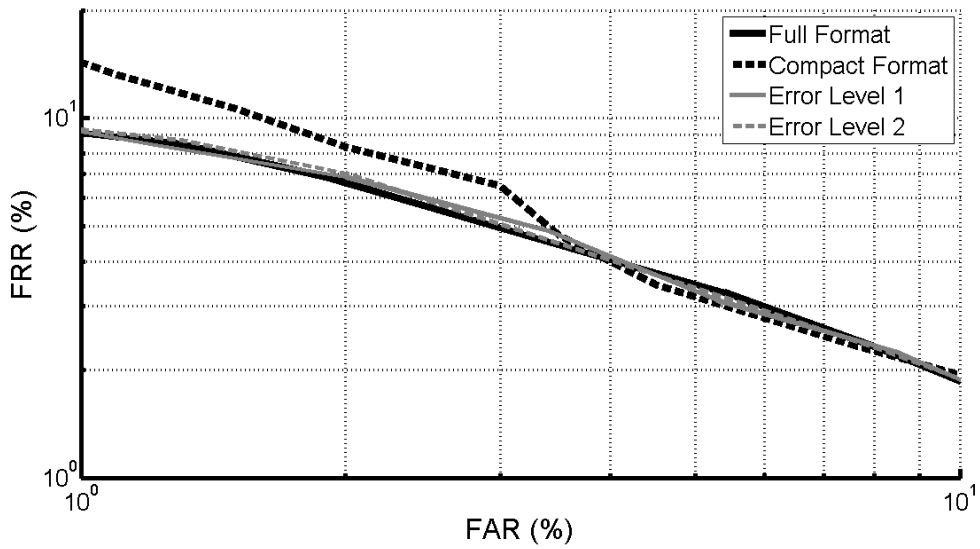


Figure 6-13 Error Rates for DTW and Different Data Formats

6.6 CONCLUSIONS

The compact format defined in 19794-7.2WD2 reaches a compression rate of close to 56%; this is mainly due to the conversion of the majority of the channels from a 2 to 1 byte range. However, the information lost as a result of this transformation can lead to increased error rates for several of the signature verification algorithms (see Figure 6-12).

The different approaches for the new compression formats have shown very different and diverse results. Versions 1 and 2, which store the original sample point values in two different ways (ordered by sample points or ordered by channels, respectively), did not demonstrate satisfactory compression ratios. Although, Versions 3 to 6, which do not store the original sample point values but the difference between consecutive sample point values, have achieved much improved compression ratios.

Both data formats Versions 3 (which stores the difference with a 2 byte resolution) and Version 4 (which stores the difference using a 1 byte resolution, where a control mechanism is used when the difference is out of the range of -126 to 126) achieved similar compression ratio results of close to 50%, where there is no loss in information from the original values, therefore, there will be no impact on the performance of the verification algorithm.

It has been shown in Figure 6-9, that the data formats Version 5 and 6 are observed to improve the performance of the compression ratio when compared with Version 4. Both Versions 5 and 6 also store the difference with a 1 byte resolution, however in this case, a predefined error level between the original values and those stored is allowed to maximize the compression ratio. These particular versions (5 and 6) imply a limited error between the original values and the recalculated (after compression) signals. The possibility of controlling the *Error Difference Level* value allows adjustment of the error level introduced by this data format to the resolution of the input device. Both versions are also observed not to impact on the performance of the verification algorithm (see Figure 6-12 and Figure-13).

Summarizing the results obtained, from the different approaches taken for a new compressed data format for signatures/sign time series, Versions 4 to 6 have shown good quality compression ratios (greater than the Compact Format) and the possibility of controlling the error introduced by the compression, thus further improving the compression ratios. This particular characteristic is of particular interest when considering the high resolutions being achieved by new signature input devices (lately devices report x and y resolutions of 5080 lpi and sampling rates of 200 Hz). This high resolution (both spatial and temporal) lead to bigger signature sample sizes where a compressed data format is envisaged to play an important role.

Regarding the compression algorithm, none of the different versions demonstrate a much improved performance over others.

These results were presented to the signature experts within the ISO/IEC JTC1 SC 37 subcommittee, whom invited the Spanish National Body to make a contribution on the compressed data format for signature time series. Based on the work presented in this Chapter 6, the author has made a recent contribution through AENOR (Spanish standard organization) for a new sub-format within 19797-7.2WD2, named Compression Data Format. This contribution was accepted, creating the new clause 9, "Compression Data Format" of the next working draft version (3rd WD) of the international standard "19794 Biometric data interchange formats – Part 7: Signature/Sign Time Series Data" [146].

Chapter 7 INTEROPERABILITY OF SIGNATURE BIOMETRICS AT DATA PROCESSED LEVEL

7.1 INTRODUCTION

In Chapter 3 (Section 3.6.2) the Signature/Sign Processed Dynamic Data Format that is defined in the 19794-11CD2 [119] and that does not imply any compression has been discussed. From this analysis, it has been shown that the Part 11 BDIR is larger than the BDIR in the 19794-7.2WD2 [115] Full Format. This lack of compression is primarily due to the quantity of information recorded in both pen and pressure strokes, repetitive data and the influence of noise on the channel values.

This Chapter provides solutions to the aforementioned drawbacks of the 19794-11CD2 data format, where reasonable compression ratios are achieved. A new definition is presented that minimizes the amount of data recorded and also removes repetitive redundant information. As well, solutions used to improve the robustness of the signature standard in the presence of channel noise have also been tested.

As in other sections throughout this Thesis, the MCyT [21], SVC2004 [22] and MyIDea [101] databases have been used to obtain the compression ratios of the different solutions tested. In the following subsections, the new data format will be explained in detail. Presented in this chapter are the different versions and data options that have been tested which are complimented with the BDIR size results that were obtained.

Moreover, the SC37 Signature experts have always aimed to develop the Part 11 so that the Part 7 can be reconstructed. This area of current interest will be addressed in Section 7.5.

Here, interpolation methods were used to perform this task, where different techniques were tested. A comparative analysis of the possibilities for reconstructing the Part7 between the 19794-11CD2 and the newly proposed data format will be presented. Further areas of investigation which will enhance interpolation results will also be indicated in this section.

7.2 PROPOSAL FOR A NEW DATA FORMAT

The new data format proposed in this Thesis attempts to solve the problems revealed in Section 5.2.2, these may be summarized as follows:

1. Excess Information Stored,
2. Repetitive Information,
3. Lack of Robustness against trembling/noise.

The excess information comes from the data stored within the pen and pressure strokes. A pen stroke is defined in 19794-11CD2 as “the movement of a pen between two singular points, defining a singular point as pen-down, turning points or pen-up”. A pressure stroke is defined as “the movement of a pen between the reversal of pen pressure direction, i.e. increasing pen pressure reversing to decreasing pen pressure and vice versa”. The information stored within the pen strokes are the values of X, Y and time channels at the pen stroke start and end , as well as minimum, maximum and average values for the x and y velocity, x and y acceleration and the pressure. Both the vector direction and the stroke lengths are also recorded. Considering all this information, the pen-stroke requires 46 Bytes.

The information contained within the pressure strokes are X, Y, Pressure and Time at the pressure stroke start and end, as well as minimum, maximum and average pressure values. For pressure-strokes 22 Bytes are required to store this information.

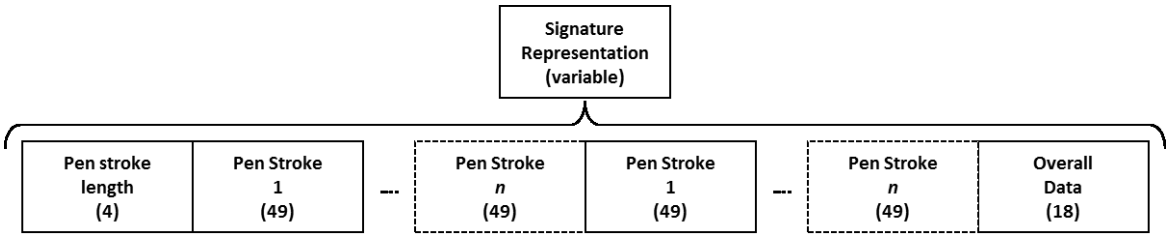


Figure 7-1 Signature Representation Data Block for 19794-11CD2

Although the number of elements stored (i.e. pen and pressure strokes) is much smaller in the Part 11, the final BDIR size remains larger than in the Part 7 Full Format. This is due to the amount of information stored within each stroke. The increased size of the recorded sample can also be accounted for by the repetitive information stored. In every stroke, the initial and final channel values are stored, thus implying duplication for consecutive strokes (e.g. for two consecutive strokes, the first stroke shares the same end channel values as the start channel values for the second stroke).

In order to solve both of these problems, a new definition and structure which completely changes the philosophy of the data to be store within the Part 11 is presented in this Thesis.

Instead of the aforementioned storage method, a reduced format is proposed. This novel proposal only stores the X, Y, Pressure and the Time channel values for those singular points (i.e. pen-down, pen-up, turning point and pressure reversal point), thus removing the stroke concept within the Part 11, see Figure 7-2.

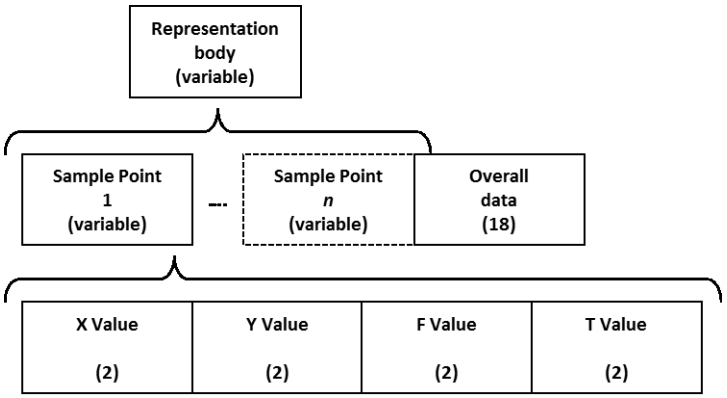


Figure 7-2 Signature Representation Data Block Proposed

With this definition, the 19794-11 can store the same quantity of information for every element stored (i.e. singular point) as the Part 7 Full Format (i.e. sample points). However, the Part 7 stores this information for every sample point whereas the Part 11 will only store values for the singular points considered. This methodology leads to a significant level of compression. By considering this technique, the first two problems identified in Chapter 5 are solved.

To deal with the third problem of noise in channel values, which implies the presence of spurious dynamic events, the diverse approaches used are presented in the following sections. These processes are based on different definitions and calculations of the singular points within the Part 11.

A singular point is defined as "*a collective event description of a pen-up, pen-down, turning point or pressure reversal point event*".

The pen-up and pen-down event are clearly defined within this standard. Pen-down is an "*event from which the pen tip touches the writing plane*", while a Pen-up is an "*event from when the pen tip does not touch the writing plane, after a pen-down event*".

These two singular points are not considered to be affected by any trembling or signal noise.

Also, a turning point may be defined as an "*event from which the direction of the pen is changed in either the x, the y, or both axes, or the sign of the curvature of the written signature/sign changes*". A similar definition is provided for the pressure-reversal point, "*event from which the pen pressure movement is reversed, i.e. point at which there is an increasing pen pressure reversing to decreasing pen pressure or vice versa*".

The definition for the turning point and pressure-reversal point are not robust when faced with noise or even slight trembling in the X, Y and Pressure channels, thus increasing the number of those points (i.e. turning and pressure-reversal points) by adding spurious singular points.

