Universidad de Valladolid

Escuela Técnica Superior de Ingeniería Informática

Departamento de Informática

Tesis doctoral:

Uso de la Firma Manuscrita Dinámica para el Reconocimiento Biométrico de Personas en Escenarios Prácticos

Presentada por Juan Manuel Pascual Gaspar para optar al grado de doctor por la Universidad de Valladolid

Dirigida por:

Dr. Valentín Cardeñoso Payo

Dr. Marcos Faúndez Zanuy

A Ana, mi esposa.

«La mayoría de las ideas fundamentales
de la ciencia son esencialmente sencillas y,
por regla general pueden ser expresadas
en un lenguaje comprensible para todos.»

ALBERT EINSTEIN

Agradecimientos

Las primeras palabras de agradecimiento quiero dedicarlas a mis directores de tesis, Valentín Cardeñoso y Marcos Faúndez sin los que este trabajo habría sido imposible, en todos los aspectos.

A los miembros del grupo ECA-SIMM, en concreto a César González que me ha ayudado mucho en los asuntos prácticos de la tesis, a David Escudero que revisó el proyecto de esta tesis y que siempre me ha dado ánimos, y especialmente a Carlos Vivaracho por sus consejos y apoyo durante los numerosos cafés que hemos tomado estos años.

A Julián Fiérrez por la revisión del proyecto de tesis y por mostrarme personalmente su valoración del trabajo realizado en las ocasiones en las que hemos coincidido.

A Bernardette Dorizzi y Sonia Garcia Salicetti por la disponibilidad gratuita de la base de datos BIOMET y a Andreas Humm por la disponibilidad de la base de datos MyIDea.

A Héctor Olmedo, por su amistad y por las conversaciones entre doctorandos que tan bien vienen para poder seguir adelante.

A mis antiguos compañeros de la Universidad Europea Miguel de Cervantes, en especial a Francisco Palacios y Roberto Méndez, cuyos consejos en los primeros años de tesis fueron muy importantes para mí.

A Jose Angel Romero, por estar siempre ahí, aunque por la distancia nos veamos menos de lo que nos gustaría.

A mi madre, porque siempre ha entendido y estimulado mi dedicación al estudio y a mis hermanos, Francisco, Carlos y Alejandro, por su cariño.

A Pedro y Nani por su paciencia y confianza. A Toñi, además, por la concienzuda e inestimable revisión del texto. A Javier y a Bull's Travels por su excelente eficacia. A los jóvenes científicos Carlos y Leen por sus constructivos comentarios e ideas para el futuro.

Finalmente a Ana, mi esposa, que es quien más ha 'sufrido' el esfuerzo y dedicación que ha supuesto la realización de esta tesis y a la que debo miles de horas de ocio de compensación los próximos años.

Resumen

El tema central de esta tesis es el estudio de la firma manuscrita dinámica para el reconocimiento biométrico de personas en escenarios prácticos. Se proponen dos sistemas originales de Verificación Automática de Firma dinámica diseñados, implementados y evaluados para ser implantados en escenarios reales. Cada uno de ellos se basa en una de las dos principales estrategias de clasificación de firma dinámica: los modelos ocultos de Markov (Hidden Markov Models, HMM) y el alineamiento temporal dinámico (Dynamic Time Warping, DTW). El trabajo se completa con la implantación de uno de estos sistemas en dos escenarios de carácter práctico: un concurso de imitaciones de firma, organizado durante las IV Jornadas de Reconocimiento Biométrico de Personas, y la participación en una competición internacional de verificación de firma dinámica (BSEC 2009: BioSecure Signature Evaluation Campaign), en la que el sistema que presentamos aquí obtuvo, en promedio, la segunda menor tasa de error sobre las evaluaciones realizadas.

El sistema basado en HMMs utiliza una nueva estrategia de diseño de modelos estructurales dependientes del usuario, con la que hemos obtenido una reducción del error en un factor de hasta seis respecto a la solución clásica con estructura universal. Empleamos un conjunto de características más reducido de lo habitual, que ha sido seleccionado siguiendo el procedimiento descrito en la memoria.

En el sistema basado en DTW se incluye un procedimiento de selección de combinaciones óptimas de características dependiente del escenario de aplicación. Esta selección contribuyó a una reducción del error de hasta el 86 % en escenarios de imitación casual y del 40.8 % en escenarios de ataque adiestrado.

Los resultados del concurso de imitaciones mostraron que las cifras de error publicadas en escenarios de ataque adiestrado dependen de la calidad de la imitación, que puede ser mejorada permitiendo el acceso no supervisado del impostor a la dinámica de la firma atacada. El sistema preparado para la competición BSEC 2009 se basa también en el sistema DTW, pero incluye como aportación una nueva técnica de normalización de puntuaciones denominada *norma EER*, que permitió disminuir significativamente el incremento del error al pasar de umbral individual a universal.

Tanto los sistemas propuestos como las experiencias realizadas demuestran que, al igual que otras modalidades biométricas más maduras tecnológicamente, las técnicas de reconocimiento biométrico basadas en firma manuscrita dinámica están en condiciones de dar el salto de las condiciones controladas de laboratorio a los entornos de aplicación reales. Los resultados que se contienen en esta tesis representan, en resumen, un paso adelante en la consolidación de la firma manuscrita dinámica como modalidad biométrica competitiva en escenarios prácticos.

Abstract

The main subject of this thesis is the usage of on-line handwritten signatures for biometric recognition in practical scenarios.

We propose two automatic Handwriting Signature Verification (HSV) systems designed, implemented and evaluated to be deployed in real scenarios. Each system is based on one of the two main classification approaches for HSV: Hidden Markov models (HMM) and Dynamic Time Warping (DTW). The work is completed with two practical experiences. The first one is a forgeries competition at the *IV Jornadas de Reconocimiento Biométrico de Personas* conference, where a HSV website was attacked under uncontrolled conditions. The second experience is the participation in the 2009 BioSecure Signature Evaluation Campaign (BSEC 2009), where our system got the second lowest average error rate over all the assessments.

The HMM system uses a new user-dependent structural proposal with which we obtained a error reduction factor of six times related to the obtained with the universal structure classical solution approach. Based on this new proposal we tuned up the system at features selection and preprocessing levels obtaining very competitive performance with less features than typical HMM systems. The DTW system was optimized by selecting an optimal feature set adapted to the application scenario which provided us a error reduction of 86 % in the casual scenario and 40.8 % in the skilled scenario. The results of the forgeries competition showed that the figures published in skilled scenarios depend on the quality of the imitation, which can be improved by allowing unsupervised access to the signature under attack. Regarding the system submitted to BSEC 2009, it was based on the DTW system, but was modified to include a new technique for score normalization called *EER norm* which allowed us to minimize significantly the error increase shifting from user-dependent to user-independent threshold.

The proposed systems and experiments show that the transition from laboratory to real application environment can be successfully carried out and set the handwritten signature at the same level as other more technologically mature biometric modalities. These results represent a step in the consolidation of on-line handwritten signature as a competitive biometric modality in practical scenarios.

Índice general

Lista de Figuras

Lista de Tablas

1 Introducción

La creciente demanda de acceso a los servicios de la Sociedad de la Información ha dado lugar en las últimas décadas a la aparición de una nueva rama de la Tecnología denominada *Autenticación biométrica* o simplemente *Biometría* (Jain *et al.*, 2004a). Un *sistema biométrico* podría definirse como *'un sistema automático que permite el reconocimiento de seres vivos a través de sus rasgos inherentes'*. Existen dos tipos de rasgos biométricos mediante los que poder realizar la autenticación biométrica:

- **Rasgos fisiológicos**: son aquellos que corresponden a características diferenciadoras del cuerpo humano de índole principalmente fisiológica. La huella dactilar, el rostro o el iris son ejemplos de este tipo de rasgos.
- **Rasgos conductuales**: son rasgos que están más relacionados con el comportamiento de la persona. A esta categoría pertenecerían por ejemplo la firma, el habla o la escritura manuscrita.

La tesis que a continuación se va a presentar trata del reconocimiento biométrico de personas mediante la modalidad dinámica de la firma manuscrita.

1.1 Motivación

Actualmente los sistemas biométricos que más han avanzado desde el punto de vista tecnológico son los basados en las modalidades de tipo fisiológico (huella, iris, cara, geometría de la mano, etc.) y ya es habitual encontrar en el mercado dispositivos basados en ellos (cámaras de vídeo con reconocimiento de cara, acceso a ordenadores portátiles mediante la huella dactilar, sistemas de identificación de iris en aeropuertos, etc.).

Por otro lado las modalidades conductuales tales como la voz, la escritura y la firma manuscrita o incluso el modo de caminar siguen siendo motivo de investigación sin que dicho esfuerzo se vea reflejado en aplicaciones prácticas para los ciudadanos.

En el caso concreto de la firma manuscrita esta falta de aplicación práctica se debe principalmente a aspectos sociales y legales aunque también a la falta de bancos de prueba para comparar y evaluar sistemas de forma objetiva. Dichos bancos de prueba son creados normalmente a partir de competiciones internacionales en las cuales se definen metodologías con las que evaluar y comparar objetivamente las soluciones de distintos autores, siendo

habitual liberar al final de la competición las bases de datos utilizadas para promover la investigación. En otras modalidades biométricas tales como huella dactilar (Cappelli *et al.*, 2006), voz (Przybocki & Martin, 2004; Przybocki *et al.*, 2006), cara (P. J. Phillips & Rizvi, 2006) e iris (Phillips, 2006a) estas competiciones se realizan regularmente, pero en el caso de firma se limita a dos (Yeung *et al.*, 2004; Dorizzi *et al.*, 2009).

Al implantar un sistema biométrico en el mercado deben tenerse en cuenta otros aspectos prácticos tales como el rendimiento computacional, el grado de esfuerzo del usuario en la inscripción o la seguridad del propio sistema. En el caso concreto de la firma, algunos de los aspectos prácticos a tener en cuenta son:

1. Uso de un número reducido de muestras de referencia.
2. Reducir el tamaño de almacenamiento de las firmas.
3. La eficiencia computacional.
4. La seguridad, buscando sobre todo robustez ante ataques con imitaciones.
5. Capacidad de adaptación a los distintos tipos de firmantes y a sus variaciones a la hora de firmar, a corto y largo plazo.
6. Evitar emplear imitaciones en la creación del modelo.
7. No basar el éxito en características muy específicas del hardware de adquisición para ganar en universalidad.

La presente tesis está principalmente dirigida a la aplicabilidad de los sistemas de firma en escenarios prácticos. Se han detectado los problemas que surgen al diseñar una solución biométrica basada en firma y se han propuesto algunas soluciones para hacerla tecnológicamente viable.

Como se verá a lo largo de esta memoria, el trabajo experimental se ha llevado a cabo bajo las siguientes condiciones:

1. Hemos usado siempre un número reducido de firmas para el registro del usuario.
2. Se ha evaluado el rendimiento obtenido con distintas combinaciones de características de la firma buscando la mejor relación rendimiento-coste de almacenamiento.
3. Las características extraídas de la firma son fáciles de calcular.
4. Se ha partido de los algoritmos de probada mayor eficacia en el campo de verificación de firma dinámica, tanto basados en modelos como en almacenamiento de plantillas.
5. Para dotar de mayor validez a los resultados se han empleado varias bases de datos de firma, creadas por distintos grupos de investigación, sumando entre todas ellas más de 500 usuarios.
6. Todas los sistemas son creados bajo condiciones realistas, sin utilizar imitaciones en su construcción.

7. El dispositivo utilizado para la captura de las firmas ha sido una tableta gráfica debido a su fácil acceso, precio asequible y buena resolución espacio-temporal durante la captura. También se han realizado pruebas eventuales con dispositivos de adquisición móvil (PDA).

A continuación enunciaremos formalmente la tesis de esta investigación describiendo sus términos fundamentales. Además, se enumeran los objetivos del trabajo, la justificación del mismo y un breve resumen de las contribuciones realizadas.

1.2 La tesis

La hipótesis de partida sobre la que se fundamenta la presente investigación es el hecho contrastado de que la firma manuscrita dinámica es un medio aceptable para el reconocimiento automático de personas, de calidad contrastada en escenarios experimentales en laboratorio y además es una técnica idónea para ser empleada en escenarios prácticos.

Desde finales de la década de los 80 han sido numerosas las publicaciones en las que se han propuesto sistemas para el reconocimiento biométrico de personas a partir de la firma manuscrita, tanto en modalidad estática como dinámica (Plamondon & Lorette, 1989; Leclerc & Plamondon, 1994; Gupta & McCabe, 1997; Nalwa, 1997; Plamondon & Srihari, 2000; Kalenova, 2003; Dimauro *et al.*, 2004; Fierrez & Ortega-Garcia, 2007; Impedovo & Pirlo, 2007). Es ésta última modalidad la que permite obtener mejores tasas de reconocimiento al disponer no sólo de la realización final de la firma (información estática) sino también de la información sobre el proceso de ejecución (información dinámica).

Además, los resultados obtenidos en competiciones internacionales (Yeung *et al.*, 2004; Dorizzi *et al.*, 2009), informes de mercado (International Biometric Group, 2007), fabricantes de hardware (Wacom, 2009; Electronics, 2007; Genius, 2007), (Dynalink, 2007) y software comercial (Cyber-SIGN, 2007; DataVision, 2007; PenOp, 2007), sirven para fundamentar nuestra hipótesis de partida.

Partiendo de la hipótesis inicial, la tesis puede enunciarse del modo siguiente:

> *Es posible diseñar sistemas de reconocimiento biométrico de personas basados en la modalidad dinámica de la firma manuscrita que proporcionen un rendimiento competitivo en escenarios de aplicación práctica.*

En otras palabras, dado que ya existen algoritmos para el reconocimiento biométrico de personas basados en la firma dinámica de reconocida eficacia en laboratorio, lo que trataremos de demostrar es que se puede mantener este rendimiento en escenarios prácticos a niveles tecnológicamente competitivos respecto a otras modalidades biométricas.

El *escenario de aplicación práctica* viene determinado por los aspectos prácticos y condiciones de trabajo expuestas en la sección 1.1. El *rendimiento competitivo* se determinará comparando los resultados obtenidos en escenarios experimentales con los de otras modalidades biométricas.

1.3 Objetivos

El *objetivo principal* es demostrar que es posible utilizar eficazmente la firma manuscrita dinámica en escenarios prácticos. Para lograrlo nos marcamos los siguientes objetivos más concretos encaminados a demostrar conjuntamente el principal:

1. Establecer las condiciones de trabajo de los sistemas de firma manuscrita dinámica en escenarios prácticos.
2. Emplear una metodología para la evaluación objetiva de sistemas de reconocimiento biométrico basados en firma dinámica que permita disponer de datos fiables para el análisis comparativo de los mismos.
3. Seleccionar un conjunto de sistemas de referencia en reconocimiento biométrico de firma dinámica basado en su eficacia en escenarios experimentales.
4. Elegir el conjunto de características de la firma a utilizar en escenarios prácticos para que los sistemas sean sencillos y computacionalmente eficientes sin perder su capacidad de reconocimiento.
5. Evaluar el rendimiento de un algoritmo de referencia de cada una de las dos principales aproximaciones metodológicas (basadas en distancias y basadas en modelos estadísticos) y comprobar su eficiencia individual en escenarios prácticos.
6. Comprobar la influencia de la calidad de las falsificaciones en el rendimiento del sistema.
7. Analizar el rendimiento del sistema en función del escenario de prueba desde el punto de vista de la seguridad (acceso con firma casual o con imitación).
8. Determinar los beneficios y el coste que supone la personalización al usuario de los sistemas de reconocimiento desarrollados.
9. Aplicar los sistemas desarrollados en escenarios prácticos tales como competiciones o sistemas tecnológicamente viables.