In order to avoid this effect, two different approaches have been studied. The first attempts to remove the noise/trembling from the channels using a low pass filter. The second makes use of a new definition for turning points and pressure-reversal points that is based on the zero crossing velocity. The channel velocity is calculated using a regression formula which reduces the influence of the noise/trembling.

7.3 PROPOSAL ANALYSIS

7.3.1 DATA FORMAT VERSIONS ANALYSED

In this section, the different data format versions analysed are described. Both data formats defined in the 19794-7.2WD2 [115] (prior to the addition of the format which is to be released and is a direct consequence of this Thesis) have been included in this analysis. The Full Format will be used to calculate compression ratios, while the Compact Format is used as a reference for the compression ratio.

The results from the 19794-11CD2 have also been included to compare the improvements obtained when using this new definition. The final versions that have been analysed focus on the impact of different parameters associated with the two different approaches used to avoid spurious turning and pressure-reversal points.

7.3.1.1 19794-7WD2 FULL FORMAT (FF)

As defined in the standard 19794-7.2WD2 Full Format, following the implementation explained in section 5.2.2.1.

7.3.1.2 19794-7WD2 COMPACT FORMAT (CF)

As defined in the standard 19794-7.2WD2 Compact Format, following the implementation explained in section 5.2.2.2.

7.3.1.3 19794-11CD2 (CD2 11)

As defined in 19794-11CD2, following the implementation explained in section 5.2.2.3.

7.3.1.4 NEW PART 11 DATA FORMAT PROPOSED VERSION 1 (V1)

In this newly proposed data format, Version 1, the main areas discussed in the introduction of this chapter will be presented and used to solve the problems encountered within the 19794-11CD2, and deal with the excess of stored and repetitive information.

The modifications introduced in section 7.2 (Figure 7-2) have been implemented in Version 1 (V1). The data storage from the stroke technique is replaced by data from the singular points. The data stored will only be:

- X and Y Position,
- Time,
- Pressure.

Apart from the new definition presented in this chapter, the same implementation explained in section 5.2.2.3 has also been followed.

7.3.1.5 NEW PART 11 DATA FORMAT PROPOSED VERSION 2 (V2)

The new signature data format proposed, i.e. Version 2, attempts to deal with the third problem identified in Chapter 5, regarding robustness against trembling/noise, where this has been revisited in section 7.2.

The data stored within every singular point is the same as that introduced in Version 1 (7.3.1.4), however, in order to avoid spurious singular points coming from trembling/noise during the signature capture process, the channels have been smoothened before calculating the singular points. This is carried out using a moving average as a low pass filter.

The moving average is the un-weighted mean of "n" data points that are centred on the point of interest:

$$ma_i = \frac{\sum_{i-\frac{n-1}{2}}^{i+\frac{n+1}{2}} x_i}{n} \tag{41}$$

According to this definition "n" can only take odd values and is generally referred to as the *Span*. The effect of the span value will be tested for this version using three different values:

- Version 2.3 will use a span value of 3,
- Version 2.5 will use a span value of 5,
- Version 2.7 will use a span value of 7.

7.3.1.6 NEW PART 11 DATA FORMAT PROPOSED VERSION 3 (V3)

In Version 3 (V3) a different approach that minimizes the noise impact on the number of singular points has been tested. In this case, instead of using a moving average filter, the zero-crossing channel velocity definition is tested. The noise is removed by using a robust regression formula to calculate the channel velocities.

The regression formula used [80] is:

$$reg(x(t), N) = \frac{\sum_{k=1}^{N} k \cdot \big(x(t+k) - x(t-k)\big)}{2\sum_{k=1}^{N} k^2} \tag{42}$$

Here "N" is the regression order.

In order to calculate the turning points for channels X, Y and F, the zero-crossing velocity definition has been used. According to this method, a turning point is regarded as the event that occurs when the velocity changes from positive to negative values or vice versa. Again, the next version will test the impact of the regression order parameter on the compression ratio. As in Version 2, different values for the regression order "N" have been tested, these are:

- Version 3.2 will test a regression order value of 2,
- Version 3.3 will test a regression order value of 3,
- Version 3.4 will test a regression order value of 4.

7.3.1.7 NEW PART 11 DATA FORMAT PROPOSED VERSION 4 (V4)

In Version 4 (V4), both approximations are used together. The values of the channels are first smoothened using the moving average (as explained in section 7.3.1.5) followed by an N order regression formula to calculate the channel velocities and to identify the singular points (as explained in section 7.3.1.6).

Different combinations for the span used within the moving average and the regression order used within the regression formula have been tested, these are:

- Version 4.3.2 will test a span value of 3 and a regression order value of 2,
- Version 4.5.3 will test a span value of 5 and a regression order value of 3,
- Version 4.7.4 will test a span value of 7 and a regression order value of 4.

7.3.1.8 SUMMARY

Table 7-1 summarizes all New Part 11 Data Format Versions analysed:

Table 7-1 Summary of New Part 11 Data Format Versions analysed

Version	ID	Definition	Span	Order (N)
19794-7.2WD2 Full Format	FF	As defined in the 19794-7.2WD2 Full Format	N/A	N/A
19794-7.2WD2 Compact Format	CF	As defined in the 19794-7.2WD2 Compact Format	N/A	N/A
19794-11CD2	CD2 11	As defined in the 19794-11CD2	N/A	N/A
New Part 11 V1	V1	excess of information and repetitive information solved by new structure and new information stored within singular points	N/A	N/A
New Part 11 V2.3	V2.3	As V1, but using a moving average (Span 3) to avoid spurious singular points by smoothing the channel values.	3	N/A
New Part 11 V2.5	V2.5	As V2.3, but using a span value of 5	5	N/A
New Part 11 V2.7	V2.7	As V2.3, but using a span value of 7	7	N/A
New Part 11 V3.2	V3.2	As V1, but a velocity regression formula has been used to calculated turning points and avoid spurious singular points. A regression order value of 2 is used.	N/A	2
New Part 11 V3.3	V3.3	As V3.2, but using a regression order value of 3	N/A	3
New Part 11 V3.4	V3.4	As V3.2, but using a regression order value of 4	N/A	4

Version	ID	Definition	Span	Order (N)
New Part 11 V4.3.2	V4.3.2	As V1, but using both methods, moving average and regression formula, to avoid spurious singular points. In this version, a span value of 3 and regression order of 2 have been used.	3	2
New Part 11 V4.5.3	V4.5.3	As V4.3.2, using a span value of 5 and regression order of 3.	5	3
New Part 11 V4.7.4	V4.7.4	As V4.3.2, using a span value of 7 and regression order of 4.	7	4

7.3.2 Analysis of Pen-Stroke Data Options

During the implementation of the different versions that analyse the BDIR size, another issue was detected.

In the 19794-11CD2 both the pressure and pen strokes include a field containing the type of pen-stoke or pressure-stroke (PST Value Field), i.e. for pen strokes, this field may have the following values:

1. From pen-down to turning point,
2. From turning point to turning point,
3. From turning point to pen-up,
4. From pen-down to pen-up.

This field allows the pen-up and pen-down event to be recognized. It can be used for signature verification algorithms which deal with signature segmentation performance based on such events. The absence of this field increases the difficulty of pen-down or pen-up events to be recognized. This may imply erroneous signature segmentations.

In order to solve this issue, the inclusion of this field is analysed and tested to determine its effect on the BDIR sizes. Two different options have been analysed which will be explained in the following subsections.

7.3.2.1 Option 1

For every singular point the same information as proposed earlier will be stored, i.e. X, Y, F and T channel values, but in this case, the type of dynamic event will be added to the Pen-Stroke Data Stored.

This new field will take the values as presented in Table 7-2.

Table 7-2 Allowed Values for type of dynamic event

Value	Meaning
0	Turning Point in X channel
1	Turning Point in Y channel
2	Turning Point in F channel
3	Pen-down
4	Pen-up

7.3.2.2 OPTION 2

In this option, two different types of strokes have to be defined (as in the 19794-11CD2):

- Pen-strokes,
- Pressure-strokes.

The pen-strokes will store the same data as proposed above, i.e. X, Y, F and T channel values. However, in this case the Pressure-strokes will also store the data proposed, i.e. X, Y, F and T channel values, in addition to the type of dynamic event. Table 7-3 shows the values allowed. This new field will be stored in 1 byte.

Table 7-3 Allowed Values for the type of dynamic event for a pressure-stroke

Value	Meaning
0	Turning Point in F channel
1	Pen-down
2	Pen-up

7.4 BIOMETRIC DATA INTERCHANGE RECORD SIZE FOR SIGNATURE/SIGN DATA FORMATS WITHIN THE 19794-11

As has been performed throughout this Thesis, only genuine signatures from the different databases (MCyT, MyIDea, and SVC2004) were used. These datasets were again divided in 5 different sub-datasets depending on the signature user's country and the character set used (Table 5-1). The results show the BDIR sizes and compression ratios for these different datasets and the 19794-7.2WD Full and Compact Format together with the different versions of the 19794-11CD2 analysed. Also a comparative graph showing the number of singular points, and its reduction, in both position and pressure will be presented.

For the instances stored following the 19794-7.2WD2, Full Format (FF in Figure 7-3), Compact Format (CF in Figure 7-3) and the 19794-11CD2 data format (CD2 11 in Figure 7-3), the implementation details described in sections 5.2.2.1-3 have been used. For the instances stored following the new formats proposed, Versions 1 to 4 (V1 to V4 in Figure 7-3), the singular points have been calculated following the techniques defined in section 7.3.

In Figure 7-3 the average BDIR sizes for all the data formats analysed are presented. It can clearly be seen how the new definitions proposed achieve much lower sizes when compared with the 19794-7.2WD2 Full Format, and much more efficient when considering the definition proposed in 19794-11CD2 in terms of compression. Also, it may be observed that the new versions also improve on the compression ratio achieved by the Compact Format defined in 19794-7.2WD2.

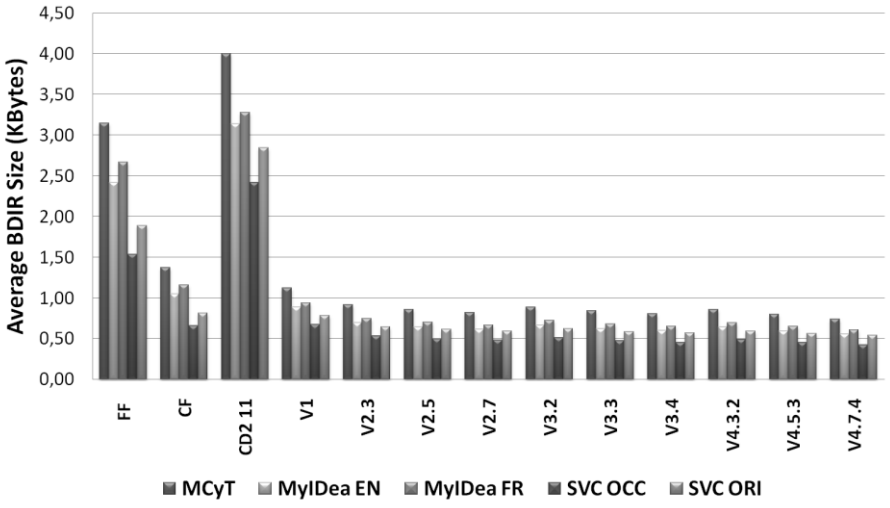


Figure 7-3 Average BDIR sizes for different data formats and datasets

By examining the compression ratios obtained the effectiveness of the new versions are clearly demonstrated, see Figure 7-4. This figure shows how the previous version, 19794-11CD2, does not imply any compression ratio, on the contrary, the sample size is seen to be greater. The new versions proposed obtain, to some extent, a higher compression ratio than the Compact Format. Once the different strategies for reducing the noise and spurious dynamic events are introduced (i.e. versions from V2 to V4), the compression ratio achieved increases, demonstrating improved effectiveness when compared with the Compact Format.

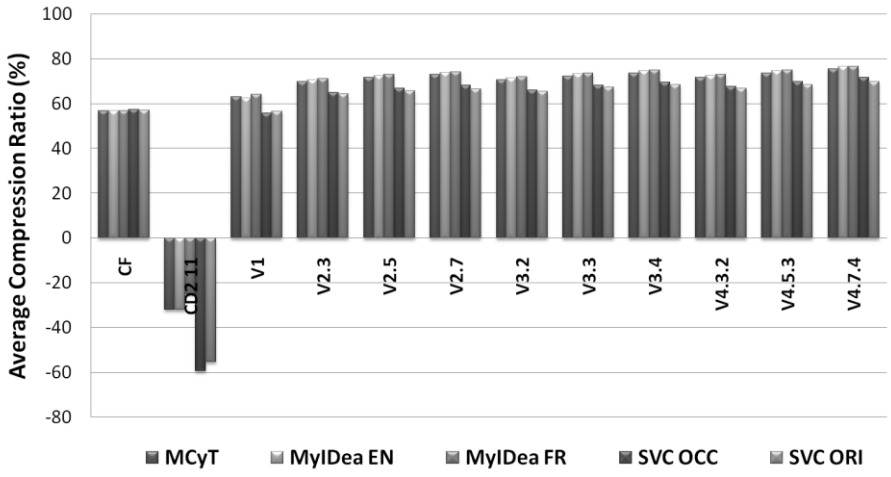


Figure 7-4 Average Compression Ratios for different data formats and datasets

It may also be observed from Figure 7-4 that the greater the parameter used to reduce the noise, the bigger the compression ratio. This parameter is the “span” for the smooth function (V2.X), the regression order (V3.X) and the combination of both (V4.X.X).

The different average number of pen singular points among all the versions tested is presented in Figure 7-5. From these results it should be highlighted that there is a greater number of pen singular points for the new definitions proposed. This is due to the fact that in these definitions, the pen singular points are calculated by considering all the sample points, including those where there is no pen contact with the tablet (pen-up movements), unlike the 19794-11CD2 which only considers the pen-down sample points. The fact that the SVC2004 only records pen-down movements explains how this dataset obtains the same number of points.

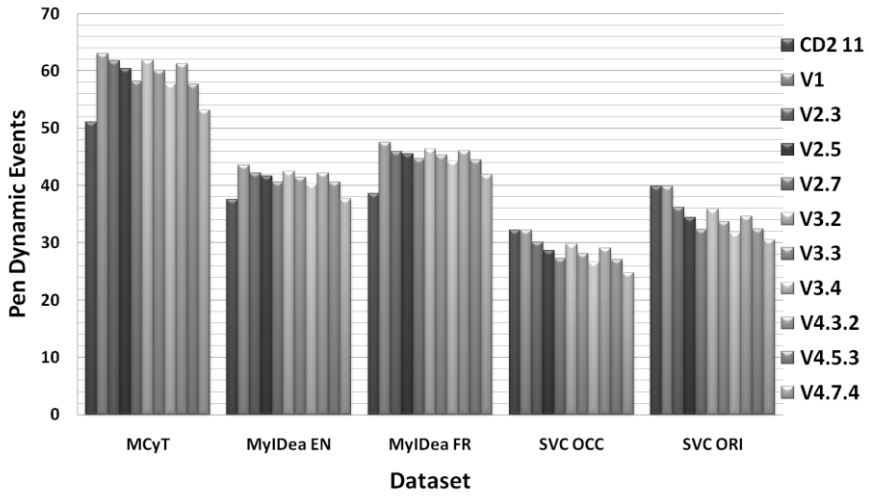


Figure 7-5 Average number of Pen Dynamic Events for different data format versions and datasets

The introduction of noise-reduction strategies is seen to reduce the number of points. This reduction is observed to be greater when the reduction parameter used is higher, where this is clearly seen in Figure 7-6 for pen singular points. Nevertheless, it will be demonstrated that the effect of these noise-reduction strategies is not as significant as for the pressure singular points.

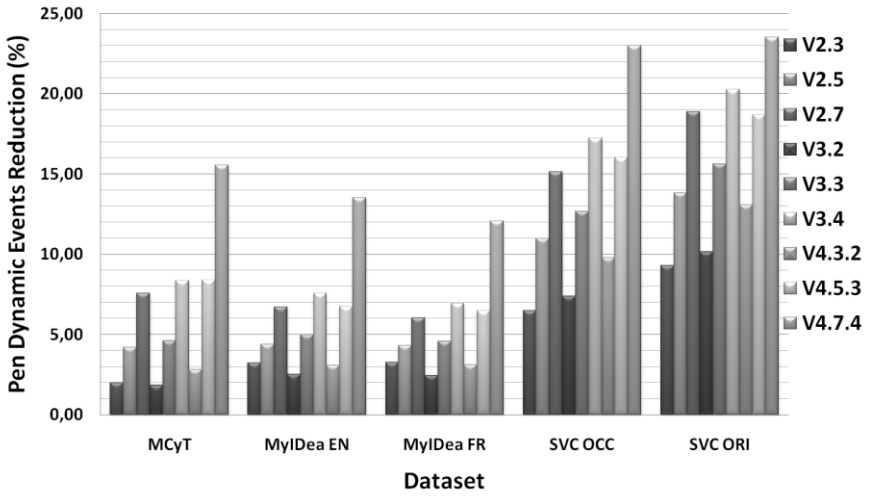


Figure 7-6 Average number of Pen Dynamic Events Reduction for different data format versions and datasets

In Figure 7-7 and Figure 7-8 the results for the pressure singular points are presented. In this case, there are no differences between the old and the new definitions as opposed to the pen singular points. This is because only pen-contact sample points are considered for

both definitions. It is worth highlighting that, in the pressure channel case, the noise-reduction strategies obtain a much lower number of singular points. The pressure signal acquired by the input devices (a pen tablet for all the databases used) has a noise level associated with it, thus implying the detection of false singular points. When using the noise-reduction strategies presented in this Thesis the number of pressure singular points is reduced significantly. For the majority of the different strategies used, a reduction of more than 30% is achieved, reaching more than 50% in some cases. Also, the reduction-parameter values have an important role when considering pressure singular points, where reductions beyond 15% have been observed in several cases.

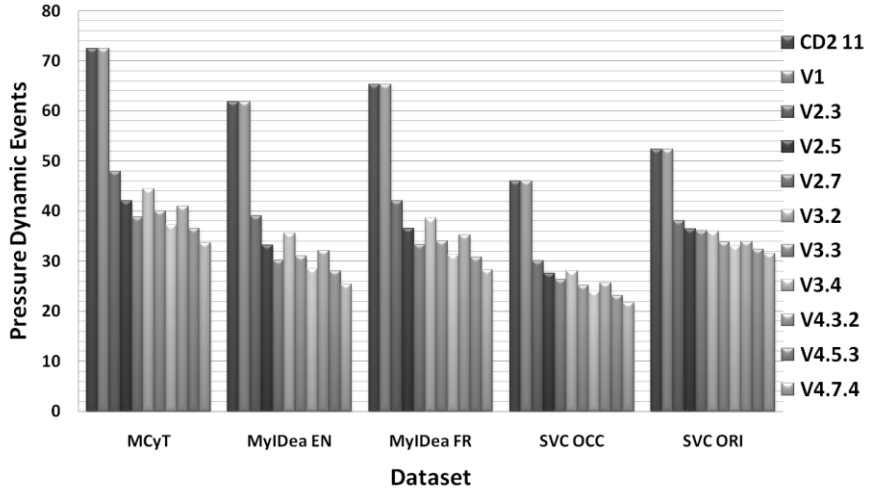


Figure 7-7 Average number of Pressure Dynamic Events for different data format versions and datasets

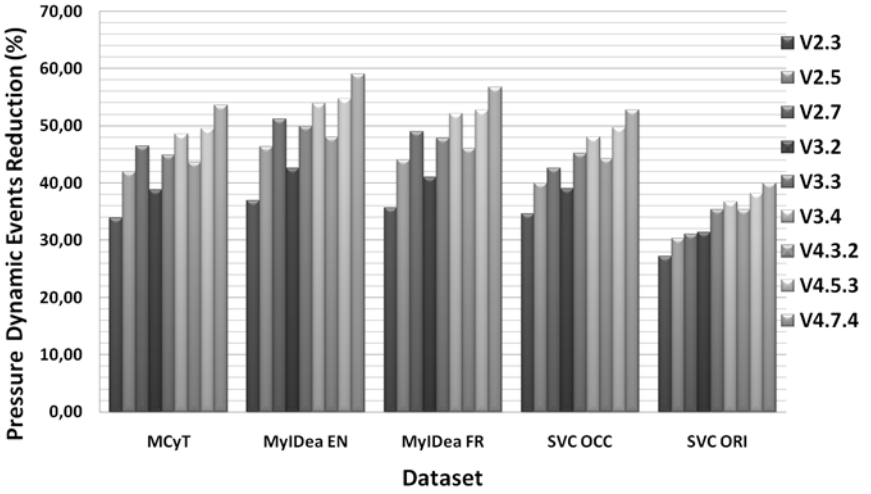


Figure 7-8 Average number of Pressure Dynamic Events Reduction for different data format versions and datasets

Regarding the different options presented to record the type of singular points, Figure 7-9 shows the average compression ratios for all the databases. It may be observed from Figure 7-9 that the impact of both options is not significant, being close to 3% for Option 1 (recording the singular point type in all the singular points) and 2% for Option 2 (recording the singular point type only in pressure singular points). In order to reduce the complexity of the data format proposed, Option 1 will be considered.

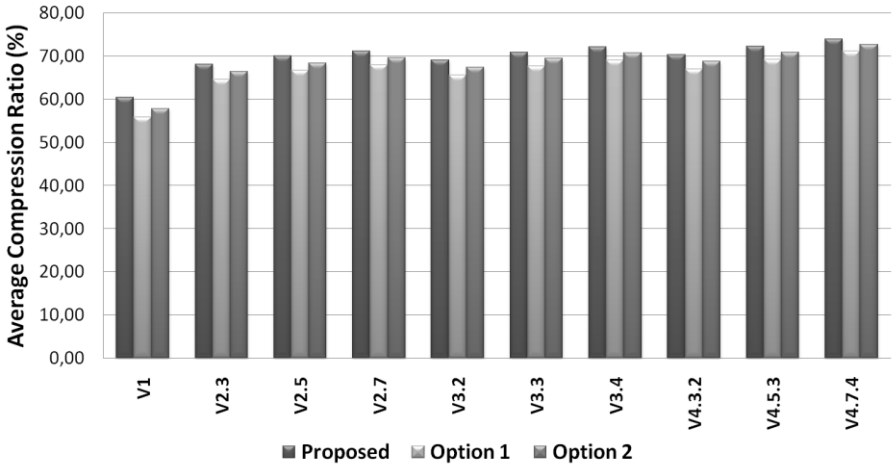


Figure 7-9 Average compression ratio for the different data formats proposed and the different options for storing singular point types.