1.4 Justificación

Los resultados obtenidos por otros autores en el campo del reconocimiento biométrico de la firma manuscrita dinámica fundamentan la base sobre la que se ha sustentado la viabilidad de este trabajo de tesis. La ausencia de bases de datos de referencia y protocolos de evaluación estándar, fundamentalmente antes del año 2003, hace difícil distinguir las mejores propuestas en base únicamente a sus tasas de error. Examinando las publicaciones existentes desde principios de los años 80 y basándonos en el número de propuestas existentes, dos son los métodos más empleados para abordar este problema, ambos tomados de la experiencia con otras modalidades biométricas, especialmente la voz.

En primer lugar, los modelos ocultos de Markov (HMM) representan un enfoque al problema mediante técnicas estadísticas que han tenido mucho éxito para modelar secuencias

temporales de patrones, sobre todo secuencias de voz. Sin embargo, esta técnica necesita usualmente de un gran número de muestras de entrenamiento para modelar correctamente la variabilidad intraclase del patrón, lo que es un hándicap para la firma en escenarios prácticos. Otros problemas de la técnica son la necesidad de un número elevado de características y la falta de métodos sistemáticos para la selección de la estructura óptima del modelo.

De todas formas, y tal y como se verá en detalle en la sección 4.2, trabajos recientes (Fierrez *et al.*, 2007) demuestran que mediante una adecuada selección de características y de la estructura del modelo, pueden obtenerse muy buenos resultados, tal y como quedó demostrado en la competición SVC 2004 (Yeung *et al.*, 2004). Dado que la firma es una modalidad biométrica que presenta una gran variabilidad entre usuarios vimos, desde el principio de esta tesis, que la adaptabilidad al usuario de la estructura del HMM podría reportar ventajas significativas en el rendimiento.

La otra técnica que se ha empleado con éxito en este campo es el alineamiento temporal dinámico (DTW). Antes de la aparición de HMM, fue la que mejores resultados proporcionaba en el campo del reconocimiento de voz, aunque la escasa disponibilidad de muestras de referencia la mantiene entre las más empleadas en el campo de la firma. De hecho, después de los resultados de SVC 2004 y BSEC 2009 (Dorizzi *et al.*, 2009) podría considerarse como la técnica de referencia a nivel del estado del arte. Trabajos como (Kholmatov & Yanikoglu, 2005) y sus resultados en SVC 2004 demuestran que es posible obtener con ella un rendimiento igual o mejor que con HMM en condiciones prácticas.

Esta técnica presenta inconvenientes como son la necesidad de almacenar las firmas de referencia, la falta de estudios sistemáticos sobre la selección de características óptimas en cada tipo de escenario de aplicación o la dificultad de actualizar los patrones de referencia de forma progresiva para incorporar la variabilidad del usuario con el paso del tiempo. En nuestro caso, la posibilidad de mejorar el rendimiento mediante un estudio de las características adaptadas al escenario de aplicación, y que tenga en cuenta las interdependencias entre características, nos indujo a la construcción de un sistema basado en esta técnica.

1.5 Contribuciones

A continuación presentamos algunas de las contribuciones más relevantes realizadas en el marco de la presente tesis:

Revisiones de la literatura. Se han revisado en detalle y desde un punto de vista crítico los trabajos de verificación de firma basados en HMM más relevantes de la literatura (sección 4.2). También se han sintetizado los resultados obtenidos por sistemas basados en otras técnicas de autores que utilizaron las mismas bases de datos que las utilizadas en esta tesis (sección 5.5).

Bancos de prueba. Hemos evaluado de forma sistemática los sistemas presentados con un protocolo experimental basado en el de la competición SVC 2004. Se ha medido el rendimiento en condiciones realistas en escenario casual y de ataque, ambos representativos de valores extremos de funcionamiento.

Se ha realizado un breve estudio estadístico del corpus MCYT extrayendo conclusiones sobre los principales parámetros globales de sus firmas.

En el marco de la competición BSEC 2009 (sección 6.2) se ha analizado el rendimiento de uno de los sistemas presentados (sección 5.2) al usar firmas obtenidas en condiciones móviles (PDA). Este mismo sistema ha sido evaluado sobre cuatro bases de datos de referencia obteniendo un rendimiento competitivo en todas ellas. Destacamos que no es habitual evaluar un mismo sistema con varias bases de datos.

Análisis de características. Hemos realizado un análisis de la eficiencia de características locales de la firma dinámica en los dominios de la posición, velocidad y aceleración. Este estudio presenta la novedad de evaluar las características no sólo desde un punto de vista individual, sino también al combinarlas entre sí. Aparte de la selección de combinaciones de características óptimas y universales, hemos proporcionado cotas teóricas de rendimiento utilizando combinaciones de características dependientes del usuario que indican que sistemas construidos con este método podrían mejorar significativamente el rendimiento actual de los sistemas de VAFD.

Sistemas de verificación. Hemos diseñado, construido y evaluado dos sistemas de verificación automática de firma dinámica basados en las dos principales orientaciones metodológicas en el dominio de aplicación. En el capítulo 4 hemos descrito un nuevo sistema basado en modelos ocultos de Markov con la novedad de poseer estructura dependiente del usuario. En el capítulo 5 hemos presentado otro sistema basado en plantillas y alineamiento temporal dinámico con el que se ha obtenido un excelente rendimiento bajo condiciones realistas.

Adaptación al usuario. Hemos introducido el nuevo concepto de *modelos ocultos de Markov dependientes del usuario* (HMM-DU) como alternativa al enfoque más tradicional basado en HMM con estructura universal o independiente del usuario (HMM-IU). Hemos medido las cotas de error obtenidas con ambos métodos bajo las mismas condiciones experimentales. Las conclusiones de este estudio han servido como base para la creación de un sistema más óptimo, basado en HMM-DU, cuya estructura es obtenida de los propios datos de entrenamiento.

Nueva normalización de puntuaciones. Hemos introducido un nuevo método de normalización de puntuaciones centradas en impostor y cliente que hemos denominado *norma EER* (sección 6.2.5). Dicha norma ha sido empleada con éxito para la competición BSEC 2009.

Aplicaciones. Durante las IV Jornadas de Reconocimiento Biométrico de Personas celebradas en Valladolid en septiembre de 2008 y en el marco del presente trabajo de tesis se organizó un concurso de imitaciones de firma para evaluar la resistencia ante ataques de un sistema de Verificación Automática de Firma (VAFD) en condiciones muy adversas. De él se han extraído varias conclusiones de cara al establecimiento de una medida de calidad de la firma para futuros trabajos.

El sistema basado en DTW del capítulo 5 ha sido utilizado en la aplicación experimental FirmWeb (sección 6.1.2). Dicha aplicación ha sido creada para investigación sobre acceso a servicios web a través de rasgos biométricos utilizando múltiples dispositivos (tableta gráfica, tablet PC, etc.)

Participación en evaluaciones internacionales. El sistema descrito en el capítulo 5 fue presentado en la competición internacional BSEC 2009 obteniendo un resultado destacable (sección 6.2) en la mayoría de las tareas en las que participó.

1.6 Esquema de la memoria

Esta memoria de tesis se ajusta a un esquema clásico, donde el marco teórico general del problema va seguido de la descripción crítica de los distintos sistemas y experiencias desarrollados en el trabajo.

En el capítulo 2 presentamos el marco teórico general del problema del reconocimiento biométrico de personas basado en la modalidad dinámica de la firma manuscrita.

Seguidamente, el capítulo 3 sirve para describir los bancos de prueba existentes para determinar el nivel del estado tecnológico de la firma dinámica como modalidad biométrica. Dentro de estos bancos de prueba distinguimos las bases de datos públicas, que sirven de marco de datos de referencia para muchos trabajos científicos, y de las cuales describimos cuatro de las más populares, todas utilizadas en esta tesis. En segundo lugar, hablamos de las competiciones internacionales, que sirven para evaluar sistemas de forma objetiva cada cierto tiempo. En el caso de la firma manuscrita describiremos en los capítulos 3 y 6 los dos precedentes existentes hasta la fecha actual.

El primer sistema desarrollado en esta tesis se describe en el capítulo 4. Se trata de un sistema de tipo estadístico basado en los modelos ocultos de Markov. En él se introduce el concepto de HMM dependiente de usuario junto a un sistema basado en este nuevo concepto.

El segundo sistema se describe en el capítulo 5. Se basa en el almacenamiento de firmas del cliente que se usan como plantillas contra las que se comparan nuevas firmas mediante el algoritmo DTW. Basado en él se llevaron a cabo las experiencias descritas en el siguiente capítulo.

Las experiencias de carácter práctico realizadas a lo largo de la tesis se han incluido en el capítulo 6. En primer lugar, se organizó una competición de imitaciones de firmas para evaluar el sistema del capítulo 5 en escenario de ataque. Seguidamente, describimos tanto los resultados oficiales de BSEC 2009 como el sistema con el que participamos, que aunque básicamente es el sistema descrito en el capítulo 5, fue necesario ampliar debido a las reglas de la competición.

El último capítulo (7) lo dedicamos a exponer las conclusiones del trabajo realizado a lo largo de esta tesis y a proponer nuevas líneas de investigación para el futuro.

Como apéndices hemos incluido un glosario con términos biométricos y acrónimos utilizados (apéndice A) y un resumen de las herramientas desarrolladas para llevar a cabo la investigación (apéndice B).

2

Reconocimiento de firma dinámica

En este capítulo expondremos los conceptos fundamentales del problema del reconocimiento biométrico basado en firma. Con ello se pretende dar a conocer de forma resumida los términos necesarios para la lectura del resto de esta memoria. Para profundizar más se recomienda la lectura de las publicaciones (Faundez-Zanuy, 2007; Fierrez & Ortega-Garcia, 2007).

En primer lugar, se describen los tipos de firma y tareas con las se puede realizar reconocimiento biométrico basado en firma. Tras ello, se enumeran los principales módulos que componen un sistema biométrico general para posteriormente pasar a analizar en más detalle las fases de extracción y tratamiento de las características, así como los métodos de clasificación más utilizados en el dominio concreto de la firma.

2.1 Tipos de firma

Dependiendo del modo de adquisición se pueden distinguir dos modalidades de firma, *estática* y *dinámica*.

- **Estática**: esta modalidad corresponde a la digitalización de una firma manuscrita a partir de una muestra obtenida en papel. Comúnmente a esta modalidad se la denomina también mediante su acepción inglesa (*off-line*).
- **Dinámica**: este tipo de firma es capturada usando dispositivos especiales con capacidad de registrar la evolución temporal de varias señales generadas por el lápiz al firmar. Además de las coordenadas posicionales, algunos de estos dispositivos proporcionan características adicionales como por ejemplo, la presión ejercida sobre el plano de escritura y los ángulos formados entre el lápiz y dicha superficie de escritura. El término inglés para esta modalidad de firma es *on-line*.

En esta tesis se ha seleccionado la modalidad dinámica por ser la más viable para ser aplicada en la práctica y por proporcionar un rendimiento aceptable.

2.2 Tareas biométricas

En Biometría podemos distinguir dos tipos de tareas con fines distintos: Identificación y Verificación Biométrica. Ambas se agrupan en el término más general *Reconocimiento*

Biométrico:

- La *Identificación Biométrica* es el proceso por el que se trata de determinar quién es un individuo, comparando sus características biométricas con las almacenadas en una base de datos. Se trata de una comparación de uno a muchos (1:N).
- La *Verificación Biométrica* busca comprobar si un sujeto que intenta acceder al sistema es quien dice ser. Para ello, el sistema comparará los datos biométricos del usuario con los almacenados previamente para ese usuario, tras lo cual decidirá si le permite o deniega el acceso. Este tipo de tarea es una comparación de tipo uno a uno (1:1).

Dado que la aplicación más frecuente y de interés práctico de la firma en escenarios cotidianos es la *Verificación*, esta tesis se ha centrado en esta tarea biométrica.

2.3 Evaluación de sistemas biométricos

Para evaluar la eficacia de un sistema biométrico es necesaria una medida que determine el rendimiento en cada una de las tareas anteriores. En el caso de la identificación biométrica es común utilizar como medida de rendimiento el porcentaje de aciertos del sistema respecto al total de identificaciones realizadas. Al evaluar esta tarea pueden seguirse dos alternativas: o bien presuponer que el sujeto a identificar existe en la base de datos (*identificación cerrada*), o bien realizar una verificación encubierta posterior a la identificación (*identificación abierta*).

Para evaluar la verificación biométrica es necesario contar con datos de dos tipos de usuarios: a) el *cliente* o sujeto que proporciona las muestras biométricas genuinas y b) el *impostor* que intenta suplantar la identidad del cliente. Existen dos tipos de impostor en función del esfuerzo o intención realizados al producir las falsificaciones:

- **Impostor casual**: es el que produce una falsificación que no intenta intencionalmente imitar el rasgo biométrico del cliente. Estas falsificaciones se denominan *falsificaciones casuales*.
- **Imitador**: es el sujeto que produce una falsificación que imita de modo intencional el rasgo biométrico del cliente. Estas falsificaciones las denominaremos simplemente *imitaciones*[1].

A partir de los conceptos de cliente e impostor se definen típicamente dos tipos de error para evaluar el rendimiento de los sistemas biométricos en modo verificación. En primer lugar, se define la *Tasa de Falsos Rechazos* (FRR, False Rejection Rate) como el porcentaje de muestras de cliente rechazadas por el sistema. En segundo lugar, la *Tasa de Falsas Aceptaciones* (FAR, False Acceptance Rate) mide el porcentaje de muestras de impostor aceptadas.

[1]Aunque una traducción literal del término inglés (*skilled forgeries*) sería 'falsificaciones entrenadas', utilizaremos el término 'imitación' que en castellano ya sugiere por si mismo esfuerzo al realizar una falsificación

El valor concreto de estos dos tipos de error depende del valor de corte o *umbral de referencia* usado para distinguir entre muestras de cliente y de impostor. Una medida muy extendida para evaluar el rendimiento de los sistemas biométricos en una única cifra es la Tasa de equierror (Equal Error Rate, EER) que se obtiene cuando FAR y FRR coinciden. Se puede representar el comportamiento gráficamente para todos los umbrales mediante las curvas ROC (Receiver Operation Curve) y DET (Detection Error Tradeoff) (Martin *et al.*, 1997).