7.5 Interpolation Error For Reconstructing 19794-7 from 19794-11 Data

One of the objectives of SC37 WG3 while developing the "19794-11 Signature/sign processed dynamic data format" was to assure that the Part 11 stores sufficient information to recreate the data stored in the Part 7 coming from capture devices.

Keeping this in mind, the current section will present several attempts carried out to recreate the Part 7 data from the Part 11 data by means of interpolation. This technique is defined as a method of constructing new data points within a certain range from a discrete set of known data points. This discrete set of points are those that are stored and defined by the Part 11, and are either pen-up, pen-down, pen or pressure turning points, where for all these points the X, Y, Pressure and Time values are stored.

Four of the most commonly used interpolation methods are:

- Nearest neighbour interpolation,
- Linear interpolation,
- Cubic spline interpolation,
- Piecewise cubic Hermite interpolation polynomial (PCHIP).

The first two methods, nearest neighbour and Liner interpolation are not used as they do not provide continuity in their derivative. As a result of this, the present study attempts to identify which method, either spline or PCHIP, is more suited to signature characteristics.

The cubic spline interpolation method is a popular choice because of its straightforward implementation producing a curve that appears seamless, being continues in its second derivative. However, the PCHIP is relatively similar in the way it constructs new data, however it does not imply a continuous second derivative.

The next section, 7.5.1, assumes that the cubic spline interpolation is better suited to the X and Y position signal characteristics and that the PCHIP is the preferred choice for Pressure signal characteristics. Here the different errors that arise from this assumption will be investigated for the 19794-11CD2 definition (without pen-up movements) and also the definition proposed in 7.2.

The experiments carried out in Section 7.5.2 will examine which interpolation method is better suited to signature characteristics, i.e. cubic spline interpolation or PCHIP, for both the X-Y position signals and the Pressure signals.

The last section will explore different strategies that can be employed to attain enhanced interpolation accuracy from the definition proposed.

7.5.1 INTERPOLATION FROM 19794-11CD2 AND THE NEW PROPOSAL

As explained above, this section will demonstrate the errors generated from the interpolation of the X and Y position and Pressure signals that arise from the data stored in the format defined by 19794-11CD2. These will be compared to the errors generated from the interpolation of the signals from the new data format proposed.

In order to perform this analysis, the information related to the different points stored within the 19794-11CD2 (starting and ending information for both pen and pressure strokes) has been gathered, and from this information the signals which are stored in the 19794-7.2WD2 Full Format have been interpolated. As the 19794-11CD2 only stores the pen-down movements, the interpolation has been carried out for each different stroke (movement between a pen-down event and a pen-up event), obtaining the corresponding time series made by the points that are time-spaced by 10 milliseconds.

The same procedure has been used for the information gathered from the new data format definition proposed, in this case, all the singular points stored from the 3 signals are used, i.e. X and Y position and Pressure. Also, the complete time sequence has been interpolated, including both pen-up and pen-down movements.

The errors generated have been calculated by comparing (only taking into account the pen-down movement, as the 19794-11CD2 does not store pen-up movements) the interpolated signals with the original signals (where these have been normalized to obtain values between 0 and 1 before transforming the 19794-7.2WD2 Full Format to the 19794-11CD2 and the new data format proposed). The differences between the original and the interpolated signals have been analysed by calculating the minimum distance from the point interpolated with respect to the corresponding point in the original signal and the 2 previous and 2 succeeding points, see Figure 7-10.

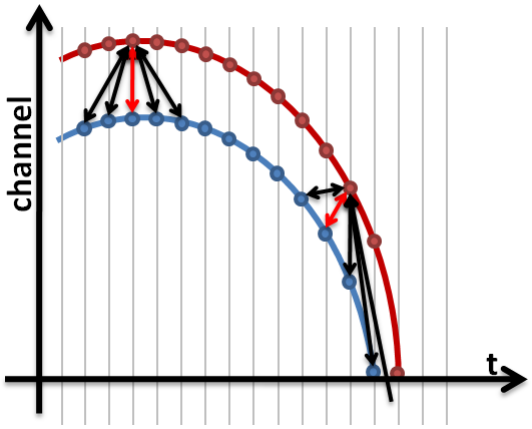


Figure 7-10 Error estimation for interpolated signal

The differences observed have been investigated, where the maximum and the main differences over the complete time sequence have been obtained. This process has been carried out for the complete set of signatures in the MCyT database. The following figures present the histograms obtained from this analysis where the maximum and mean errors may be observed.

In Figure 7-11 to Figure 7-14 the % errors generated using both formats, i.e. the 19794-11CD2 and the new proposal, are presented. It can be seen that the new proposal (light bars) demonstrates a greater frequency of low errors, while the 19794-11CD2 has a greater frequency of higher errors. This particular result signifies that the new proposal is not only much more compact, but can also store more useful information for interpolation of the original signals.

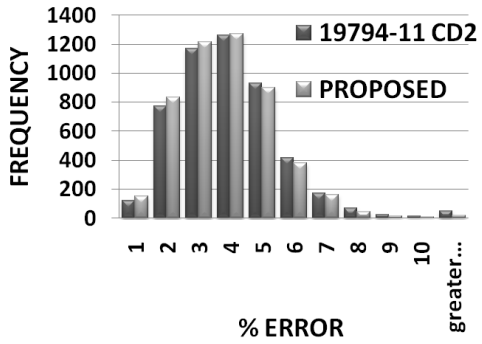


Figure 7-11 Maximum Error for the interpolation of the X-Y Position Signal, comparison between the 19794-11CD2 and the new Proposal

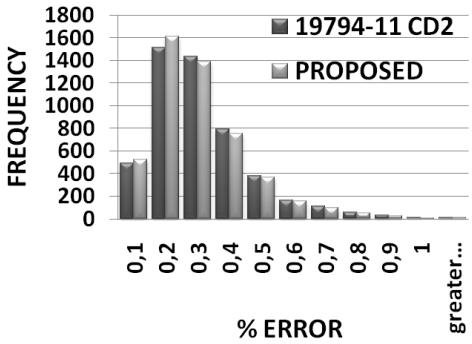


Figure 7-12 Mean Error for the interpolation of the X-Y Position Signal, comparison between the 19794-11CD2 and the new Proposal

This is particularly true for the mean pressure error, see Figure 7-14, where it can be observed that the number of errors over 0.5% is much greater for the 19794-11CD2. This result can be explained by the fact that the new proposal stores the Pressure for each and every singular point, regardless of where the singular points come from i.e. a X, Y or Pressure signal, whereas in the 19794-11CD2 the pressure information related to the starting and ending points are only stored within the pressure strokes and not within the pen strokes (where only the maximum, minimum and mean pressure values are stored).

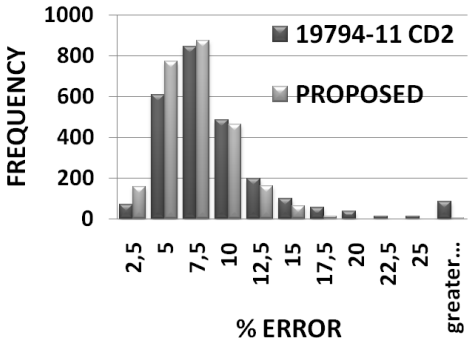


Figure 7-13 Maximum Error for the interpolation of the Pressure Signal, comparison between the 19794-11CD2 and the new Proposal

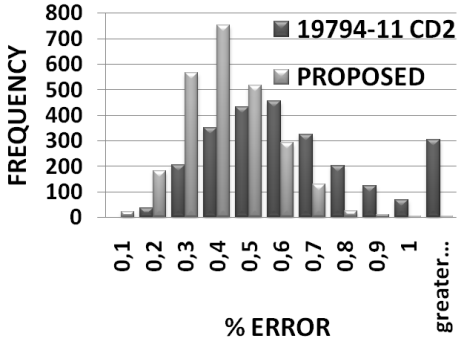


Figure 7-14 Mean Error for the interpolation of the Pressure Signal, comparison between the 19794-11CD2 and new Proposal

7.5.2 PCHIP AND SPLINE STUDY

At the beginning of this section it was stated that the Spline and PCHIP interpolations are relatively similar in the way they construct new data, but differ in some of their characteristics. These differences imply that the spline interpolation is better suited to the X and Y position signal characteristics and that the PCHIP interpolation is better suited to pressure signals.

The spline interpolation produces smoother results, and produces more accurate results when the data consists of values from a smooth function, while the PCHIP presents no overshooting and less oscillation when the data is not smooth. Thus, it is expected that the spline provides better results for signals such as the X and Y positions, where the time series is smoother and that the PCHIP is expected be more adequate for the Pressure signals.

In this section both interpolation methods are used to recreate signature signals (X and Y position, and pressure), as an attempt to identify if the aforementioned assumption is correct regarding the usage of each interpolation technique.

A similar methodology explained in the previous section has been used, i.e. changing the interpolation methods but maintaining the data stored within the new data format proposed, where the difference in this study is that both the pen-down and pen-up movements have been analysed.

In Figure 7-15 and Figure 7-16 the results obtained for the X and Y position signal recreation are presented, the maximum % error and mean % error frequencies are presented. It can clearly be seen that, taking into account all the sample points (both pen-up and pen-down movements), the cubic spline Interpolation technique obtains better results when compared to the PCHIP interpolation technique. This confirms the assumption that the

cubic spline Interpolation is better suited to the characteristics of both the X and Y temporal signals, this is due to the fact that these signals are normally smooth. Also this Interpolation technique assures that the derivatives of the signals are continuous.

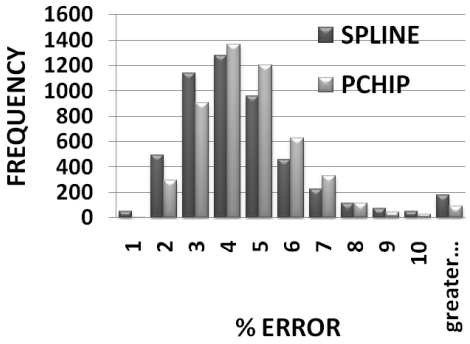


Figure 7-15 Maximum Error for the interpolation of the X-Y Position Signal, comparison between SPLINE and PCHIP interpolation methods

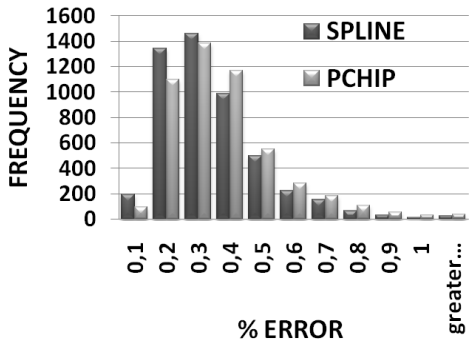


Figure 7-16 Mean Error for the interpolation of the X-Y Position Signal, comparison between SPLINE and PCHIP interpolation methods

Regarding the Pressure signals, the PCHIP demonstrates enhanced performance when recreating the original data from the information stored in the new data format proposed, see Figure 7-17 and Figure 7-18. The original pressure signal captured by the input device is observed to change more abruptly, and is not necessarily continuous in its derivative.

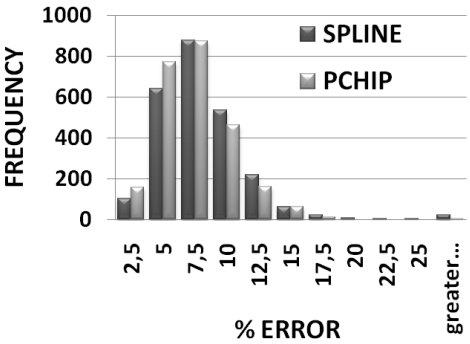


Figure 7-17 Maximum Error for the interpolation of the Pressure Signal, comparison between SPLINE and PCHIP interpolation methods

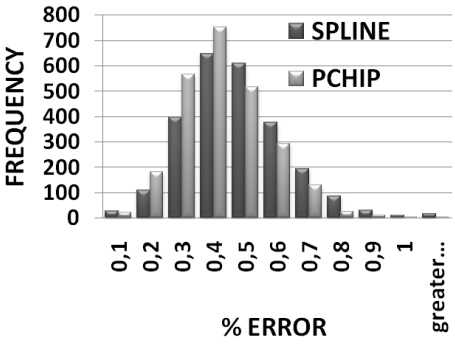


Figure 7-18 Mean Error for the interpolation of the Pressure Signal, comparison between SPLINE and PCHIP interpolation methods

7.5.3 ENHANCING INTERPOLATION

The previous sections of this chapter have shown the errors that are generated from interpolating the original data captured by the signature input devices from the data stored in the new data format proposed, and also from the data format defined in the 19794-11CD2.

From this analysis it has been shown that interpolation provides excellent results in terms of mean error, demonstrating a mean error of less than 0.5% for the majority of the signatures stored in the MCyT database. However, the maximum errors obtained are not as effective, where most of the signatures are below 5% for the position signals and 10% for the pressure signals.

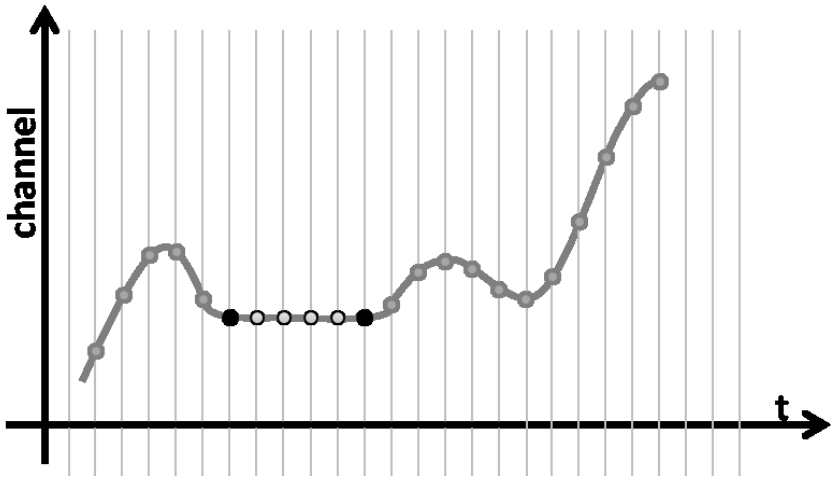


Figure 7-19 Stroke maintaining a certain constant value

In this section, results from a new singular point definition are presented. On observing the interpolation errors obtained, it has been seen that several of the errors that occur arise from the strokes, where the signals maintain a certain constant value (sample points between the two black points in Figure 7-19). The definition of singular points state the following "*event from which the direction of the pen is changed in either the x, the y, or both axes, or the sign of the curvature of the written signature/sign changes*". A similar definition is provided for the pressure-reversal point, "*event from which the pen pressure movement is reversed, i.e. point at which there is increasing pen pressure reversing to decreasing pen pressure or vice versa*". Taking into account these two definitions, only one of the extremes of those areas (black sample points in Figure 7-19) could fall within this definition.

In this section a new definition is proposed in order to specify both singular point extremes for the aforementioned areas. The new definition is as follows:

A turning point is an event in which the sign of the curvature of the written signature/sign changes. There can be 3 different cases for the sign of the curvature to change:

1.- Changing from positive to negative, in this case the turning point shall be the first point from which the curvature sign is different from the previous one.

2.- Changing either from positive or negative to zero, in this case the turning point shall be the first point with a value of 0 for its curvature.

3.- Changing from zero to either positive or negative, in this case the turning point shall be the last point with a value of 0 for its curvature.

These definitions not only take into consideration the point where the curvature changes from increasing to decreasing (or vice versa), but also for the point at which the curvature changes from 0 to increasing or decreasing (or vice versa).

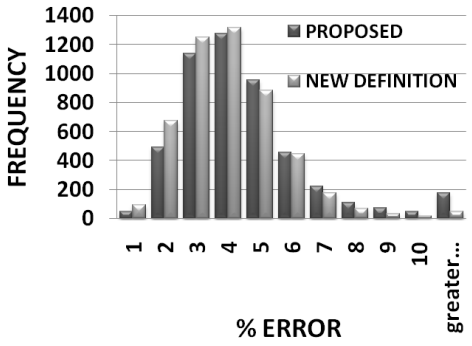


Figure 7-20 Maximum Error for the interpolation of X-Y Position Signal, comparison between SPLINE and PCHIP interpolation methods

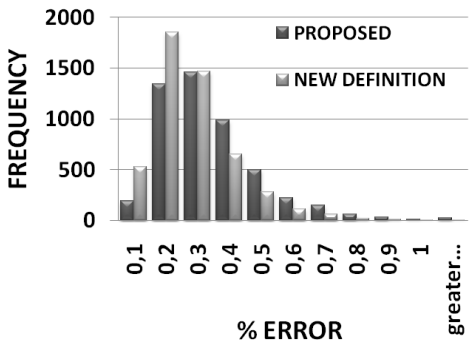


Figure 7-21 Mean Error for the interpolation of X-Y Position Signal, comparison between SPLINE and PCHIP interpolation methods

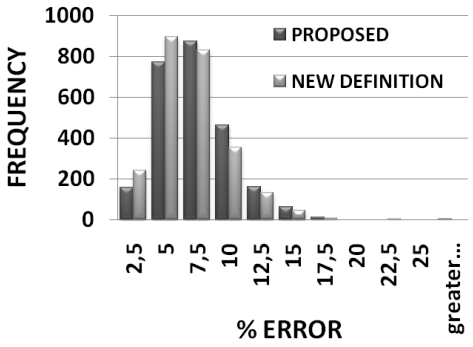


Figure 7-22 Maximum Error for the interpolation of the Pressure Signals, comparison between SPLINE and PCHIP interpolation methods

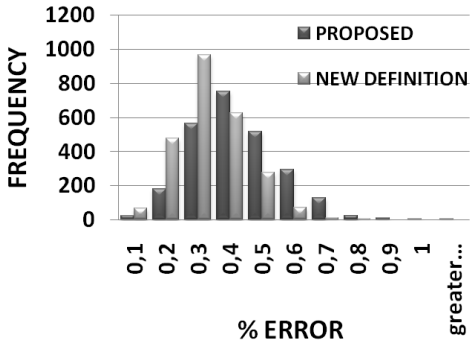


Figure 7-23 Mean Error for the interpolation of the Pressure Signals, comparison between SPLINE and PCHIP interpolation methods

In Figure 7-20 to Figure 7-23 it can be seen how the new definition improves on the error interpolation in all cases, for both maximum and mean error as well as X-Y signals and pressure. It may be concluded here that this new definition allows the interpolation of signature signals more accurately.

7.6 CONCLUSIONS

It has been shown in the previous section of this chapter that the new data format proposed in this Thesis for the Part 11 of the project 19794 implies a significant reduction in the BDB data size, where it has been observed to be lower than that obtained with the Compact Data Format defined in 19794-7. It has also been seen that the new definition provides better reconstruction of Part 7 data from the data stored in Part 11. This new definition, along with the results shown in this chapter, were presented to the SC37 WG3 and have been accepted as the new version of the committee draft 19794-11CD3.

Regarding the use of different strategies designed to avoid spurious dynamic events, different techniques have been studied, these are; smoothing of the sample point channels and the use of a regression formula for the velocity. The results show a reduction of close to 10% of the BDB data size (where the reduction is directly proportional to the span used on the smooth function or the regression formula order). The best results have been obtained when both strategies were combined. Again, these results were presented to the SC37 WG3, where the introduction of a noise-reduction strategy has been agreed upon. The method chosen, in part due to its straightforwardness, was the moving average with a 5 point span parameter.

The differences observed regarding the BDIR size when considering the two options of storing the type of dynamic events are not significant, around to 2-3%. Option 1 is observed to increase the BDB data size, but is less complex as it does not need to distinguish between pressure and pen singular points. This option was chosen by the SC37 WG3 biometric experts to be incorporated within the 19794 part 11.

Finally the studies for interpolating the data in Part 7 Full Format from part 11 was analyzed with the biometric experts in the SC37 WG3, and the new definition for singular points was added to part 11.

Summarizing, the 19794-11CD2 was implemented and analyzed in chapter 5, where several negative aspects were indicated that make this data format unsuitable for the Part 11 definition and targets, i.e. compression and adequate amount of information. In this chapter a new data format has been presented and different options tested. The results of this work were presented to the SC37 WG3 experts, whom accepted the proposed data format along with all the suggestions regarding improvements to the 19794-11. All these changes were accepted and implemented in the latest version of the Part 11: 19794-11FDIS [147].

Chapter 8 CONCLUSIONS AND FUTURE WORK

8.1 CONCLUSIONS

This Thesis has presented the research work carried out over the last 4 years on improving the viability of deploying on-line automatic signature verification systems such as ID documents. These systems are restricted in their data storage capacity and computational power. ID documents are designed to be used for a wide range of applications. By adopting international standards their interoperability is guaranteed.

The conclusions and contributions of this work have been commented in each chapter. In this final chapter a summary of all the conclusions reached is presented.

To introduce new potential readers to the topics debated in this Thesis, a brief overview of biometric systems has been presented. The emphasis here has been of a descriptive nature for the different phases contained within a biometric system, these are: enrolment, identification and verification. Following this overview, evaluation techniques for the performance of biometric systems were detailed. Readers were also introduced to the international standardisation organisations which lead biometrics standard developments. A brief introduction to the most relevant standard for this Thesis, ISO/IEC 19794 project, regarding biometric data interchange formats has been presented.

In chapter 3, the state of the art for on-line signature verification has been provided. Also, a detailed overview is presented on the current signature standards under development within the ISO/IEC in which the work of this Thesis has contributed, i.e. “19794-7 Signature/sign time series data” and “19794-11 Signature/Sign Processed Dynamic”.

Following the introductory chapters, the two major contributions of this Thesis to automatic signature verification have been detailed. The first, found in Chapter 4, discusses how to address the storage and computational load requirements. The second contribution of this Thesis has focussed on improving current signature standards. To accomplish these improvements, a viability analysis of the current standard was carried out and presented in chapter 5. The limitations that were identified have been addressed in chapters 6 and 7.

To reduce the size of the user model and the computational load of the algorithms in automatic signature verification, 4 different feature selection techniques were analyzed and applied to 2 different algorithms. Both of these signature verification algorithms have been detailed in chapter 4. The first is based on the GMM, where a total of 143 features were considered. The second algorithm investigated was based on the DTW, which is one of the most successful techniques for on-line signature verification systems.