2.4 Arquitectura de un sistema biométrico

La figura 2.1 muestra los principales módulos de un sistema biométrico desde un punto de vista general (Ross *et al.*, 2006):

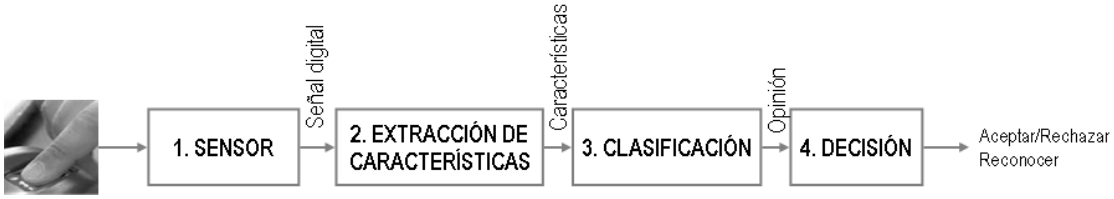


FIGURA 2.1: Módulos de un sistema biométrico

1. **Sensor**: es el dispositivo de captura de los datos biométricos. Las firmas de las bases de datos utilizadas en este trabajo de tesis han sido obtenidas con una tableta gráfica WACOM la cual lleva consigo un lápiz especial con el que se recoge la información espacial y temporal ejercida por el lápiz al moverlo sobre la tableta.

2. **Extracción de características**: usualmente es necesario realizar un preprocesamiento de las muestras biométricas antes de ser utilizadas. Mediante este preprocesamiento se acondicionan los datos del sensor y se generan los *vectores de características*. En el caso de la firma manuscrita estas características pueden ser de dos tipos (Plamondon & Lorette, 1989):

   - *Estáticas*: también llamadas *parámetros* o características *globales*, son aquellas que se extraen de la firma en su conjunto. Ejemplos de este tipo de características son los estadísticos básicos de posición, velocidad y aceleración del lápiz, número de trazos, relación de aspecto, etc.
   - *Dinámicas*: también llamadas *funciones* o características *locales*. Representan a la firma mediante una secuencia temporal de longitud variable de vectores de características. Algunos ejemplos pueden ser la posición, la velocidad y la aceleración instantánea del bolígrafo durante el proceso de firma.

3. **Comparador**: también llamado módulo de *clasificación*, mide la semejanza entre el patrón de entrada y la información previamente almacenada del cliente. Puede realizarse mediante diferentes técnicas de reconocimiento de patrones tales como

algoritmos basados en plantillas o basados en modelos de tipo estadístico. A los resultados obtenidos a la salida de este módulo se les denomina *puntuaciones* u *opiniones* (en inglés 'scores').[2]

4. **Decisión**: módulo en el que se genera la decisión final del sistema biométrico a partir de las puntuaciones de la etapa de clasificación. En el caso de identificación biométrica se realizan tantas comparaciones como patrones de cliente existente en la base de datos. En el caso de la verificación biométrica sólo se necesita una comparación, aunque es necesario un valor de corte o *umbral de decisión* con el que determinar si se acepta o rechaza el patrón de entrada.

2.5 Características

Dada la mayor eficacia de los algoritmos basados en características locales (funciones) frente a los basados en características globales (parámetros) (Leclerc & Plamondon, 1994) los sistemas desarrollados usan características locales.

Normalmente la selección de las características se hace en función de criterios heurísticos o de la propia experiencia del autor. Para que estos sistemas funcionen con suficiente precisión es necesario emplear un conjunto de características más apropiado que el proporcionado por defecto por el hardware de adquisición. En un sistema práctico de firma manuscrita es deseable que este número de características sea pequeño a fin de reducir los requerimientos de almacenamiento y el tiempo de cómputo.

Respecto al *acondicionamiento de las características* podemos encontrar dos tipos de normalización:

1. **Normalización geométrica**: se realiza para que los datos sean invariantes ante transformaciones geométricas (traslación, escalado y rotación).

2. **Normalización estadística**: aunque las características estén normalizadas respecto a transformaciones geométricas es habitual realizar una transformación adicional de tipo estadístico para que las distintas características tengan órdenes de magnitud similares al ser incluidas en los vectores de características.

2.6 Clasificación

El reconocimiento automático de firma dinámica suele abordarse mediante uno o varios algoritmos pertenecientes a los siguientes tipos de técnicas de clasificación de patrones (Tapiador Mateos & Sigüenza Pizarro, 2005):

1. **Métodos basados en alineamiento de características**: consistentes en la comparación entre la muestra de entrada y un prototipo de referencia almacenado (denominado *plantilla*). El método perteneciente a esta categoría más empleado para el reconocimiento de firma dinámica es el Alineamiento Temporal Dinámico (DTW, Dynamic Time Warping).

[2] A lo largo de este documento el autor usará el término *puntuación*.

2. **Métodos basados en modelos estadísticos**: con estos métodos los patrones de referencia son usados para construir un modelo probabilístico/estadístico. Los modelos ocultos de Markov (HMM, Hidden Markov Models) son el algoritmo de referencia de este tipo en el campo de reconocimiento de firma dinámica (Dimauro *et al.*, 2004).

3. **Métodos basados en fronteras de decisión**: estos algoritmos se basan en la creación de fronteras entre clases a partir de la optimización de un determinado criterio de error entre los resultados buscados y los obtenidos. A este tipo pertenecen las redes neuronales, los árboles de decisión y las máquinas de soporte vectorial.

Las dos primeras categorías anteriores presentan la ventaja respecto a la tercera de no necesitar muestras negativas para la generación de los prototipos de referencia. Este hecho tiene especial importancia en firma ya que, incluso por implicaciones de tipo legal, no deberían utilizarse imitaciones de firmas para la creación de los prototipos de referencia.

No obstante, y puesto que los mejores resultados obtenidos de modo objetivo hasta la fecha (Yeung *et al.*, 2004), han sido producidos con métodos basados en el alineamiento de características (Yanikoglu & Kholmatov, 2003; Kholmatov & Yanikoglu, 2005) y en métodos estadísticos (Fierrez-Aguilar *et al.*, 2005a), se han utilizado estos sistemas como referentes del estado del arte.

2.7 Métodos estadísticos

Mediante técnicas de modelado estadístico las firmas de un usuario son almacenadas en un modelo parametrizado de las mismas con valores específicos del usuario. Las dos técnicas más representativas de esta categoría para reconocimiento de firma son los Modelos de Mezclas de Gaussianas (GMM, Gaussian Mixture Models) (Titterington *et al.*, 1985) y los Modelos Ocultos de Markov (HMM, Hidden Markov Models) (Rabiner, 1989). No obstante, dado que la primera de ellas puede verse como un caso particular de la segunda, la mayoría de los sistemas de tipo estadístico están basados en HMM.

Los HMMs son una técnica probabilística ampliamente usada para modelar patrones dependientes del tiempo. Ha sido aplicada con éxito en reconocimiento de habla (Rabiner, 1989), escritura (Hu *et al.*, 1996) y verificación de firma dinámica (Yang *et al.*, 1995; Fierrez-Aguilar, 2006). Un modelo oculto de Markov puede ser descrito a grandes rasgos como un grafo de estados interconectados mediante una topología determinada por medio de una matriz de probabilidades de transición entre estados. La probabilidad de emisión de un estado es usualmente modelada como una superposición de distribuciones de gaussianas.

El número de gaussianas asociadas a cada estado constituyen los *parámetros estructurales* del modelo. Establecer estos parámetros estructurales es un problema altamente dependiente de la tarea a resolver. Usualmente es realizada por expertos en el dominio de aplicación o mediante prueba y error, ya que es difícil encontrar algoritmos generalistas con los que inferir los parámetros estructurales a partir de los datos a modelar.

Una vez establecida la estructura del modelo, éste es entrenado mediante el algoritmo clásico de Baum-Welch (Rabiner, 1989). Este algoritmo determina los valores de las probabilidades de transición, los pesos y los momentos estadísticos de las gaussianas que proporcionen la máxima verosimilitud para la secuencia de observaciones temporales de

entrenamiento. Estos valores representan lo que denominamos *parámetros estadísticos* del modelo.

2.8 Alineamiento de características

Las técnicas basadas en alineamiento de características miden la distorsión o deformación que es necesario realizar sobre una firma para alinearla con otra de referencia. En el caso de firmas dinámicas el alineamiento se realiza entre los puntos obtenidos al capturar la firma. Así pues, para alinear dos firmas de un mismo autor, se requerirá menos deformación que cuando se alinean dos firmas pertenecientes a personas distintas.

La técnica de alineamiento más utilizada en el reconocimiento de la firma dinámica es el alineamiento temporal dinámico (DTW, *Dynamic Time Warping*) (Rabiner *et al.*, 1978). El algoritmo DTW permite realizar un alineamiento óptimo entre dos secuencias de vectores de distinta longitud mediante programación dinámica. De dicho alineamiento se obtiene una medida de distancia entre los dos patrones temporales.

2.9 Resumen

En este capítulo hemos revisado los conceptos necesarios para la lectura de esta memoria de tesis. En el apéndice A el lector dispone además de un glosario con términos necesarios para la lectura del resto de esta memoria.

En el siguiente capítulo describiremos los bancos de prueba existentes en el campo del reconocimiento de la firma. Con ello obtendremos una visión del estado de madurez tecnológica de esta modalidad biométrica.

3
Bancos de prueba

Uno de los recursos que más hacen avanzar la tecnología biométrica son los *bancos de prueba*. Distinguimos dos tipos: a) bases de datos y b) competiciones o evaluaciones por terceros.

La disponibilidad pública de bases de datos biométricas de referencia permite a los investigadores evaluar sus propuestas con los mismos conjuntos de datos, aunque la falta de protocolos estándar de evaluación pueden provocar diferencias significativas entre sus resultados. En este sentido, las competiciones o evaluaciones objetivas realizadas por terceros establecen reglas y protocolos de evaluación que los autores adoptan para medir el rendimiento de sus sistemas. La combinación de ambos recursos permite con el tiempo avanzar el estado del arte a nivel científico y tecnológico.

En este capítulo presentamos cuatro bases de datos de firma dinámica sobre las que existen numerosos trabajos con los que poder comparar los sistemas que presentamos en esta tesis. En la segunda parte del capítulo mostramos los resultados obtenidos en la competición internacional de reconocimiento de firma SVC 2004. Los resultados obtenidos en ella, junto a los resultados de la segunda competición de este mismo tipo (BSEC 2009) descrita en el capítulo 6, dan una buena visión del estado del arte en el área.

3.1 Bases de datos

En los distintos trabajos realizados en el marco de la tesis se han empleado hasta cuatro bases de datos de firma dinámica:

- MCYT (Ortega-Garcia *et al.*, 2003b)
- BIOMET (Garcia-Salicetti *et al.*, 2003)
- MyIDEA (Dumas *et al.*, 2005)
- SVC 2004 (Yeung *et al.*, 2004)

Todas ellas fueron adquiridas por distintos grupos de investigación o bien como resultado de la liberación al público de las bases de datos de competiciones. Las hemos seleccionado para nuestro trabajo porque todas fueron adquiridas con el mismo tipo de tableta gráfica (de la marca WACOM) de modo que poseen un conjunto de características comunes lo que que nos facilitaría la reutilización y comparación de sistemas. Este tipo de tableta gráfica proporciona las cinco características siguientes (fig. 3.1):

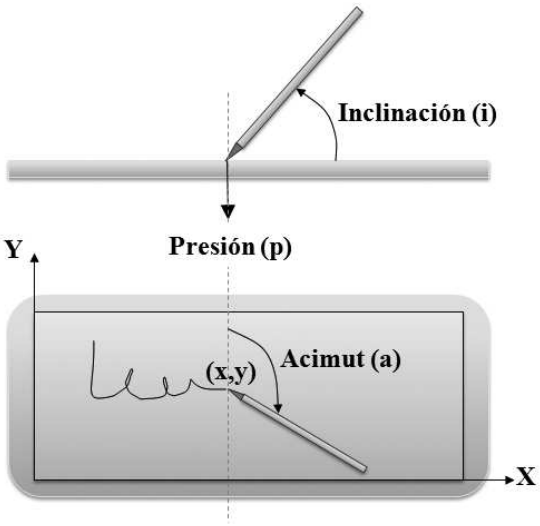
- Posición en el eje X: entre 0 y 12700 (0 - 127 mm).
- Posición en el eje Y: entre 0 y 9700 (0 - 97 mm).
- Presión: entre 0 y 1023
- Acimut: entre 0 y 3600 (0 - 360º)
- Inclinación: entre 300 y 900 (30º - 90º)

En la tabla 3.1 se adelantan algunas de las características básicas de la composición de las cuatro bases de datos utilizadas. Las columnas 3ª, 4ª y 5ª muestran los datos por usuario mientras que las columnas 6ª y 7ª se refieren a la base de datos en su totalidad. Las cifras mostradas en la tabla son los de las publicaciones de referencia, aunque en nuestros experimentos se eliminaron algunos usuarios por falta de datos.

Tabla 3.1: Composición de las cuatro bases de datos utilizadas

Base de datos	**Origen**	**Nº auténticas**	**Nº imitaciones**	**Total/usuario**	**Nº usuarios**	**Total firmas**
MCYT	Occidental	25	25	50	330	16500
SVC2004	Occidental y chinas	20	20	40	40	1600
BIOMET	Occidental	15	17	32	91	2912
MyIDea	Occidental	18	36	54	70	3780
	Media/**Total**	*19.5*	*24.5*	*44*	**531**	**24792**

A continuación describimos de forma más detallada cada una de estas bases de datos.



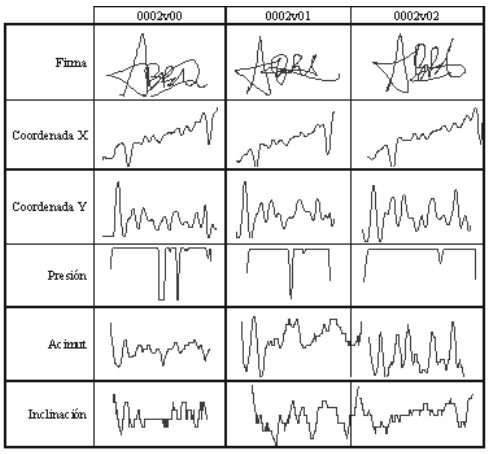


FIGURA 3.1: Tableta gráfica utilizada y características capturadas

3.1.1. BIOMET

La base de datos BIOMET (Garcia-Salicetti *et al.*, 2003) es una base de datos formada por cinco modalidades biométricas (audio, cara, mano, huella y firma). Las firmas fueron adquiridas con una tableta gráfica WACOM Intuos2 A6 a una frecuencia de muestreo de 200Hz. Para normalizar la frecuencia con respecto al resto de bases de datos, realizamos una normalización temporal de las firmas a 100Hz eliminando uno de cada dos puntos. Las características de la firma eran las mismas que las de la base de datos MCYT.

El proceso de captura de rasgos fue dividido en tres sesiones espaciadas tres y cinco meses. En la publicación de referencia se indica que el número de participantes fue de 130 para la primera, 106 en la segunda y 91 en la tercera. No obstante, para el caso de la firma sólo se disponen de datos de 84 usuarios para las tres sesiones.