Four different feature selection techniques were applied to these algorithms. The feature selections techniques were based on the Fisher Ratio (FR), the Principal Component Analysis (PCA), a combination of the FR and PCA, and finally, the Hellinger Distance (HD). Results show that the Principal Component Analysis obtained improved results when compared to the more commonly used Fisher Ratio. Furthermore, the combination of both techniques, FR and PCA, achieved increased performance when compared to both techniques used individually. The Hellinger Distance, a novel feature technique in the signature context, was also tested. The results obtained demonstrated superior performance when compared to the previous feature selection methods for low levels of overlap between the genuine and forgery feature distributions, such as the DTW algorithm proposed.

The aim of this work, i.e. reduction of both the user model size and the computational load, was achieved using feature selection techniques for both algorithms. In particular, the achievements obtained using the GMM algorithm, where both a feature vector of 28 elements and an extremely small 13 element feature vector were proposed and proven using a GMM consisting of only 4 Gaussian functions. The error rates obtained were in agreement with the state-of-the-art results previously published by authors using GMMs. However, in this case the number of Gaussian functions and the feature vector size were much lower. This result demonstrates a GMM algorithm with a reduced user model and a much lower computational load.

Using the DTW algorithm, a reduced feature vector of 6 elements was proposed and maintained an acceptable error rate of lower than 4% for the skilled forgeries, with a remarkable performance for the random forgery scenario, where an EER of 0.5% was achieved. This result shows that the feature selection technique was successfully applied.

However, the pressure data remained part of the feature vector along with the x and y axes data. As a result of this, the user model reduction and computational load did not arise from a reduced feature vector. Instead, the objectives were reached by reducing the number of sample points to 256, where the x and y axes, pressure and temporal data are stored. Using this amount of sample points it was demonstrated that sufficient information is stored for the same level of error rates as other alternatives using larger numbers of sample points. Also, it has been demonstrated that by storing the data within 128 sample points the error rates were still reasonable. Systems using this amount of sample points require even less storage space and computational load. Another outcome from these results is that the storage need and comparison time are fixed by these sample points, avoiding issues that may arise by signatures with high sample points.

Once the first objective for signature verification systems was achieved, the analysis of current international standardized data formats was then carried out. In chapter 5 a viability analysis of these international standards was presented. The results from this analysis showed the average signature sample size for different databases and standardized data formats. From these results it could be seen how the 19794-7.2WD2 compact format has a compression ratio of 56% while the 19794-11CD2, although describing itself as a compact format, does not imply any compression. In reality the signature sample size is increased. Also, the results from the algorithm's performance demonstrated a loss in information from both compact formats (19794-7.2WD2 Compact Format and 19794-11CD2) and how this loss impacts the algorithm's performance.

In chapter 6, a new compression data format was proposed to avoid reducing on the performance of algorithms when using the 19794-7.2WD2 Compact Format. This new compressed data format is based on lossless compression data algorithms. An evaluation of such lossless algorithms was carried out, along with different options for ordering and compressing the data acquired by the input device so as to maximize the compression ratio achieved by the compression techniques. A lossless compression data format was proposed, achieving a similar compression ratio to the 19794-7.2WD2 Compact Format but without losing any information. Moreover, a near lossless compression version was proposed, which outperforms the compression ratio of the compact format with minimum information loss. This near lossless proposal includes a mechanism which allows the level of information lost to be preset. This feature is of particular interest when considering the excessive resolution that current signature input devices are achieving, especially for the sample frequency, which is leading to larger signature sample data. The near lossless compression data format also demonstrated no impact on the algorithms performance. This version was proposed and accepted by the ISO/IEC SC37 WG3 experts to be included within the project 19794-7. It has become the third sub-format within this international standard (which is expected to be published in 2013).

Returning again to the problems encountered in the 19794-11CD2, chapter 7 has covered both the lack of compression and the information lost. A new data format structure

was proposed. By reducing the data stored, avoiding duplicated data and providing a new singular point definition the problems related to the lack of compression are solved. Different strategies to avoid spurious singular points were analyzed, selecting because of its simplicity, a 5 point moving average filter. Also, to reduce the amount of information lost in the 19794-11, an analysis of the possibility for interpolating the data in the 19794-7 Full Format from the 19794-11 was put forward; this again provided a new definition for singular points. The new data format obtains reasonable compression ratios, lower than the compression ratio obtained with the 19794-7.2WD2 Compact Format. This new data format definition was also presented to the ISO/IEC SC37 WG3 experts and accepted as the new committee draft for the 197974-11, 19794-11CD3. This standard is expected to be published in 2012.

8.2 FUTURE WORK

Automatic online signature verification is continually being established as a mature biometric modality. The verification error rates achieved for both random and skilled forgeries are sufficiently low allowing its use in a wide range of applications. This biometric modality has always been considered as an ideal candidate for implementation in commercial systems, but until now, its use has not been generalized.

Nowadays, many shopping commercial systems provide hardware that can be used for both online and offline verification. However, the offline version is the most commonly used, where the information is processed in a back-office, and is used as a means of double-checking money transactions.

There are two main factors which could indicate that this modality will be used massively in the near future. The first is related to the release of new specific input devices for capturing online signatures that offer comfortable and high quality acquisitions. Secondly, paperless processes are being introduced and encouraged in all companies as a result of their inherent savings. These two factors make signature input devices more and more available in many applications (e.g. shopping centres, post-offices) and will provide the signature biometric community with an ideal opportunity to finally introduce online signature verification to improve security.

The use of touch-screen mobile devices, such as tablets and smart phones, has increased enormously and become very widespread in recent years. This technology possesses the possibility of being used as signature input devices. The rapid and wide spread nature of capacitive touch-screen devices can provide signature acquisition for the majority of users almost anywhere.

The different areas discussed above show that this field of research is currently very active and has a promising future. Signature techniques are bringing new opportunities for online verification and are rapidly developing towards real systems. This growth and future applications are also bringing new issues that require research. Several of these topics are detailed below:

Implementation of online signature verification in real scenarios:

- To verify a signature sample against the identity claimed, a user model is always required. This user model can be stored within a token such as a smart card. Furthermore, the comparison score can be calculated within it. This Thesis has made progress in facilitating the storage and algorithm implementation in such devices, however the implementation should be carried out and evaluated to verify its real potential in real scenarios.

- The error rates reported in published works are based on public databases acquired in controlled scenarios, typically office-based scenarios. The implementation of automatic signature verification systems in more realistic scenarios should be addressed. In such cases, signature algorithms have to deal with a greater level of intra-user variability. These scenarios should include, as an example, standing signature acquisition, which is the case for the majority of point of sale scenarios. The collection of databases in these types of environments should be done, allowing researchers to investigate improvements to algorithms under such circumstances.

- The error rates achieved for random forgeries are another important aspect of online signature verification systems. The error rate levels compared with those achieved from skilled forgeries are still too high. The possibility of false accept errors when a user is signing with his/her own signature while attempting to impersonate other identities reduces the reliability of signature verification systems, especially when security levels are of top priority. In this case, the error rates should be as close to zero as possible. Lately, research in this area has mainly focussed on skilled forgeries. More thorough investigations on random forgeries should be addressed.

- A further two issues that arise from signature applications in real scenarios are ageing and user model updates. Signatures are observed to evolve over time, as users get older their motion capabilities change. This should be reflected in the user model by performing regular updates. Only a limited amount of research has been carried out in this area, requiring a more thorough overview. In order to research such issues, public databases where signature ageing is taken into account should be acquired. This must comply with data protection laws, complicating this line of investigation.

Online signature verification in mobile devices:

- As previously stated, new mobile devices will enhance the area of application of signature acquisition for a wide range of different scenarios, such as: office, home, standing, sitting, etc. The act of signing on these devices is not as natural as on desktop input devices, i.e. digital tablets. Hence, new databases are required which include signatures acquired with these particular devices for different scenarios. The BioSecure database has presented a reasonable attempt at testing mobile devices, however the input device chosen was a PDA, which is currently not the most commonly used mobile input device.

- More research is required on the acquisition quality and how it affects the verification algorithms. The acquisition process of the new capacitive touch-screen devices is different to the more commonly used digital tablets. Instead of capturing points at specific sampling rates, these devices use an event-based capturing process, which does not imply a homogeneous sampling rate. Also, the information provided for the pressure is not as good as those captured using digital tablets. The lack of quality pressure information and the non-homogeneous nature of the system may lead to the requirement for specifically developed algorithms.

These are only just some of the future areas of research that can be addressed and continued from the work presented in this Thesis. Obviously there are many others that should be considered by future researchers, such as: sociology, usability, combination with other modalities, etc. All new efforts in this the field of Biometrics is welcomed by the scientific community to offer citizens the best and most up to date possible solutions.

Chapter 9 REFERENCES

[1] ISO/IEC TR 24741, *"Information Technology - Biometrics Tutorial"*, ISO/IEC, Geneva, 2006.

[2] T. Dunstone and N. Yager, "*Biometric System and Data Analysis: Design, Evaluation, and Data Mining",* Springer Publishing Company, 2008.

[3] International Biometric Group (IBG), "Biometric Market and Industry Report, 2006-2010," International Biometric Group, 2005.

[4] S. Cole, *"History of Fingerprint Pattern Recognition",* Automatic Fingerprint Recognition Systems, Springer New York, pp. 1-25, 2004.

[5] D. Maltoni*, et al.*, "*Handbook of Fingerprint Recognition",* Springer London, 2009.

[6] R. Bolle*, et al.*, "*Guide to Biometrics",* Springer-verlag, 2003.

[7] A. A. Moenssens, "*Fingerprint techniques",* Chilton Book Co, Philadelphia, 1971.

[8] B. H. Juang and L. R. Rabiner, *"Automatic speech recognition - A brief history of the technology development,"* Elsevier Encyclopedia of Language and Linguistics, 2005.

[9] W. Bledsoe, *"The Model Method in Facial Recognition",* Panoramic Research, Palo Alto California, 1964.

[10] IEEE, "*Certified Biometrics Professional (CBP) Learning System",* IEEE, 2009.

[11] R. McCabe, *"Fingerprint Interoperability Standards",* in *Automatic Fingerprint Recognition Systems*, Eds: N. Ratha and R. Bolle, Springer New York, pp. 433-451, 2004.

[12] N. Duta, *"A survey of biometric technology based on hand shape"* in Pattern Recognition, vol. 42, pp. 2797-2806, 2009.

[13] ANSI-NIST, *"Fingerprint identification - data format for information interchange",* American national standards institute, New York, 1986.

[14] J. G. Daugman, *"High confidence visual recognition of persons by a test of statistical independence",* IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 15, pp. 1148-1161, 1993.