En cada sesión cada usuario contribuyó con cinco firmas auténticas con lo que el total de firmas de clientes fue de 15. Respecto a las imitaciones y según la documentación oficial las imitaciones se realizaron como sigue: en la primera sesión el usuario realizó cinco imitaciones de las firmas de otro participante; en las dos sesiones restantes el usuario pasó a realizar sólo tres veces imitaciones pero de dos usuarios distintos. En total cada usuario contribuyó con un total de 17 imitaciones. No obstante en el CD de distribución recibido sólo se disponían de 12 imitaciones por usuario. Además, faltan firmas de algunos de los usuarios, en concreto 8 firmas originales y 59 imitaciones, con lo que el total de firmas es de: 84 usuarios x (15 originales + 12 imitaciones) - 67 perdidas = 2201 firmas. Por lo tanto, cada usuario tiene de media 14.90 firmas auténticas y 11.30 imitaciones.

3.1.2. MyIDea

MyIDea (Dumas *et al.*, 2005) es una base de datos multimodal resultado de la colaboración entre la Universidad de Friburgo en Suiza, la Escuela de Ingeniería de Friburgo (Suiza) y el Grupo de Escuelas de Telecomunicaciones (GET) en Paris.

La base de datos de firmas fue adquirida con una tableta gráfica A4 Intuos2 de WACOM (Wacom, 2009). Los datos registrados por esta tableta son los mismos que los de la base de datos MCYT (sección 3.1.4) y la frecuencia de muestreo también es de 100Hz.

Según la documentación publicada por sus autores, la base de datos está formada por 70 usuarios[1]. Para cada usuario se registraron 18 firmas a lo largo de tres sesiones (6 firmas por sesión). Además se recogieron imitaciones de las firmas realizadas por los mismos usuarios participantes. Estas imitaciones fueron realizadas en dos escenarios de ataque diferentes. En un primer escenario el imitador sólo disponía de la imagen de la firma a imitar. En el segundo tenía además acceso a la dinámica de la firma mediante un software dedicado.

Siguiendo este protocolo cada usuario realizó en cada sesión 6 imitaciones sin acceso a la dinámica y otras 6 con acceso a la dinámica. En total cada usuario contribuyó con un total de 12 imitaciones/sesión x 3 sesiones = 36 imitaciones. Este valor junto a las 18 firmas originales forma un total de 54 firmas por usuario.

Debido a los usuales problemas de adquisición de las bases de datos multimodales estos números no fueron del todo uniformes. De los 73 usuarios recibidos en el CD de

[1] Aunque el objetivo inicial era de 104.

distribución[2], cuatro de ellos no disponían de imitaciones, por lo que fueron eliminados de nuestros experimentos. Además, no todos contaban con el mismo número de firmas auténticas ni de imitaciones. La media de firmas auténticas por usuario fue de 17.17 (en vez de 18), mientras que el número medio de imitaciones fue de 33.04 (en vez de 36). Así pues, el número total de firmas auténticas fue de 1185 y el de imitaciones de 2280.

No existe información detallada del espacio entre sesiones de captura, aunque según consta en la publicación de referencia el intervalo era variable, desde días a semanas, para simular condiciones realistas.

3.1.3. SVC

En 2004 se celebró la primera Competición Internacional de Verificación de Firma (Signature Verification Competition, SVC) la cual supuso un hito al permitir por primera vez la evaluación objetiva de sistemas. Para ello los organizadores liberaron previamente a la competición un subconjunto de desarrollo formado por 40 usuarios. Éste ha permanecido desde entonces como sistema de referencia para promover la investigación y comparación de sistemas.

Las firmas fueron adquiridas en dos sesiones espaciadas usualmente en una semana y en cada una de ellas cada usuario realizó 10 firmas auténticas. Las imitaciones de un mismo usuario fueron realizadas por al menos cuatro personas distintas.

Algunas características de esta base de datos hacen especialmente complicado obtener buenos resultados con ella:

- Por razones de privacidad los organizadores sugirieron a los usuarios que no utilizaran su firma original.
- Las firmas fueron realizadas con un lápiz sin tinta sobre la tableta gráfica (marca WACOM) de modo que no había retroalimentación visual durante la realización de la firma.
- Los imitadores disponían de un software de visualización de la dinámica de la firma antes de realizar las imitaciones.

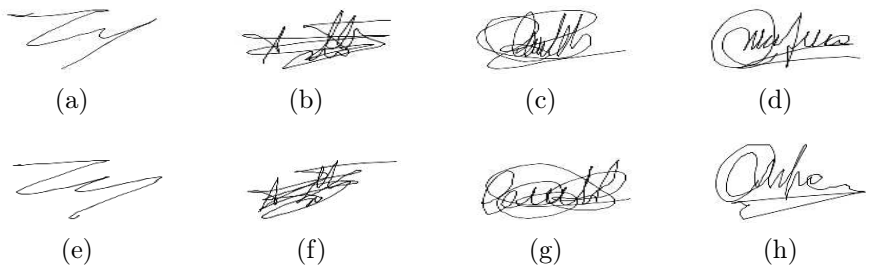
Las mayoría de las firmas fueron realizadas por personas chinas, aunque las firmas están tanto formato occidental como Chino, puesto que muchas de ellas utilizan firmas realizadas en caligrafía occidental (en idioma Inglés) en sus vidas cotidianas.

3.1.4. MCYT

Es la mayor de las bases de datos utilizadas. Fue adquirida por una iniciativa coordinada entre cuatro universidades españolas en el marco de un proyecto financiado[3]. La distribución de usuarios por Universidad fue la siguiente:

[2]Se recibieron 73 en vez de los 70 a los que se refiere la publicación.

[3]Proyecto MCyT2000: *Aplicación de la Identificación de Personas Mediante Multimodalidad Biométrica en Entornos de Seguridad y Acceso Natural a Servicios de Información.* MCyT. TIC2000-1669-C04.



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

(e) (f) (g) (h)

FIGURA 3.2: Algunas firmas de ejemplo extraídas de la base de datos MCYT: (a,b,c,d) son originales y (e,f,g,h) imitaciones

- Universidad Politécnica de Madrid (UPM). La componen 145 usuarios, numerados del 0 al 144.
- Universidad de Valladolid (UVA). Formada por 75 usuarios, numerados del 200 al 274.
- Universidad del Pais Vasco/Euskal Herriko Unibertsitatea, Escuela Superior de Ingenieros (EHU). Contiene 75 usuarios, numerados del 300 al 374.
- Escuela Universitaria Politécnica de Mataró (EUPMT). La forman 35 usuarios, numerados del 400 al 435.

El proceso de adquisición de corpus se realizó en cinco sesiones espaciadas en el tiempo varios meses. Además de las firmas del usuario, esta base de datos contiene imitaciones de su firma realizada por otros usuarios (fig. 3.2). En cada sesión el donante realizó cinco firmas propias y reprodujo de la forma más fiel posible una firma de otro usuario ya registrado en el sistema (imitando un total de cinco firmas diferentes en las cinco sesiones). Para llevar a cabo la imitación se suministró al usuario únicamente una imagen de la firma, sin tener acceso a la dinámica de ejecución. Cada firmante realizó un total de 50 firmas, 25 propias y 25 imitaciones (cinco repeticiones de otros cinco usuarios). El corpus utilizado en este trabajo consta de 330 usuarios y 16500 instancias de firma, la mitad de ellas auténticas (8250) y la otra mitad imitaciones (8250).

Datos estadísticos

Dado que la base de datos más completa es MCYT y con el objetivo de completar la sección se ha hecho un pequeño estudio estadístico de los parámetros globales de sus firmas. Los parámetros globales estudiados han sido la *duración temporal*, la *longitud* y la *velocidad* media de las firmas.

De los datos reflejados en la tabla tabla 3.2 observamos que la longitud de la firma[4] ha resultado ser el parámetro más estable entre usuarios (con menos desviación estándar) aunque también el más sencillo de imitar, ya que la media de longitud de las imitaciones está muy cercana a la media de longitud real de las firmas auténticas. La duración temporal

[4]La longitud es la suma de desplazamientos instantáneos entre cada dos puntos (es decir, la longitud de la firma 'estirada').

TABLA 3.2: Resumen de parámetros globales básicos de la base de datos MCYT

Auténticas (Imitaciones)	**Media**	**Desviación estándar (%)**	**Máximo**	**Mínimo**
Longitud (cm)	23.98 (24.25)	18 % (38 %)	47.51 (71.28)	6.12 (3.44)
Duración (s)	5.79 (7.15)	41 % (74 %)	20.20 (53.06)	0.57 (0.50)
Velocidad (cm/s)	5.71 (4.59)	28 % (56 %)	19.46 (23.96)	1.29 (0.51)

y la velocidad muestran mayor desviación entre los distintos. Como era de esperar el tiempo medio de las imitaciones es mayor que el de las firmas auténticas puesto que se realizan a menor velocidad.

Histogramas. Respecto a la duración de las firmas, puede observarse en la figura 3.3(a) que el 89 % de las firmas dura 10 segundos o menos, siendo la duración media de 5.79 segundos.

Otro dato obtenido es la longitud L de las firmas calculada como la suma de la distancia Euclídea entre los N puntos que la representan (ecuación 3.1). En la figura 3.3(b) puede observarse la distribución estadística de estas longitudes. Se observa que tiene la forma de distribución normal ligeramente sesgada hacia longitudes mayores. Como dato curioso destacar que el 63 % de las firmas tiene entre 20 y 30 cm longitud.

$$L = \sum_{i=1}^{N-1} \sqrt{(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2} \tag{3.1}$$

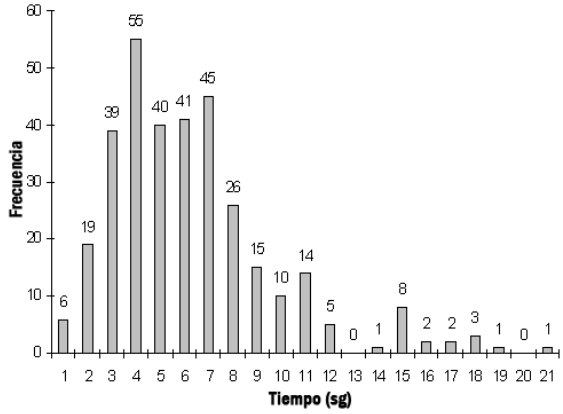
Por último, respecto a la distribución en velocidad de las firmas (figura 3.3(c)), sólo señalar que el 80 % de las firmas se realiza con velocidades entre 3 y 8 cm/sg.

Diagramas de dispersión. Para finalizar el estudio se han obtenido los diagramas de dispersión de las tres variables globales (figura 3.4) realizando varios ajustes funcionales sobre los datos para determinar su correlación.

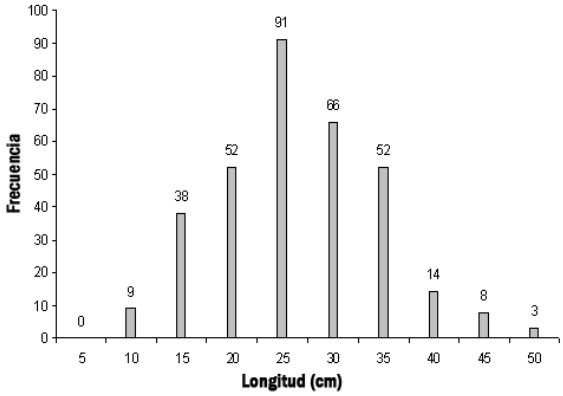
En la figura 3.4(a) vemos que la duración y la longitud tienen cierto grado de correlación ($R^2 = 0{,}4779$ mediante ajuste por función potencial). Podemos interpretar que existe una lógica tendencia a que cuanto más tiempo haya durado una firma de mayor longitud será. Sin embargo las firmas realizadas en menos tiempo recorren proporcionalmente más espacio que las que tardan más. No obstante, la dispersión de los datos impediría que conocida la longitud de la firma determináramos con certeza el tiempo que ha tardado en realizarse.

La figura 3.4(b) que corresponde al diagrama de dispersión tiempo-velocidad es la que presenta el mayor grado de correlación ($R^2 = 0{,}6641$ también mediante ajuste por función potencial). Se deduce de esta gráfica que las firmas de corta duración temporal son realizadas por firmantes que podemos considerar más 'ágiles', ya que realizan su firma a mayor velocidad media que aquellos que han tardado más tiempo. Puede deberse a que estos últimos no tengan demasiado hábito en firmar, realicen la firma de forma muy meticulosa o tengan algún tipo de problema de motilidad por edad o enfermedad.

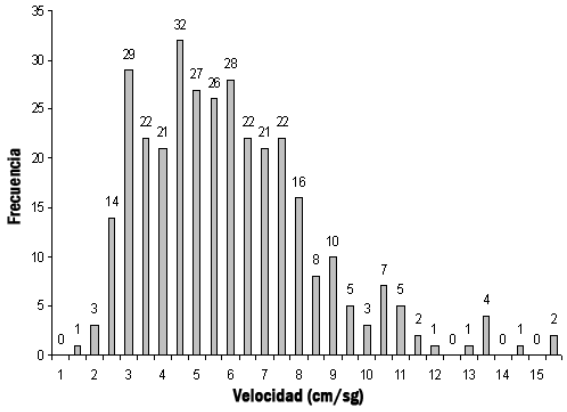
El último de los diagramas (3.4(c)) muestra que curiosamente no existe casi correlación entre la longitud de la firma y su velocidad de ejecución. Sería imposible a la vista de ello determinar la velocidad de una firma a partir de su longitud y al contrario.



(a) Histograma de duraciones (1 - 21 sgs)

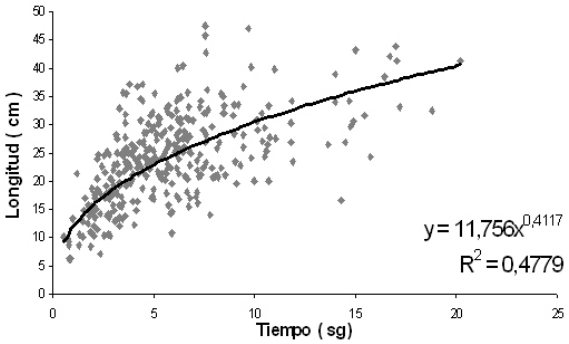


(b) Histograma de longitudes

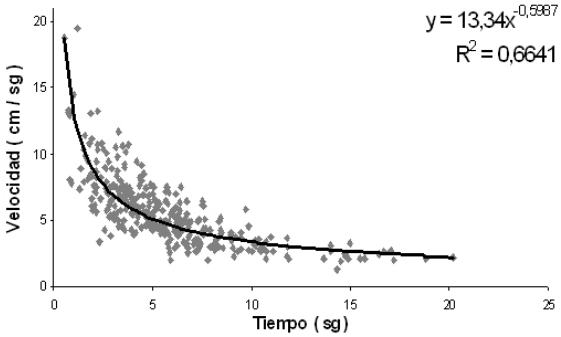


(c) Histograma de velocidades

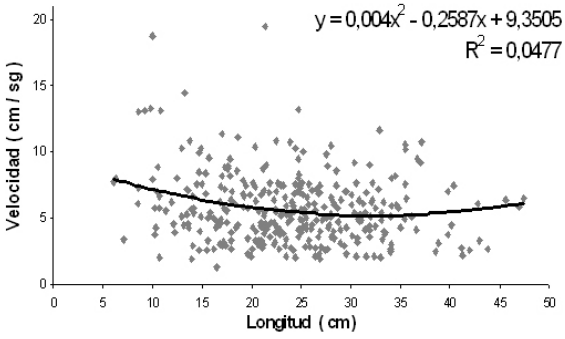
FIGURA 3.3: Histogramas de los parámetros globales básicos de MCYT



(a) Tiempo - Longitud



(b) Tiempo - Velocidad



(c) Longitud - Velocidad

FIGURA 3.4: Diagramas de dispersión de las tres características globales de MCYT

3.2 Competiciones

Uno de los recursos más importantes para calibrar el estado del arte de la tecnología Biométrica son las competiciones y evaluaciones externas. No sólo porque permiten realizar un evaluación objetiva de sistemas bajo las mismas condiciones, sino también porque las bases de datos utilizadas suelen liberarse una vez finalizado el evento para promover la investigación.