[15] W. Zhao*, et al.*, *"Face recognition: A literature survey"* ACM Comput. Surv.*,* vol. 35, pp. 399-458, 2003.
[16] A. K. Jain*, et al.*, "*Biometrics, Personal Identification in Networked Society: Personal Identification in Networked Society"* Kluwer Academic Publishers Norwell, MA, USA, 1998.
[17] A. J. Mansfield and J. L. Wayman, *"Best Practices in Testing and reporting Perfomance of Biometric Devices (V 2.01)*", Biometrics Working Group, UK, 2002.
[18] ISO/IEC IS 19795-1:2006, *"Information technology - Biometric performance testing and reporting - Part 1: Principles and framework"*, ISO/IEC, Geneva, 2006.
[19] ISO/IEC IS 19795-2:2007, *"Information technology -- Biometric performance testing and reporting -- Part 2: Testing methodologies for technology and scenario evaluation",* ISO/IEC, Geneva, 2007.
[20] D. Petrovska-Delacretaz*, et al.*, "*Guide to Biometric Reference Systems and Performance Evaluation*", Springer-Verlag, London, 2009.
[21] J. Ortega-Garcia*, et al.*, *"MCYT baseline corpus: a bimodal biometric database", IEEE Proceedings Vision, Image and Signal Processing,* vol. 150, pp. 395-401, 2003.
[22] D. Y. Yeung*, et al.*, *"SVC2004: First international signature verification competition"* in Proceedings Biometric Authentication*,* vol. 3072, pp. 16-22, 2004.
[23] ISO/IEC Guide2:2004, *"Standardization and related activities -- General vocabulary",* ISO/IEC, Geneva, 2004.
[24] International Organization for Standardization, Available: www.iso.org, 2011
[25] International Electrotechnical Commission, Available: www.iec.ch, 2011
[26] ISO/IEC Directives 1, *"Directives, Part 1: Procedures for the technical work (8th Ed.)",* ISO/IEC, Geneva, 2011.
[27] ISO/IEC IS 19785-1:2006, *"Information technology - Common Biometric Exchange Formats Framework - Part 1: Data element specification",* ISO/IEC, Geneva, 2006.
[28] ISO/IEC FDIS 19794-1, *"FDIS Information Technology - Biometric data interchange formats – Part 1: Framework"* ISO/IEC, Geneva, 2010.
[29] A. Osborn, "*Questioned Document",* Boyd Printing Co, New York, 1929.
[30] A. Mauceri, *"Feasibility studies of personal identification by signature verification",* RADC TR 65 33, Space and Information System Division, North American Aviation Co, 1965.
[31] N. M. Herbst and J. H. Morrissey, *"Signature verification method and apparatus",* US Patent 3983535, 1976.
[32] N. M. Herbst and C. N. Liu, *"Automatic Signature Verification Based on Accelerometry",* IBM Journal of Research and Development*,* vol. 21, pp. 245-253, 1977.
[33] R. F. Farag and Y. T. Chien, *"Online Signature Verification"* in Proceedings of the International Conference on Online Interactive Computing, Brunel University, London, 1972.
[34] R. Plamondon and G. Lorette, *"Automatic signature verification and writer identification - the state of the art" Pattern Recognition,* vol. 22, pp. 107-131, 1989.
[35] F. Leclerc and R. Plamondon, *"Automatic signature verification: the state of the art-1989-1993",* IEEE Transactions on Pattern Recognition and Artificial Intelligence*,* vol. 8, pp. 643-660, 1994.
[36] R. Plamondon and S. N. Srihari, *"Online and off-line handwriting recognition: a comprehensive survey",* IEEE Transactions on Pattern Analysis and Machine Intelligence*, ,* vol. 22, pp. 63-84, 2000.
[37] G. Gupta and A. McCabe, *"A review of dynamic handwritten signature verification",* Department of Computer Science, James Cook University Townsville,1997.

[38] G. Dimauro, *"Recent Advancements in Automatic Signature Verification",* International Workshop on Frontiers in Handwriting Recognition, pp. 179-184, 2004.

[39] M. Faundez-Zanuy, *"Signature recognition state-of-the-art"*, IEEE Aerospace and Electronic Systems Magazine vol. 20, pp. 28-32, 2005.

[40] G. K. Gupta, *"The State of the Art in the On-Line Handwritten Signature Verification"* in Academic Press, Faculty of Information Technology, Australia, 2006.

[41] D. Impedovo and G. Pirlo, *"Automatic Signature Verification: The State of the Art"*, IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*,* vol. 38, pp. 609-635, 2008.

[42] S. J. Elliott, *"Development of a biometric testing protocol for dynamic signature verification"* in International Conference on Control, Automation, Robotics and Vision*,* vol. 2, pp. 782-787, 2002.

[43] C. Gruber*, et al.*, *"A Flexible Architecture for Online Signature Verification Based on a Novel Biometric Pen"* in IEEE Mountain Workshop on Adaptive and Learning Systems, pp. 110-115, 2006.

[44] C. Hook*, et al.*, *"A Novel Digitizing Pen for the Analysis of Pen Pressure and Inclination in Handwriting Biometrics"* in Lecture Notes in Computer Science. vol. 3087, Springer, pp. 283-294, 2004.

[45] R. Baron and R. Plamondon, *"Acceleration measurement with an instrumented pen for signature verification and handwriting analysis",* IEEE Transactions on Instrumentation and Measurement, vol. 38, pp. 1132-1138, 1989.

[46] W. Jeen-Shing*, et al.*, *"An Inertial-Measurement-Unit-Based Pen With a Trajectory Reconstruction Algorithm and Its Applications"*, IEEE Transactions on Industrial Electronics*,* vol. 57, pp. 3508-3521, 2010.

[47] Genius Tablets, Available: www.geniustablets.com, 2011.

[48] Wacom Tablets, Available: www.wacom.com, 2011.

[49] Y. Sato and K. Kogure, *"Online signature verification based on shape, motion, and writing pressure"* in 6th Int. Conf. on Pattern Recognition, pp. 823 - 826, 1982.

[50] R. Plamondon and M. Parizeau, *"Signature verification from position, velocity and acceleration signals: a comparative study"* in 9th International Conference on Pattern Recognition, vol. 1 , pp. 260-265 , 1988.

[51] R. Martens and L. Claesen, *"On-line signature verification by dynamic time-warping"* in Proceedings of the 13th International Conference on Pattern Recognition, vol. 3, pp. 38-42, 1996*.*

[52] J. Fierrez-Aguilar*, et al.*, *"Target Dependent Score Normalization Techniques and Their Application to Signature Verification"* in IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews, vol.35, pp.418-425, 2005.

[53] M. Wirotius*, et al.*, *"Distance and matching for authentication by on-line signature"* in Fourth IEEE Workshop on Automatic Identification Advanced Technologies, pp. 230-235, 2005.

[54] B. Kar*, et al.*, *"DTW Based Verification Scheme of Biometric Signatures"* in IEEE International Conference on Industrial Technology, pp. 381-386, 2006.

[55] O. Henniger and S. Muller, *"Effects of Time Normalization on the Accuracy of Dynamic Time Warping"* in First IEEE International Conference on Biometrics: Theory, Applications, and Systems, pp 1-6, 2007.

[56] C. Won-Du and S. Jungpil, *"Modified Dynamic Time Warping for Stroke-Based On-line Signature Verification"* in Ninth International Conference on Document Analysis and Recognition, pp. 724-728, 2007.

[57] M. I. Khalil*, et al.*, *"Enhanced DTW based on-line signature verification"* in 16th IEEE International Conference on Image Processing, pp. 2713-2716, 2009.

[58] J. Pascual-Gaspar*, et al.*, *"Practical On-Line Signature Verification"* in Lecture Notes in Computer Science, Vol. 5558, pp 1180-1189, 2009.

[59] A. Kholmatov and K. Yanikoglu, *"Identity authentication using improved online signature verification method",* in Pattern Recognition Letters, Elsevier Science Inc. New York, vol. 26, pp. 2400-2408, 2005.

[60] A. K. Jain*, et al.*, *"On-line signature verification"* in Pattern Recognition Letters, Elsevier Science Inc. New York, vol. 35, pp. 2963-2972, 2002.

[61] M. Papaj and E. Hermanowicz, *"Identity verification using complex representations of the handwritten signature"* in 2nd International Conference on Information Technology pp. 79-82, 2010.

[62] M. Wirotius*, et al.*, *"Selection of points for on-line signature comparison"* in Ninth International Workshop on Frontiers in Handwriting Recognition, pp. 503-508, 2004.

[63] B. Li*, et al.*, *"On-Line Signature Verification Based on PCA (Principal Component Analysis) and MCA (Minor Component Analysis)"* in Lecture Notes in Computer Science, 2004, Vol. 3072, pp. 1-28, 2004.

[64] H. Feng and C. C. Wah, *"Online signature verification using a new extreme points warping technique"* in Pattern Recognition Letters, vol. 24, pp. 2943-2951, 2003.

[65] B. Wirtz, *"Stroke-based time warping for signature verification"* in Proceedings of the Third International Conference on Document Analysis and Recognition, vol. 1, pp. 179-182, 2005.

[66] T. K. Worthington*, et al.*, *"Ibm dynamic signature verification"* in International Conference on Computer Security, pp. 129-154, Amsterdam, 1985.

[67] Y.-C. Cheng and S.-Y. Lu, *"Waveform Correlation by Tree Matching"* in IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 299-305, 1985.

[68] S. Y. Lu, *"A Tree-Matching Algorithm Based on Node Splitting and Merging"* in IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 249-256, 1984.

[69] M. Parizeau and R. Plamondon, *"A comparative analysis of regional correlation, dynamic time warping, and skeletal tree matching for signature verification"* in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 12, pp. 710-717, 1990.

[70] N. Houmani*, et al.*, *"BioSecure Signature Evaluation Campaign 2009 (BSEC'2009): Results"* TELECOM & Management SudParis, Dept. EPH, Every, France, 2009.

[71] G. Dimauro*, et al.*, *"Component-oriented algorithms for signature verification"* International Journal of Pattern Recognition and Artificial Intelligence*,* vol. 8, pp. 771-794, 1994.

[72] W. Liang*, et al.*, *"Signature Verification using Integrated Classifiers,"* in Chinese Conference on Biometric Recognition, China, 2003.

[73] B. Fang and Y. Tang, *"Reduction of Feature Statistics Estimation Error for Small Training Sample Size in Off-Line Signature Verification"* in Lecture Notes in Computer Science, Springer, Vol. 3072, pp. 1-11, 2004.

[74] S. Krawczyk and A. Jain, *"Securing Electronic Medical Records Using Biometric Authentication"* in Audio- and Video-Based Biometric Person Authentication, Springer, vol. 3546, pp. 435-444, 2005.

[75] L. Baum and T. Petrie, "*Statistical Inference for Probabilistic Functions of Finite State Markov Chains"* in The Annals of Mathematical Statistics, vol. 37, pp. 1554-1563, 1966.

[76] L. Rabiner and B. Juang, *"An introduction to hidden Markov models"* in IEEE ASSP Magazine, vol. 3, pp. 4-16, 1986.

[77] J. G. A. Dolfing, *"Handwriting recognition and verification, a Hidden Markov approach",* Philips Electronics N.V., 1998.

[78] J. G. A. Dolfing*, et al.*, *"On-line signature verification with hidden Markov models"* in Fourteenth International Conference on Pattern Recognition, vol.2, pp. 1309-1312, 1988.

[79] R. S. Kashi*, et al.*, *"On-line handwritten signature verification using hidden Markov model features"* in Proceedings of the Fourth International Conference on Document Analysis and Recognition, vol. 1, pp. 253-257, 1994.

[80] V. Bao Ly*, et al.*, *"On Using the Viterbi Path Along With HMM Likelihood Information for Online Signature Verification"* in IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, vol. 37, pp. 1237-1247, 2007.

[81] Q. Zhong-Hua*, et al.*, *"A Hybrid HMM/ANN Based Approach for Online Signature Verification"* in International Joint Conference on Neural Networks, pp. 402-405, 2007.

[82] L. Yang*, et al.*, *"Application of hidden Markov models for signature verification"* in Pattern Recognition, vol. 28, pp. 161-170, 1995.

[83] J. Richiardi and A. Drygajlo, *"Gaussian Mixture Models for on-line signature verification"* Proceedings of the 2003 ACM SIGMM workshop on Biometrics methods and applications, Berkley, California, 2003.