Son varias las evaluaciones llevadas a cabo en distintos rasgos biométricos tales como cara −FERET (Phillips *et al.*, 2000), FRVT (Phillips *et al.*, 2007) o FRGC (Phillips, 2006b)−, huella dactilar −FVC (Cappelli *et al.*, 2004, 2007; Maio *et al.*, 2006)− o locutor −NIST SRE (Przybocki & Martin, 2004; Przybocki *et al.*, 2006; Kajarekar *et al.*, 2009)−, promovidos por distintas instituciones, gobiernos y universidades.

En cambio en el caso de la firma sólo existen dos precedentes de este tipo bastante distanciados en el tiempo. El primero, organizado en el año 2004, fue la primera Competición Internacional en Verificación de Firma (SVC). Más recientemente, en el año 2009, se celebró la BioSecure Signature Evaluation Campaign (BSEC), a la cual presentamos el sistema descrito en el capítulo 5.

A continuación describiremos brevemente ambas competiciones, las reglas que las rigieron así como sus resultados más relevantes, con el fin de mostrar el nivel en el que se encontraba la tecnología del reconocimiento de firma.

3.2.1. SVC 2004

SVC 2004 surgió con el objetivo de permitir la comparación de los sistemas de los participantes de forma sistemática bajo las mismas reglas y misma base de datos. Para ello los organizadores crearon una base de datos formada por 100 usuarios, de los cuales 40 fueron liberados para que los participantes desarrollaran sus sistemas, y el resto se dedicaron para la evaluación final. Los detalles de esta base de datos ya han sido descritos en la sección 3.1.3.

Objetivos. El objetivo de esta primera competición era el permitir a los investigadores comparar el rendimiento de sus sistemas de forma sistemática sobre una base de datos de prueba común y bajo unas mismas reglas de evaluación.

Protocolo. La competición fue dividida en dos tareas. En la primera el sistema debía utilizar únicamente las coordenadas geométricas mientras que con el segundo se disponía de las características adicionales de la tableta (presión y ángulos del lápiz).

Para el entrenamiento del sistema se utilizaron cinco firmas de entrenamiento escogidas de la primera sesión. Para evaluar la tasa de falsos rechazos se utilizaron las firmas genuinas de la segunda sesión. La tasa de falsas aceptaciones en el caso de escenario casual se determinó seleccionando 20 firmas escogidas aleatoriamente de otros 20 usuarios. En el caso de la prueba con imitaciones se utilizaron las 20 imitaciones disponibles de las firmas de cada usuario. Con todo ello se obtuvieron las puntuaciones de cada sistema y se determinó el error en forma de EER. Este proceso fue repetido 10 veces escogiendo cada vez un conjunto de firmas de entrenamiento diferente. Aunque no se indica de forma explícita en

Tabla 3.3: Equipos participantes en SVC 2004

ID Equipo	Institución	País	Miembro(s)	Tarea(s)
3		Australia	V. Chandran	1,2
4	Anónimo			1,2
6	Sabanci University	Turkey	Alisher Kholmatov Berrin Yanikoglu	1,2
8	Anónimo			2
9	Anónimo			1,2
12	Anónimo			1
14	Anónimo			1,2
15	Anónimo			1
16	Anónimo			1
17	Anónimo			1,2
18	Anónimo			1,2
19	Biometrics Research Laboratory Universidad Politécnica 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 Télécommunications	France	Bao Ly Van Sonia Garcia-Salicetti Bernadette Dorizzi	2

el artículo de referencia parece que la evaluación del error fue calculado usando umbral individual. Así lo parece puesto que en los resultados se muestra no sólo el valor promedio del EER de cada sistema sino también su desviación estándar, así como el máximo EER de cada sistema.

Resultados. Un total de 15 equipos participaron en la primera tarea y 12 equipos en la segunda (tabla 3.3).

El error obtenido fue bastante dispar en valores absolutos entre las bases de datos de desarrollo (3.4 y 3.5) y de evaluación (tablas 3.6 y 3.7). No obstante, en ambas bases de datos, los mejores sistemas fueron los de los equipos 6 y 19 que estaban basados en DTW y HMM respectivamente. Las menores tasas de error obtenidas sitúan al estado del arte en imitaciones en el rango 2-3 %. En escenario casual este valor baja al rango 1-2 %.

TABLA 3.4: Resultados de SVC 2004 en la tarea 1 con el conjunto de desarrollo

	10 genuinas + 20 imitaciones			**10 genuinas + 20 casuales**		
ID Equipo	**Media**	**Desv.Estándar**	**Máximo**	**Media**	**Desv.estándar**	**Máximo**
6	**5.50 %**	7.73 %	30.00 %	3.65 %	4.80 %	40.00 %
26	6.45 %	10.41 %	50.00 %	3.49 %	4.53 %	29.63 %
24	7.33 %	7.71 %	35.00 %	2.93 %	3.72 %	20.00 %
15	9.80 %	13.90 %	70.00 %	2.90 %	3.60 %	20.00 %
14	11.10 %	11.11 %	50.00 %	3.36 %	4.36 %	20.00 %
19c	11.98 %	17.65 %	95.00 %	**2.87 %**	3.68 %	16.67 %
19b	11.99 %	17.66 %	95.00 %	2.88 %	3.68 %	16.67 %
18	14.34 %	16.11 %	90.00 %	4.29 %	5.45 %	30.00 %
19a	14.91 %	18.98 %	96.67 %	2.90 %	3.64 %	15.00 %
16	15.67 %	13.24 %	60.00 %	2.89 %	3.64 %	20.00 %
17	16.45 %	11.79 %	60.00 %	4.66 %	5.22 %	40.00 %
4	18.99 %	13.95 %	56.52 %	11.57 %	13.28 %	70.00 %
3	25.83 %	22.11 %	95.00 %	6.58 %	9.20 %	55.00 %
12	31.32 %	18.09 %	85.00 %	11.67 %	9.58 %	41.43 %

TABLA 3.5: Resultados de SVC 2004 en la tarea 2 con el conjunto de desarrollo

	10 genuinas + 20 imitaciones			**10 genuinas + 20 casuales**		
ID Equipo	**Media**	**Desv.Estándar**	**Máximo**	**Media**	**Desv.Estándar**	**Máximo**
19b	**6.90 %**	9.45 %	50.00 %	**3.02 %**	3.65 %	15.00 %
19c	6.91 %	9.42 %	50.00 %	**3.02 %**	3.65 %	15.00 %
6	6.96 %	11.76 %	65.00 %	3.47 %	4.23 %	20.00 %
29	7.64 %	12.62 %	60.00 %	4.45 %	6.84 %	50.00 %
19a	8.90 %	11.72 %	71.00 %	3.08 %	3.74 %	15.00 %
14	11.29 %	13.47 %	70.00 %	4.41 %	5.35 %	28.57 %
18	15.36 %	13.88 %	60.00 %	6.39 %	7.03 %	45.00 %
17	19.00 %	14.43 %	70.00 %	4.29 %	4.69 %	30.00 %
3	20.01 %	18.44 %	76.19 %	5.07 %	8.13 %	44.44 %
4	21.89 %	17.05 %	73.33 %	8.75 %	9.71 %	48.72 %

TABLA 3.6: Resultados de SVC 2004 en la tarea 1 con el conjunto de evaluación

	10 genuinas + 20 imitaciones			**10 genuinas + 20 casuales**		
ID Equipo	**Media**	**Desv.Estándar**	**Máximo**	**Media**	**Desv.Estándar**	**Máximo**
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 %

TABLA 3.7: Resultados de SVC 2004 en la tarea 2 con el conjunto de evaluación

	10 genuinas + 20 imitaciones			**10 genuinas + 20 casuales**		
ID Equipo	**Media**	**Desv.Estándar**	**Máximo**	**Media**	**Desv.Estándar**	**Máximo**
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 %

3.3 Resumen

En este capítulo hemos descrito los bancos de prueba existentes en el campo del reconocimiento de firma. Las cuatro bases de datos descritas (MCYT, BIOMET, SVC y MyIDea) junto a las competiciones realizadas hasta la fecha (SVC 2004 y BSEC 2009) constituyen un recurso inestimable para el investigador. Otras bases de datos multimodales con datos de firma de más reciente adquisición son BIOSECURE (Ortega-Garcia *et al.*, 2009) y BiosecurID (Fierrez *et al.*, 2009), que junto a la base de datos de BSEC 2009, añaden datos no sólo capturados con tableta gráfica, sino también con PDA lo que sin duda ayudará al desarrollo de la firma como modalidad biométrica en la práctica.

En el próximo capítulo se expondrá el primero de los sistemas desarrollados en esta tesis. Dicho sistema se ha basado en un nuevo concepto de *modelos ocultos de Markov dependientes del usuario* que es una de las principales aportaciones de esta tesis.

4

Modelos Ocultos de Markov

En capítulos anteriores se ha comentado que una de las principales técnicas para el reconocimiento de firma son los modelos ocultos de Markov (HMM). Por ello, uno de los dos sistemas desarrollados en esta tesis se ha basado en este método de carácter estadístico.

El capítulo se divide en tres bloques de contenido. El primero, formado por las secciones 4.1 y 4.2, sirve para introducir al lector en la forma en la que se aborda el problema de verificación de firma con HMM. El segundo bloque, que comprende las secciones 4.3 a 4.7, introduce el concepto de HMM dependiente del usuario y evalúa el rendimiento de este nuevo enfoque aplicado a la construcción de sistemas de verificación de firma. El tercer y último bloque lo forma la sección 4.8 y consiste en la descripción de un sistema basado en el nuevo concepto introducido pero con algunas optimizaciones para mejorar el rendimiento. Además, en este tercer bloque, se propone un método de selección de la estructura personal del HMM a partir de los datos de entrenamiento.

Parte de los resultados que se verán en este capítulo han sido publicados (Pascual-Gaspar & Cardeñoso-Payo, 2006; 2007a; 2007).

4.1 Introducción

Los modelos ocultos de Markov han sido aplicados al reconocimiento de escritura y firma manuscritas con buenos resultados desde mediados de los años 90 (Cappé, 2001). Los distintos sistemas publicados en la literatura de estos años difieren entre sí en varios aspectos como por ejemplo, en las técnicas de preprocesado de datos, en las características de la firma que utilizan, en las técnicas de normalización, etc., aunque la mayoría de ellos comparten una misma estrategia desde un punto de vista arquitectónico ya que están basados en un conjunto de *parámetros estructurales* total o parcialmente independiente del usuario.

Quiere esto decir que la estructura óptima del modelo es seleccionada entre un conjunto finito de posibles alternativas como la que proporciona el mejor rendimiento en promedio sobre todos los usuarios de la base de datos de evaluación. Con esta solución se necesitan un número relativamente elevado de características, que además lleven implícita información sobre la dinámica de la firma. Ello se debe a que con este enfoque todas las firmas se modelan con una única configuración estructural del modelo, independientemente de la longitud (número de puntos muestreados) de la firma del usuario. A este 'tipo' de modelos

ocultos de Markov les denominaremos de aquí en adelante *HMM Independientes del Usuario* (HMM-IU). Al utilizar HMM-IU la diferencia entre los modelos de distintos usuarios residirá únicamente en sus *parámetros estadísticos*, es decir medias y varianzas de las distribuciones gaussianas de los estados y probabilidades de transición entre estados. Estos parámetros estadísticos son específicos de cada usuario y se calculan típicamente mediante el algoritmo de Baum-Welch (Rabiner, 1989).

Una aproximación diferente a la anterior consiste en la personalización de la estructura del HMM a la cantidad de información de las firmas del usuario. Este enfoque basado en *HMM Dependientes del Usuario* (HMM-DU) creemos que permitiría una mejor representación de la naturaleza temporal de las firmas y por tanto un mejor rendimiento. De hecho, una vez definidas las variables experimentales tales como el número de firmas de entrenamiento y el conjunto de características, los HMM-IU pueden verse como un caso particular de los HMM-DU.

El trabajo que se describe en este capítulo analiza las posibilidades de los HMM-DU en la tarea de verificación de firma dinámica y tiene los dos siguientes objetivos principales:

1. Determinar la cota teórica del rendimiento de los HMM-DU en condiciones de uso realistas. Los resultados obtenidos servirán para determinar si el nuevo enfoque produciría mejoras suficientemente importantes en el rendimiento comparado con los sistemas basados HMM-IU.

2. Desarrollar un método para encontrar la estructura óptima de los HMM-DU a partir de los datos a modelar, firmas dinámicas en nuestro caso. Ello nos permitiría construir un sistema realista basado en HMM-DU orientado a escenarios prácticos.

4.2 Trabajos precedentes

Antes de pasar a describir en detalle nuestro trabajo repasaremos, en orden cronológico, algunas publicaciones importantes en el campo de la Biometría de la firma dinámica basada en HMM y que han servido de referencia en nuestra investigación.

Uno de los primeros trabajos de aplicación de HMM a la verificación de la firma dinámica fue realizado por Yang *et al.* (1995). En él se describen varios experimentos realizados con una pequeña base de datos de 31 usuarios y diferentes configuraciones de HMM-IU. Los mejores resultados se consiguieron con un modelo formado por 6 estados y normalizando la longitud de la firma a 300 puntos. Los errores de este sistema fueron de 1.75 % de FRR y 4.44 % de FAR en prueba con firmas casuales. Se utilizaron 8 firmas para entrenar los modelos y una única característica de la firma con valores cuantizados.

Kashi *et al.* (1997) describen un sistema basado en HMM semicontinuo (HMM sin número de estados fijo) que usaba 6 firmas para el entrenamiento y que sólo necesitaba de dos características locales. Fue evaluado con firmas de la base de datos Murray-Hill (Nelson *et al.*, 1994). Obtuvieron un ERR inicial de 5 %, que fue reducido a 2.5 % con la inclusión de un experto basado en 23 parámetros globales. No se especifica el método de evaluación aunque el texto parece indicar que la prueba fue realizada con firmas imitadas.

Dolfing *et al.* (1998) proponen un sistema basado en HMM en el que el número de estados era establecido como 0.8 veces el número de vectores de características de las firmas

de entrenamiento. Dividieron la firma en segmentos cada uno de los cuales estaba formado por 32 componentes. Evaluaron el sistema con una base de datos propietaria obteniendo un valor de EER de 1.9 % usando 15 firmas de entrenamiento, un número ciertamente elevado para aplicaciones prácticas.

Rigoll & Kosmala (1998) basados en su experiencia en reconocimiento de firma escrita crearon un sistema de verificación de firma dinámica basado en HMM con número de estados variable, entre 14 y 24 según la longitud de la firma. El sistema es evaluado sobre una pequeña base de datos de 14 personas que contiene firmas imitadas. Obtienen un 1 % de falsos rechazos sin falsas aceptaciones, aunque utilizando 16 firmas para entrenar los modelos. El número de características utilizadas fue de 27.

Wessels & Omlin (2000) exponen un sistema híbrido formado por HMM con número de estados obtenido a partir de un mapa auto-organizado de Kohonen. El HMM tiene topología izquierda-derecha y es entrenado con 15 firmas. Los resultados sobre una base de datos de 50 usuarios son de 13 % de FAR con 0 % de FRR empleando únicamente características geométricas.

Fuentes *et al.* (2002) evaluaron su sistema con la base de datos Philips comparando su trabajo con Rigoll & Kosmala (1998). Para ello usaron 15 firmas de entrenamiento y HMM discreto como técnica de modelado. Con 17 características locales y un número variable de estados (entre 6 y 12 dependiendo de la longitud de la firma) su sistema obtuvo tasas de error de 3.26 % en FAR y 19.05 % en FRR con firmas imitadas. Este error fue reducido a 4.62 % y 8.25 % en FAR y FRR respectivamente mediate fusión con una red neuronal basada en características globales. La fusión de ambos sistemas fue llevada a cabo con SVM.

Yoon *et al.* (2002) utilizan características polares en un sistema basado en HMM de 5 estados. La base de datos de prueba estaba formada por 100 usuarios y contaba con 20 firmas por usuario. Para el entrenamiento utilizaron 15 firmas y las 5 restantes para la prueba. El EER obtenido fue de 2.2 % con firmas casuales.

Igarza *et al.* (2003) mejoraron el sistema propuesto por Yang *et al.* (1995) a nivel de preprocesamiento y evaluaron el sistema sobre 150 usuarios de la base de datos MCYT con firmas imitadas. Con HMM-IU de 6 estados y normalizando las firmas en longitud obtuvieron tasas de EER de 9.25 % usando 9 firmas para el entrenamiento. Posteriormente (Igarza *et al.*, 2004), este sistema fue mejorado en el preprocesamiento, incluyendo hasta un total de 9 características locales, obteniendo un EER de 4.66 %.

Shafiei & Rabiee (2003) proponen un sistema basado en HMM y segmentación variable a partir de los puntos perceptualmente importantes de la firma. Usaron modelos con 10 gaussianas por estado y número de estados variable. El número de estados era determinado empíricamente a partir del número de segmentos de la firma siendo la relación estados-segmentos de 1 a 2. Entrenando el modelo de cada usuario con 5 firmas y 7 características locales obtuvieron un EER de 11.5 % con falsificaciones casuales y 22.5 % con firmas imitadas. La base de datos fue adquirida ad hoc.

Muramatsu & Matsumoto (2003) se basan en coordenadas polares para su sistema a partir de las coordenadas geométricas de la trayectoria de la firma. Sobre una base de datos de 16 usuarios obtienen un EER de 2.6 % con firmas imitadas. El número de estados del modelo es obtenido a partir los cambios de ángulos de la dirección de la firma.

Fierrez-Aguilar posee varias publicaciones donde se evalúa su sistema basado en HMM-

IU en distintas circunstancias. En una de ellas evalúa el sistema bajo diferentes estrategias de normalización de puntuaciones utilizando 14 características locales (Fierrez-Aguilar *et al.*, 2005c). Con el corpus de desarrollo de la base de datos SVC 2004 (Yeung *et al.*, 2004) y cinco firmas de entrenamiento (multisesión) obtiene un EER con firmas casuales de 0.50 % y de 5.79 % con firmas imitadas. El umbral de decisión para medir el error fue dependiente del usuario.

En otras de las publicaciones el sistema basado en características regionales (HMM) se fusiona con un experto basado en características locales, basado en DTW. El error del subsistema basado en HMM con cinco firmas de entrenamiento fue de 0.34 % con firmas casuales y de 8.04 % con firmas imitadas. Existen diferencias con la publicación anterior porque las firmas de entrenamiento fueron tomadas de la primera sesión (monosesión). El rendimiento de este sistema es excelente en escenario casual y de hecho fue el ganador de la competición SVC 2004 en dicho apartado (Fierrez-Aguilar *et al.*, 2005a).

Igualmente, el mismo sistema fue evaluado sobre la base de datos MCYT. Con cinco firmas de entrenamiento el EER fue de 2.51 % con firmas imitadas y 0.59 % con firmas casuales con umbral de decisión dependiente del usuario. En dicho trabajo estos valores fueron reducidos a 2.12 % y 0.24 % mediante fusión con un experto basado en parámetros globales (Fierrez-Aguilar *et al.*, 2005b).

Por último, en el trabajo de Fierrez *et al.* (2007), se sintetiza el trabajo descrito en los artículos anteriores incluyendo pruebas con la base de datos MCYT. En dicho artículo se midió el error con 10 firmas de entrenamiento sobre un subconjunto de MCYT dando valores de error de 0.09 % con firmas casuales y 0.78 % con firmas imitadas.

Ly *et al.* (2007) continúan el trabajo iniciado por Fuentes *et al.* (2002) y mejoran el sistema inicial añadiendo fusión intramodal entre dos medidas de distancia generadas por el mismo HMM. Por un lado la verosimilitud generada por el HMM (la medida habitual) y por otro la distancia entre las secuencias de estados recorridas (camino de Viterbi). Este sistema es intensamente probado sobre 4 bases de datos y el error cometido en promedio entre todas ellas es de 4.5 %. El número de firmas de entrenamiento es de cinco y el vector de características estaba formado por 25 características estáticas y dinámicas. Garcia-Salicetti *et al.* (2007) realizaron un estudio de fusión de sistemas resultando que el mejor resultado con 5 firmas de entrenamiento sobre MCYT fue obtenido con la fusión del sistema de Van-Bao y Fierrez-Aguilar, ambos basados en HMM, con un EER de 3.40 % con firmas imitadas.

Para finalizar Martinez-Diaz *et al.* (2008) aplicaron y ampliaron los conceptos de HMM-DU iniciados en el trabajo expuesto en este capítulo con dispositivos portátiles (PDA). Sobre la base de datos BIOSECURE los autores consiguieron reducir los valores de EER obtenidos con HMM-IU de 7.3 % a 5.2 % con firmas casuales y de 20.5 % a 15.8 % con firmas imitadas. Señalar que en este tipo de escenarios los errores suelen ser mayores que con firmas adquiridas con tableta gráfica debido a una menor ergonomía del dispositivo y una peor calidad de muestreo.

En la tabla 4.1 se resumen los anteriores trabajos junto a sus características más relevantes.

De la revisión de estos trabajos extraemos las siguientes conclusiones:

a) Relativo a la configuración del HMM existe acuerdo en la comunidad en que la topología que mejores resultados proporciona es la topología 'izquierda-derecha' (LTR).

Tabla 4.1: Sistemas relevantes de VAFD basados en HMM desde 1995

#	**Autor**	**Año**	**Base de datos**	**Usuarios**	**FE**[a]	**NC**[b]	**Sistema**	**Estados**[c]	**Prueba**	**Error**
1	Yang	1995	Ad hoc	31 (496)	8	1	HMM(LTR)	H=6	Casuales	[FAR, FRR] = [4.44, 1.75]
2	Kashi	1997	Murray-Hill	59 (867)	6	2	HMM Semicont. (LTR)	No aplica	Imitac.	EER = 5.0
3	Dolfing	1998	Philips	51 (4770)	15	32	HMM (LTR)	H=0.8 × segm.	Imitac.	EER = 1.9
4	Rigoll	1998	Ad hoc	14 (340)	16	27	HMM (LTR)	H=[14..24]	Imitac.	[FAR, FRR] = [0.0, 1.0]
5	Wessels	2000	Ad hoc	50 (5000)	15	7	HMM(LTR) + SOM	H=25 M=2	Imitac.	[FAR, FRR] = [13.0, 0.0]
6	Fuentes	2002	Philips	51 (4770)	15	17	HMM(LTR)	H=[6..12]	Imitac.	[FAR, FRR] = [3.26, 19.05]
7	Yoon	2002	Ad hoc	100 (2000)	15	3	HMM(LTR)	H=5	No esp.	EER = 2.2
8	Igarza	2003	MCYT	150 (7500)	9	5	HMM(LTR)	H=6	Imitac.	EER = 9.25
9	Shafiei	2003	Ad hoc	69 (1632)	5	7	HMM	H=0.5 × segm. M=10	Imitac.	[FAR, FRR] = [4,0, 12.0]
10	Muramatsu	2003	Ad hoc	14 (5018)	5 a 25	1	HMM(LTR)	H=Nº cambios direcc	Imitac.	EER = 2.6
11	Igarza	2004	MCYT	100 (5000)	No esp.	9	HMM(LTR)	H=6	Imitac.	EER = 4.66
12	Fierrez	2005c	SVC2004	40 (800)	5	14	HMM(LTR)	H=2 M=32	Ambas	EER[Rd, Sk] = [0.50, 5.79]
13	Fierrez	2005a	SVC2004	40 (800)	5	14	HMM(LTR)	H=2 M=32	Ambas	EER[Rd, Sk] = [0.34, 8.04]
14	Fierrez	2005b	MCYT	330 (16500)	5	14	HMM(LTR)	H=2 M=32	Ambas	EER[Rd, Sk] = [0.59 , 2.51]
15	Van-Bao	2007	Varias[d]	275(aprox. 12000)	5	25	HMM(LTR)	H=T / (30×M) M=4	Imitac.	EER = 4.5
16	Garcia-Salicetti	2007	MCYT	330 (16500)	5	-	Fusión #12 y #15	-	Imitac.	EER = 3.40
17	Fierrez	2007	MCYT	195 (9250)	10	14	HMM(LTR)	H=2 M=32	Ambas	EER[Rd, Sk] = [0.09 , 0.78]
18	Martinez-Diaz	2008	BIOSECURE	50 (2000)	5	12	HMM-UD	H,M dep. de usuario	Ambas	EER[Rd, Sk] = [5.2 , 15.8]

[a]FE: Firmas de entrenamiento
[b]NC: Número de características
[c]H: Num. estados, M: Num gaussianas
[d]MCYT, Philips, BIOMET y SVC2004

Aunque algunos trabajos evalúan el uso de otras topologías los mejores resultados siempre se obtienen con aquella. Respecto a los parámetros estructurales (número de estados y de gaussianas) no existe acuerdo en el método de selección de ambos. Algunos sistemas utilizan una búsqueda sistemática más o menos descrita en los trabajos de la configuración independiente del usuario (HMM-IU) con la que obtienen sus mejores resultados dado el conjunto de características y base de datos empleada. Otros sin embargo proponen valores empíricos del número de estados y gaussianas del modelo sin demasiada justificación.

b) Debido a la naturaleza estadística de los HMM el número de datos de entrenamiento suele ser bastante elevado para aplicaciones prácticas. En muchos de los trabajos anteriores el número de firmas necesarias para construir el modelo del usuario es usualmente mayor que el que necesitan otros algoritmos basados en plantillas (tales como DTW) en los que se podría empezar a utilizar el sistema incluso con una única firma de entrenamiento.

c) Exceptuando algún caso aislado, el número de características empleadas por los sistemas también es elevado lo cual conlleva una mayor demanda computacional así como mayores necesidades de almacenamiento. Además, cuanto más complejas o especializadas sean las características utilizadas, mayor es el coste del dispositivo de adquisición, comprometiéndose también la aplicabilidad del sistema en múltiples entornos.

d) El error de los sistemas al utilizar pocas firmas de entrenamiento es bastante elevado por lo que muchos de los sistemas realizan técnicas de fusión intramodal (local-local o local-global) para mejorar su rendimiento. Cabe decir que aunque la fusión mejore los resultados no hay que olvidar que añade complejidad al resultado final.

El sistema que proponemos en este capítulo pretende ser una ayuda en la resolución de los inconvenientes anteriores y así facilitar la implantación de los sistemas de VAFD basados en HMM en entornos prácticos. El sistema que proponemos buscará obtener un rendimiento competitivo en condiciones de aplicación práctica realistas. Concretamente, se trabajará con un número muy reducido de firmas de entrenamiento (tres por usuario) y un conjunto de características sencillo y reducido. Para conseguirlo nos basaremos en un nuevo enfoque al problema basado en HMM dependiente del usuario (HMM-DU) y que a continuación presentamos.

4.3 HMM dependiente de usuario

Ilustremos la idea subyacente detrás de los HMM-DU mediante la figura 4.1. En ella puede observarse la correlación entre los parámetros estructurales para un segmento ficticio de firma (fig. 4.1-a). En el resto de ilustraciones pueden verse cinco formas diferentes de modelar dicho segmento con distintas combinaciones de estados y número de gaussianas por estado. En todos los casos el modelo consta del mismo número de grados de libertad, ($N_S \times N_G = 16$). El caso (b), que es equivalente a un GMM, consta de un único estado y por tanto no tiene en cuenta la dinámica de la firma a través de las probabilidades de

transición entre estados. En el último caso (f) la estrategia es totalmente opuesta puesto que utiliza un total de 16 estados para modelar la dinámica del segmento de firma. Como cada estado consta de una única gaussiana el número de grados de libertad sigue siendo el mismo.

Así pues los distintos parámetros estructurales de los HMM tienen fines diferentes. El número de gaussianas N_G modela la intravariabilidad espacial de las muestras de referencia del usuario. Por otro lado el número de estados N_S se encarga de modelar la intravariabilidad temporal. Al aplicar HMM a la VAFD es especialmente importante seleccionar los valores de estos dos parámetros para captar tanto la especificidad (intervariabilidad) como la inestabilidad (intravariabilidad) inherente de las firmas de un mismo usuario en los dos dominios, espacial y temporal.

Ante esta disyuntiva cabía preguntarse qué influencia tienen los parámetros estructurales en el rendimiento final del sistema. Para ello, se planificaron una serie de experimentos bajo las mismas condiciones para aislar el resultado de las otras variables en juego. En concreto utilizamos un número fijo y reducido de firmas de entrenamiento (tres) y el mismo conjunto de características de la firma. En las siguientes secciones detallamos los experimentos realizados empezando por la adquisición y el preprocesamiento de las firmas.

4.4 Adquisición y preprocesamiento

Todos los experimentos de este capítulo fueron realizados con las firmas de la base de datos MCYT (Ortega-Garcia *et al.*, 2003b), descrita en detalle en el capítulo 3.