[84] W. Liang*, et al.*, *"On-line signature verification with two-stage statistical models"* in Eighth International Conference on Document Analysis and Recognition, vol. 1, pp. 282-286, 2005.

[85] A. Ahrary*, et al.*, *"On-Line Signature Matching Based on Hilbert Scanning Patterns"* in *Advances in Biometrics*. vol. 5558, Springer, pp. 1190-1199, 2009.

[86] A. I. Al-Shoshan, *"Handwritten Signature Verification Using Image Invariants and Dynamic Features"* in International Conference on Computer Graphics, Imaging and Visualisation, pp. 173-176, 2006.

[87] R. Bajaj and S. Chaudhury, *"Signature verification using multiple neural classifiers"* in Pattern Recognition, vol. 30, pp. 1-7, 1997.

[88] A. Fallah*, et al.*, *"A new online signature verification system based on combining Mellin transform, MFCC and neural network", Digital Signal Processing,* vol. 21, pp. 404-416, 2011.

[89] W. Quen-Zong*, et al.*, *"On-line signature verification using LPC cepstrum and neural networks"* IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, vol. 27, pp. 148-153, 1997.

[90] J. Bromley*, et al.*, *"Signature Verification Using a Siamese Time Delay Neural Network"* in Neural Information Processing Systems, pp. 737—744, 1993.

[91] L. L. Lee, *"Neural approaches for human signature verification,"* in 3rd International Conference on Signal Processing, vol. 2, pp. 1346-1349, 1996.

[92] X. Xiao and G. Leedham, *"Signature verification using a modified Bayesian network"* in Pattern Recognition, vol. 35, pp. 983-995, 2002.

[93] C. Gruber*, et al.*, *"Signature Verification with Dynamic RBF Networks and Time Series Motifs"* in Tenth International Workshop on Frontiers in Handwriting Recognition, 2006.

[94] M. Tanaka*, et al.*, *"Determination of Decision Boundaries for Online Signature Verification"* in Knowledge-Based Intelligent Information and Engineering Systems, Springer, vol. 2773, pp. 401-407, 2003.

[95] M. Fuentes*, et al.*, *"On line signature verification: Fusion of a Hidden Markov Model and a neural network via a support vector machine"* in Eighth International Workshop on Frontiers in Handwriting Recognition, pp. 253-258, 2002.

[96] A. Mendaza-Ormaza*, et al.*, *"On-line Signature Biometrics using Support Vector Machine"* on Lecture Notes in Informatic BIOSIG 2009: Biometrics and Electronic Signatures, GI Edition, Darmstad, 2009.

[97] S. Fauziyah*, et al.*, *"Signature verification system using Support Vector Machine"* in 6th International Symposium on Mechatronics and its Applications, pp. 1-4, 2009.

[98] S. Emerich*, et al.*, *"On-line signature recognition approach based on wavelets and Support Vector Machines"* in IEEE International Conference on Automation Quality and Testing Robotics, pp. 1-4, 2010.

[99] C. Gruber*, et al.*, *"Online Signature Verification With Support Vector Machines Based on LCSS Kernel Functions"* in IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, vol. 40, pp. 1088-1100, 2010.

[100] S. Garcia-Salicetti*, et al.*, *"BIOMET: A Multimodal Person Authentication Database Including Face, Voice, Fingerprint, Hand and Signature Modalities"* in *Audio- and Video-Based Biometric Person Authentication, Springer,* vol. 2688, pp. 1056-1056, 2003.

[101] B. Dumas*, et al.*, *"MyIdea - Multimodal Biometrics Database, Description of Acquisition Protocols"* in Third COST 275 Workshop , Hatfield (UK), pp. 59-62, 2005.

[102] J. Ortega-Garcia*, et al.*, *"The Multiscenario Multienvironment BioSecure Multimodal Database (BMDB),"* in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 32, pp. 1097-1111, 2010.

[103] *Swiss National Center of Competence in Research on Interactive Multimodal, Information Management (IM2)*. Available: http://www.im2.ch

[104] *BioSecure Network of Excelence,* Available: http://biosecure.it-sudparis.eu/AB/.

[105] J. Fierrez-Aguilar*, et al.*, *"Fusion of Local and Regional Approaches for On-Line Signature Verification"* in Advances in Biometric Person Authentication, Springer, vol. 3781, pp. 188-196, 2005.

[106] C. Vivaracho-Pascual*, et al.*, *"Feature Selection in a Low Cost Signature Recognition System Based on Normalized Signatures and Fractional Distances"* in Advances in Biometrics, Springer, vol. 5558, pp. 1209-1218, 2009.

[107] *BioSecure Benchmarking Framework*, Biosecure. Available: http://share.int-evry.fr/svnview-eph/, 2009

[108] S. Schimke*, et al.*, *"Using Adapted Levenshtein Distance for On-Line Signature Authentication"* in 17th International Conference on Pattern Recognition, pp. 931-934, 2004.

[109] J. Fiérrez-aguilar*, et al.*, *"HMM-based on-line signature verification: Feature extraction and signature modeling"* in Pattern Recognition Letters, vol. 28, pp. 2325-2334, 2007.

[110] D. A. Reynolds and R. C. Rose, *"Robust text-independent speaker identification using Gaussian mixture speaker models"* in IEEE Transactions on Speech and Audio Processing, , vol. 3, pp. 72-83, 1995.

[111] L. Rabiner and B.-H. Juang, *"Fundamentals of Speech Recognition",* Prentice Hall, 1993.

[112] *BioSecure Signature Evaluation Campaign*, BSEC'2009. Available: http://biometrics.it-sudparis.eu/BSEC2009/

[113] S. Garcia-Salicetti*, et al.*, *"A Novel Personal Entropy Measure confronted with Online Signature Verification Systems' Performance" EURASIP Journal on Advances in Signal Processing,* 2008.

[114] ISO/IEC IS 19794-7:2007, *"Information Technology - Biometric data interchange formats – Part 7: Signature/sign time series data"* ISO/IEC, Geneva, 2007.

[115] ISO/IEC WD2 19794-7, *"2nd Working Draft 19794-7 Information Technology - Biometric data interchange formats – Part 7: Signature/sign time series data"*, ISO/IEC, Geneva, 2010.

[116] ISO/IEC CD2 29109-7, "2nd CD Information technology -- Conformance testing methodology for biometric data interchange formats defined in ISO/IEC 19794 - Part 7: Signature/sign time series data," ISO/IEC, Geneva, 2010.

[117] O. Miguel-Hurtado*, et al.*, *"Analysis on compact data formats for the performance of handwritten signature biometrics"* in 43rd Annual 2009 International Carnahan Conference on Security Technology,pp. 339-346, 2009.

[118] ISO/IEC IS 19785-3:2007, *"Information technology - Common Biometric Exchange Formats Framework - Part 3: Patron format specifications"* ISO/IEC, Geneva, 2007.

[119] ISO/IEC CD2 19794-11, *"2nd CD Information Technology - Biometric data interchange formats - Part 11: Signature/Sign Processed Dynamic",* ISO/IEC, Geneva, 2009.

[120] R. Sanchez-Reillo, *"Hand geometry pattern recognition through Gaussian mixture modelling"* in 15th International Conference on Pattern Recognition, vol. 2, pp. 937-940, 2000.

[121] A. P. Dempster*, et al.*, *"Maximum Likelihood from Incomplete Data via the EM Algorithm"* in Journal of the Royal Statistical Society. Series B (Methodological), vol. 39, pp. 1-38, 1977.

[122] J. Ortega Garcia, *"Tecnicas de mejora de voz aplicadas a sistemas de reconocimiento de locutores,"* PhD., Señales, Sistemas y Radiocomunicaciones, Universidad Politecnica de Madrid, Madrid, 1995.

[123] N. Kambhatla, *"Local Models and Gaussian Mixture Models for Statistical Data Processing",* Institute of Science & Technology, Oregon, 1966.

[124] G. K. Gupta and R. C. Joyce, *"A Study of Some Pen Motion Features in Dynamic Handwritten Signature Verification"* Department of Computer Science, James Cook University, Townsville, Australia, 1997.

[125] L. L. Luan, *"Reliable On-Line Human Signature Verification Systems"* in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 18, pp. 643-647, 1996.

[126] H. Ketabdar*, et al.*, "*Global Feature Selection for On-line Signature Verification"* in International Graphonomics Society 2005 Conference, pp. 625—629, 2005.

[127] H. Sakoe and S. Chiba, *"Dynamic programming algorithm optimization for spoken word recognition"* in IEEE Transactions on Acoustics, Speech and Signal Processing, , vol. 26, pp. 43-49, 1978.

[128] D.-Y. Yeung*, et al.*, *"SVC2004: First International Signature Verification Competition"* in Biometric Authentication, Springer, vol. 3072, pp. 1-30, 2004.

[129] G. K. Gupta and R. C. Joyce, *"Using position extrema points to capture shape in on-line handwritten signature verification"* in Pattern Recognition, vol. 40, pp. 2811-2817, 2007.

[130] L. Nanni and A. Lumini, *"A novel local on-line signature verification system"* in Pattern Recognition Letters, vol. 29, pp. 559-568, 2008.

[131] F. K. Soong and A. E. Rosenberg, *"On the use of instantaneous and transitional spectral information in speaker recognition"* in IEEE Transactions on Acoustics, Speech and Signal Processing, , vol. 36, pp. 871-879, 1988.

[132] Richiardi, *"Local and Global Feature Selection for On-line Signature Verification"* in International Conference on Document Analysis and Recognition, pp. 625-629, 2005.

[133] B. Li*, et al.*, "Online signature verification based on null component analysis and principal component analysis," in Pattern Analysis & Applications, vol. 8, pp. 345-356, 2006.

[134] K. V. Mardia*, et al.*, "*Multivariate Analysis",* Academic Press, 1992.

[135] M. Hazewinkel, "*Encyclopaedia of Mathematics",* vol. 1, Springer-Verlag, 2002

[136] R. O. Duda*, et al.*, "*Pattern Classification"* John Wiley & Sons, New York, 2001.
[137] H. Lei and V. Govindaraju, *"A comparative study on the consistency of features in on-line signature verification"* in Pattern Recognition Letters, vol. 26, pp. 2483-2489, 2005.
[138] A. K. Jain*, et al.*, "*Handbook of Biometrics",* Springer-Verlag New York, 2007.
[139] Mathlab, "*Interpolation ToolBox*"
[140] C. de Boor, "*A Practical Guide to Splines"*,Applied Mathematical Sciences, Springer, 1978.
[141] *7-Zip*. Available: http://www.7-zip.org/ , 2010
[142] *GZip*. Available: http://www.gzip.org/ , 2010
[143] M. Dipperstein, *Lempel-Ziv-Welch (LZW) Encoding Discussion and Implementation,* Available: http://michael.dipperstein.com/lzw/ , 2010.
[144] BZip, BZip2 Implementation v1.0.5, . Available: http://www.bzip.org/, 2008.
[145] *RFC 1952, GZIP file format specification version 4.3*. Available: http://tools.ietf.org/html/rfc1952 , 1996.
[146] ISO/IEC WD3 19794-7, *"3nd Working Draft 19794-7 Information Technology - Biometric data interchange formats – Part 7: Signature/sign time series data"* ISO/IEC, Geneva, 2011.
[147] ISO/IEC FDIS 19794-11, *"FDIS 19794-11 Information Technology - Biometric data interchange formats – Part 11: Signature/Sign Processed Dynamic"* ISO/IEC, Geneva, 2011.