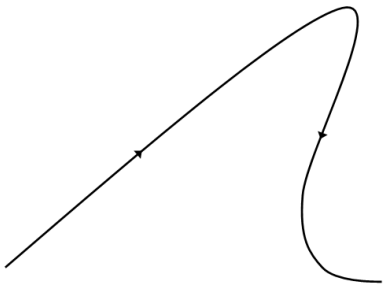
Dado que se pretendía evaluar la influencia de los parámetros estructurales no se utilizó un conjunto de características optimizado, sino que se tomaron directamente las cinco características básicas proporcionadas por la tableta gráfica:

- Posición en el eje X: entre 0 y 12700 (0 - 127 mm).
- Posición en el eje Y: entre 0 y 9700 (0 - 97 mm).
- Presión: entre 0 y 1023
- Acimut: entre 0 y 3600 (0 - 360º)
- Inclinación: entre 300 y 900 (30º - 90º)

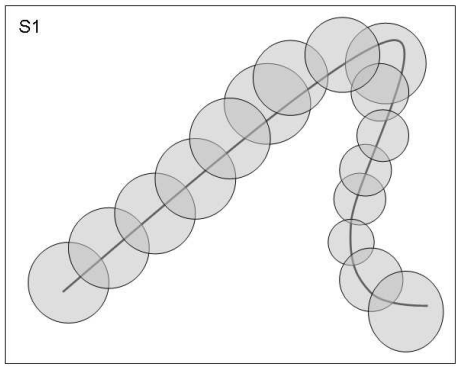
Como se explica en el capítulo 3 en el proceso de adquisición de la base de datos la tableta gráfica fue dividida en una rejilla de celdas de 3.75 cm $\times$ 1.75 cm de forma que un mismo usuario firmaba en distintas zonas. Este procedimiento sólo afectaba a los valores absolutos de las coordenadas geométricas x, y por lo que sus valores fueron corregidos eliminado los mínimos:

$$x' = x - x_{min} \tag{4.1}$$
$$y' = y - y_{min} \tag{4.2}$$

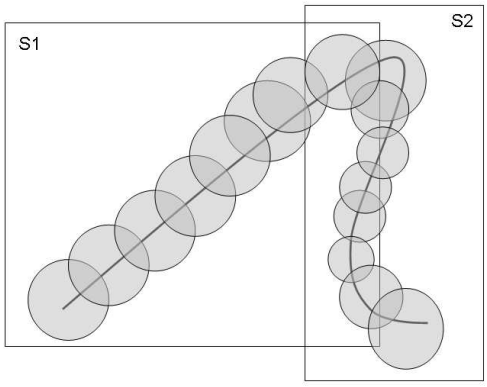
Tanto el conjunto de características empleado como la normalización realizada sobre ellas fueron mejorados en el sistema expuesto más adelante en este capítulo (4.8) puesto que se comprobó que tenían un efecto importante en el rendimiento.



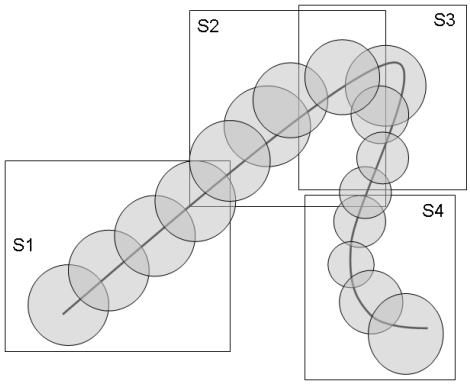
(a) Segmento de firma a modelar



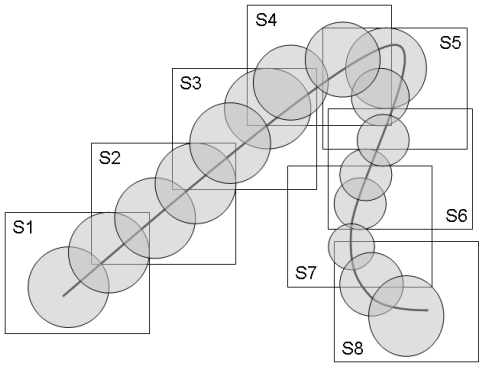
(b) $N_S = 1, N_G = 16$



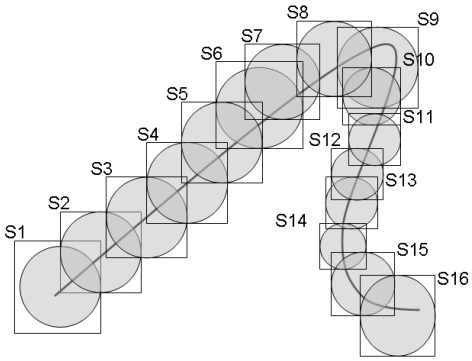
(c) $N_S = 2, N_G = 8$



(d) $N_S = 4, N_G = 4$



(e) $N_S = 8, N_G = 2$



(f) $N_S = 16, N_G = 1$

FIGURA 4.1: Ejemplo de seis configuraciones estructurales para modelar un segmento de firma mediante HMM (N_S: Núm. de estados, N_G: Núm. de gaussianas por estado).

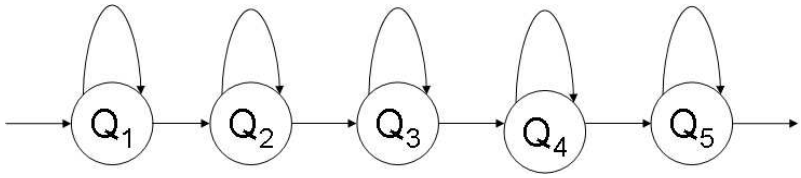


FIGURA 4.2: Topología de un HMM LTR de cinco estados sin saltos

4.5 Experimentos

Para explicar el procedimiento experimental seguido expresaremos un experimento como una función (ec. 4.5) de las siguientes variables:

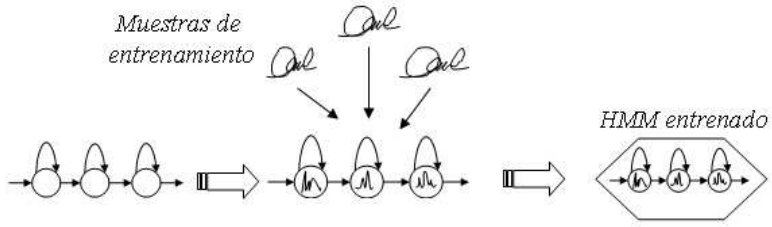
- la identidad del usuario: N_U
- el número de estados: N_S
- el número de gaussianas por estado: N_G
- el tipo de imitación: I_K $(K = [Sk|Rd])$. *Sk:firmas imitadas, Rd:firmas casuales.*

$$E(N_U, N_S, N_G, I_K) \rightarrow EER \tag{4.3}$$

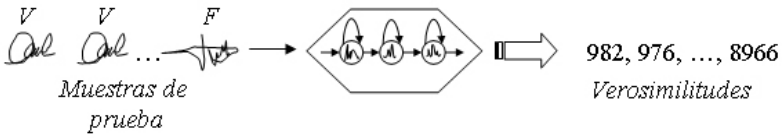
Se realizaron experimentos en base a las variables anteriores para evaluar el rendimiento de los modelos independientes (HMM-IU) versus los modelos dependientes del usuario (HMM-DU). En ambos casos se mantuvieron los siguientes parámetros experimentales que, aunque afectan al rendimiento, no eran objetivos de nuestro estudio:

- **La topología del HMM**. Debido la naturaleza intrínsecamente lineal de la ejecución de la firma escrita, y su probada eficacia se ha utilizado la topología izquierda-derecha (LTR, Left-To-Right). En la topología utilizada sólo son posibles transiciones de un estado al siguiente o a sí mismo (fig. 4.2).
- **Número de firmas de entrenamiento**. En un sistema realista el número de muestras de entrenamiento debe ser reducido para evitar el rechazo del usuario en el registro. Por ello sólo se emplearon tres firmas para generar el modelo de referencia.
- **Número de gaussianas por estado**. Aunque un HMM puede tener asociado un número de gaussianas diferente por cada uno de sus estados, en este trabajo y en la mayoría de las publicaciones existentes por simplicidad se establece un mismo número para todos los estados. Aunque este número es una constante del modelo su valor N_G ha sido una de las variables de estudio (tomando valores entre 1 y 5).

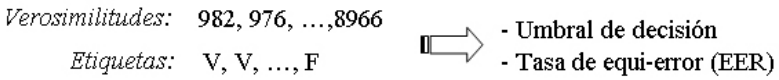
La metodología seguida en la ejecución de las pruebas, según el tipo de impostor utilizado, fue la siguiente: para la prueba en escenario casual se tomaron las 22 firmas genuinas no usadas para la creación del modelo y 250 muestras de impostor casual de diferentes



(a) Entrenamiento del modelo



(b) Fase de prueba



(c) Evaluación del error

FIGURA 4.3: Etapas de un experimento

usuarios escogidos de forma aleatoria. Para la prueba en escenario seguro se usaron las 22 muestras genuinas no utilizadas en el entrenamiento y las 25 firmas imitadas disponibles por usuario.

El error es expresado mediante la tasa de equierror (EER) como suele ser habitual en el dominio de la firma dinámica. Para el cálculo del error se utilizó umbral individual y el valor final del error se obtuvo promediando los errores de cada usuario. En la figura 4.3 puede verse un esquema del procedimiento seguido.

A continuación se dan detalles específicos del procedimiento experimental según el tipo de HMM evaluado.

4.5.1. Experimentos con HMM-IU

Para determinar el rendimiento de los HMM-IU se evaluaron 49 configuraciones diferentes obtenidas al variar el número de estados N_S y el número de gaussianas N_G entre 1 y 64 en potencias de 2 (4.4):

$$M_{IU} = (N_S, N_G)_{7x7} \tag{4.4}$$

Calculamos el valor del EER medio con cada configuración y además anotamos el número de usuarios para los cuales el modelo no pudo ser entrenado por falta de datos. En este caso se utilizaron falsificaciones casuales para medir el error.

```
bucle de estados: N_S = {1, 2, 4, 8, 16, 32, 64}
  bucle de gaussianas: N_G = {1, 2, 4, 8, 16, 32, 64}
    bucle de usuarios: N_U = [1..330]
      bucle de tipo de impostor: I_K = {I_Rd}
        EER ← E(N_U, N_S, N_G, I_K)
```

FIGURA 4.4: Pruebas con HMM-IU: se realizaron un total de 16.170 (7×7×330×1) experimentos

```
bucle de usuarios: N_U = [1..330]
  bucle de estados: N_S = [1..111]
    bucle de gaussianas: N_G = [1..5]
      bucle de tipo de impostor: I_K = {I_Rd, I_Sk}
        EER ← E(N_U, N_S, N_G, I_K)
```

FIGURA 4.5: Plan de experimentos con HMM-DU: se realizaron un total de 366.300 ($330 \times 111 \times 5 \times 2$) experimentos

4.5.2. Experimentos con HMM-DU

Con los modelos dependientes del usuario evaluamos la influencia de la selección de los parámetros estructurales N_S y N_G de forma particularizada a cada usuario. Los modelos fueron entrenados bajo las mismas condiciones experimentales anteriores para poder comparar las dos aproximaciones. Se evaluaron 555 diferentes configuraciones $M_{DU} = (N_S, N_G)_{111x5}$, variando el número de estados entre 1 y 111 y el número de gaussianas entre 1 y 5 (fig. 4.5). De todos los modelos probados se escogió el que cometía el menor error. En caso de que varios modelos produjeran el mismo error se seleccionó el de un menor número de estados. En esta segunda colección de pruebas se midió el error tanto con falsificaciones casuales como con firmas imitadas.

Los resultados obtenidos en estos experimentos se muestran en la siguiente sección.

4.6 Resultados

Mostraremos los resultados obtenidos clasificados del siguiente modo:

- **Resultados con HMM-IU**: obtenidos con todos los usuarios del corpus, falsificaciones casuales y modelos independientes del usuario.
- **Resultados con HMM-DU**: obtenidos con todos los usuarios del corpus, tanto con falsificaciones casuales como con firmas imitadas con modelos dependientes del usuario.
- **Resultados por subcorpus**: Resultados sobre los subcorpus de cada una de las universidades participantes en la adquisición de la base de datos MCYT. Sobre falsificaciones casuales y modelos dependientes de usuario.

4.6.1. Resultados con HMM-IU

La tabla 4.2 muestra los errores de las 49 configuraciones evaluadas de modelos independientes del usuario. Para cada una de ellas el número entre paréntesis representa el número de usuarios para el cual el modelo pudo ser correctamente inicializado mediante el algoritmo de entrenamiento de Baum-Welch. Como era previsible este número disminuía conforme aumentaba el número de grados de libertad del modelo debido a que las firmas no poseen suficiente cantidad de información para poder entrenar al modelo. De hecho las seis configuraciones con más grados de libertad no fueron válidas para ningún usuario, por lo que se muestran sus celdas vacías.

TABLA 4.2: EER con los modelos HMM-IU en escenario casual para distintas combinaciones de estados (Est.) y gaussianas. En negrita se muestra la mejor tasa de error válida para todos los usuarios.

	Gaussianas						
Est.	**1**	**2**	**4**	**8**	**16**	**32**	**64**
1	$36.43_{(330)}$	$35.19_{(330)}$	$33.08_{(330)}$	$31.88_{(330)}$	$30.19_{(330)}$	$28.64_{(330)}$	$28.29_{(324)}$
2	$35.55_{(330)}$	$33.93_{(330)}$	$31.77_{(332)}$	$30.11_{(324)}$	$28.21_{(306)}$	$27.61_{(287)}$	$29.92_{(260)}$
4	$34.80_{(330)}$	$32.40_{(327)}$	$29.84_{(319)}$	$27.88_{(310)}$	$26.76_{(277)}$	$29.53_{(210)}$	$37.47_{(90)}$
8	$31.11_{(330)}$	$29.71_{(330)}$	$26.80_{(312)}$	$25.83_{(287)}$	$27.28_{(241)}$	$35.91_{(104)}$	$42.59_{(7)}$
16	$24.20_{(330)}$	$23.74_{(321)}$	$22.38_{(306)}$	$22.90_{(254)}$	$30.47_{(99)}$	$55.88_{(9)}$	
32	$\mathbf{16.29}_{(330)}$	$16.27_{(309)}$	$15.85_{(259)}$	$16.36_{(107)}$	$21.04_{(8)}$		
64	$11.82_{(324)}$	$11.48_{(262)}$	$8.94_{(107)}$	$8.56_{(7)}$			

El primer dato mostrado en cada celda es la tasa de equierror (EER) y el segundo (entre paréntesis) el número de modelos que pudieron ser entrenados con dicha configuración. Como se ve existen muchas configuraciones para las que no pudo entrenarse correctamente el modelo para todos los usuarios debido a falta de datos.

El comportamiento anterior demuestra que un sistema de VAFD no sólo debe ser preciso, sino también robusto, ya que debe ser capaz de ser inicializado ante nuevas firmas no presentes en el corpus. En nuestro caso, el modelo con estructura válida para todos los usuarios del corpus que ha generado el mejor rendimiento es el de 32 estados y una gaussiana por estado, con un EER de 16.29 %. Sin embargo, otros modelos con el mismo número de grados de libertad no fueron válidos para todos los usuarios. Por ello, un sistema realista de VAFD basado en HMM-IU, puede llegar a no ser entrenado si el número de grados de libertad del modelo es muy elevado, por lo que para asegurarse validez universal debería optar por una configuración más simple, aunque ello repercutiera en el rendimiento.

La figura 4.6 muestra de forma gráfica los datos de la tabla 4.2. Se aprecia que fijado el número de estados del modelo existe un valor óptimo del número de gaussianas. Este comportamiento induce a pensar que existe una fuerte dependencia entre el rendimiento y la estructura del modelo. En particular parece que el aumento del número de estados da lugar

a mejores resultados que si aumentamos el número de gaussianas. Este mismo resultado lo obtiene Faundez-Zanuy (2007), donde los mejores resultados se consiguen también con una única gaussiana por estado, aunque con 12 estados por modelo.

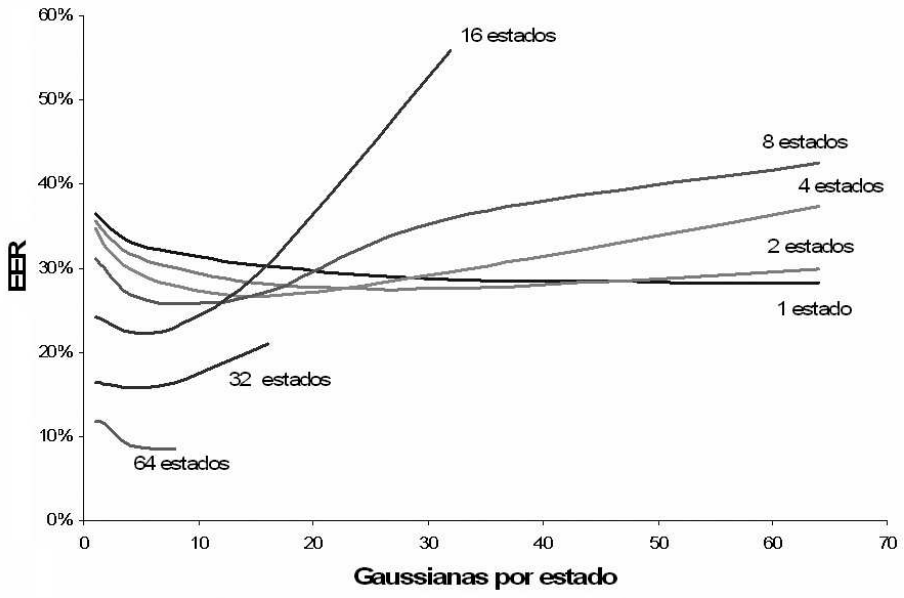


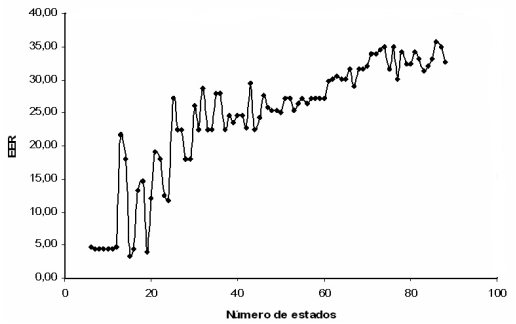
FIGURA 4.6: Error con los HMM-IU en función del número de gaussianas para distintos valores del número de estados

4.6.2. Resultados con HMM-DU en escenario casual

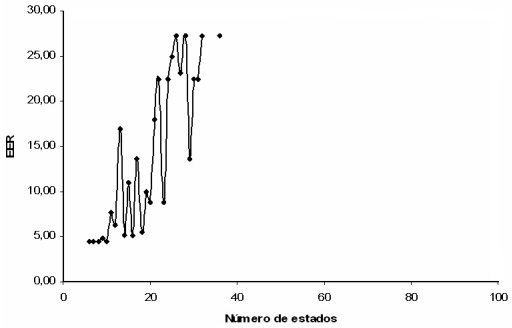
Antes de pasar a ver los resultados globales con HMM-DU veamos la dependencia entre los parámetros estructurales y el rendimiento para algunos usuarios. La figura 4.7 muestra la evolución del error para tres usuarios de la base de datos escogidos a modo de ejemplo. En ellas se representa el error (eje de ordenadas) respecto al número de estados del modelo (eje de abscisas) para tres valores prefijados del número de gaussianas (N_G: 1, 2 y 3). Como se observa la forma de las curvas es diferente para cada usuario, alcanzando mínimos de error para valores del número de estados diferentes. Vemos que mediante una adecuada selección del número de estados y gaussianas puede reducirse drásticamente el error de forma particularizada para cada usuario.

Vemos también que el valor límite de estados a partir del cual no es posible entrenar el modelo disminuye conforme aumenta el número de gaussianas por estado. En la figura 4.8 hemos representado la relación entre el valor del límite de estados respecto el número de gaussianas en promedio para todos los usuarios del corpus. Esta relación es bastante lineal pudiendo ser aproximada por la recta (4.5), que podría emplearse para acotar la búsqueda exhaustiva del modelo óptimo de cada usuario en un sistema real.

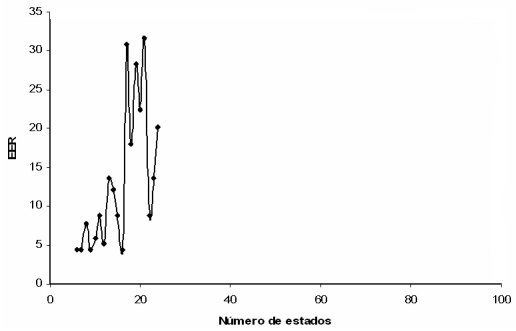
$$N_{Smax} = -15,3N_G + 125,9 \tag{4.5}$$



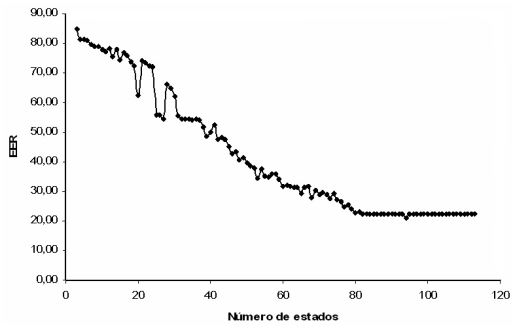
(a) Usuario 1, $N_G = 1$



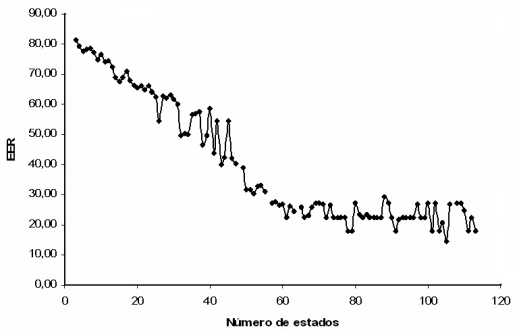
(b) Usuario 1, $N_G = 2$



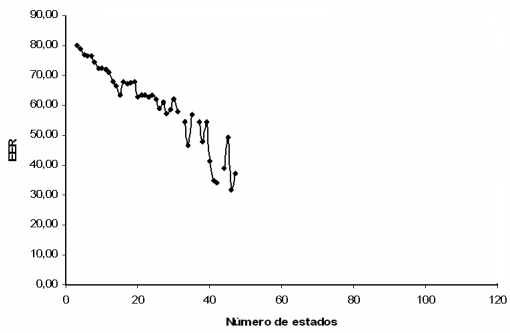
(c) Usuario 1, $N_G = 3$



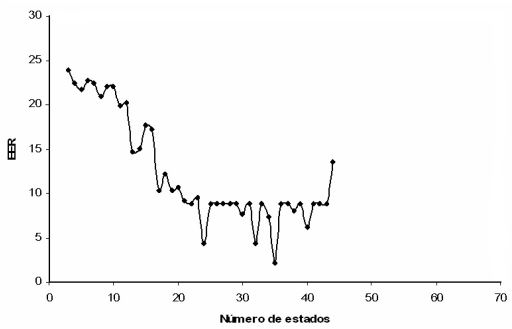
(d) Usuario 312, $N_G = 1$



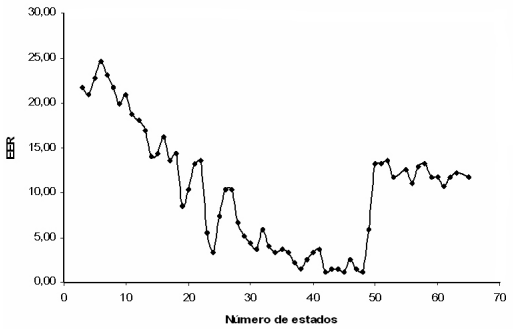
(e) Usuario 312, $N_G = 2$



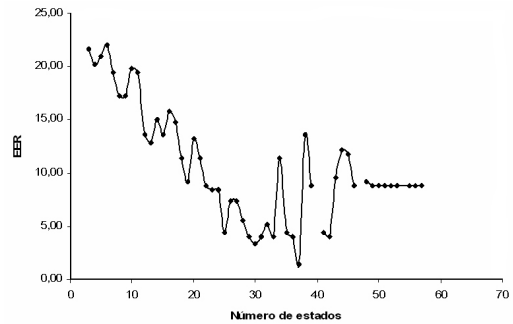
(f) Usuario 312, $N_G = 3$



(g) Usuario 431, $N_G = 1$



(h) Usuario 431, $N_G = 2$



(i) Usuario 431, $N_G = 3$

FIGURA 4.7: Evolución del error respecto al número de estados para tres valores del número de gaussianas (N_G)

Distribución del error respecto a los parámetros estructurales

Veamos ahora los errores obtenidos en las pruebas con los HMM-DU respecto a los valores de los parámetros estructurales. La tabla 4.10(a) muestra estas tasas de error en el escenario casual. Las cinco primeras filas son los errores medios cuando el número de gaussianas es prefijado, es decir dejando únicamente el número de estados como parámetro dependiente del usuario. La última fila de la tabla muestra el error medio de los modelos con valores óptimos tanto del número de estados como del de gaussianas. Mediante esta doble optimización se reduce el error en poco más de un 1 % en valor absoluto respecto al mejor de los casos con sólo optimización del número de estados.

La figura 4.9(b) representa la distribución del número óptimo de gaussianas. Se obtuvo que en el 61 % de los casos los modelos óptimos eran los que tenían pocas gaussianas por estado ($N_{Gopt} = 1, 2$) lo cual apunta a que la componente temporal del modelo, es decir el número de estados, tiene más poder discriminativo que la componente espacial (gaussianas).

La figura 4.9(c) representa la distribución del número de estados de los modelos óptimos. En este caso parece que no existe ninguna tendencia, sino que más bien da la sensación que el número limite de estados seleccionado (111) resultó escaso en la experimentación. Para la construcción del sistema propuesto más adelante en este capítulo este límite fue aumentado de forma suficiente.

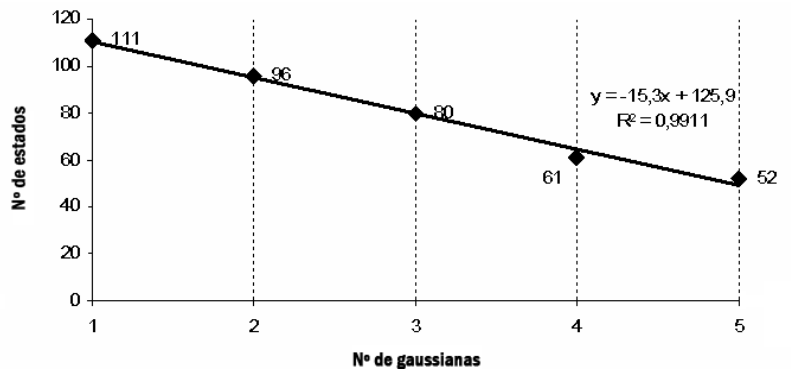


FIGURA 4.8: Promedio del número de estados límite para valores de $N_G = 1, 2, 3, 4, 5$

Distribución del error por usuario

La figura 4.9(d) muestra de forma más concisa el funcionamiento real de los modelos a nivel particular de cada usuario. Cabe destacar de ella lo siguiente:

a) Para un 28 % de los usuarios se obtuvo acierto pleno (EER=0 %).

b) Un 86 % de los usuarios tuvieron un EER menor del 5 %

c) Sólo existieron dos usuarios para los que el error fue mayor del 15 % siendo el peor resultado de toda la prueba de 15.81 %

d) El 80 % del error se concentró en el 31 % de los usuarios.

4.6.3. Resultados con HMM-DU en escenario seguro

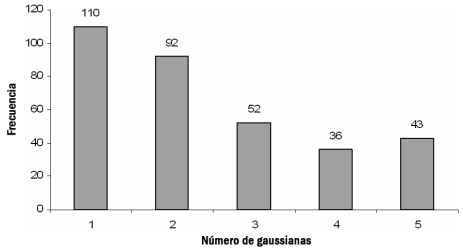
Como se ha comentado anteriormente la base de datos MCYT posee imitaciones de sus firmas por lo que se consideró probar la robustez de los HMM-DU en este tipo de escenario en el que prima la capacidad de rechazo antes posibles ataques. Los resultados obtenidos se reflejan en la figura 4.10.

En este escenario no existen grandes diferencias respecto a los resultados con impostor casual obteniéndose al igual que antes los mejores resultados con los modelos de pocas gaussianas por estado. Cabría destacar que el error cometido es ligeramente menor que con impostor casual. Este fenómeno (ciertamente extraño) puede deberse a las características utilizadas ya que algunas de ellas, sobre todo las componentes angulares, se ha comprobado que empeoran la capacidad de reconocimiento de la firma. Es de esperar que con una mejor selección de características se obtenga un menor error en el escenario casual que en el escenario seguro, como así ocurrió en el sistema propuesto al final de este capítulo.

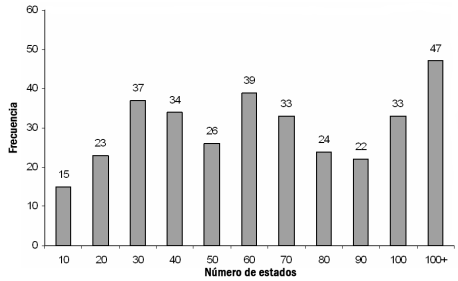
En las gráficas de la figura 4.10 se muestran los histogramas de gaussianas y de estados óptimos junto a la distribución de errores por usuario en este escenario. El histograma de estados muestra que parece que en este escenario existe cierta tendencia a que los modelos óptimos tengan un número de estados menor que con impostor casual. Esto mismo se refleja en el histograma de gaussianas (4.10(c)), que en este caso son en promedio más

N_G	% **EER**
1	3.83
2	3.46
3	4.08
4	4.63
5	4.96
$N_{Gopt.}$	**2.33**

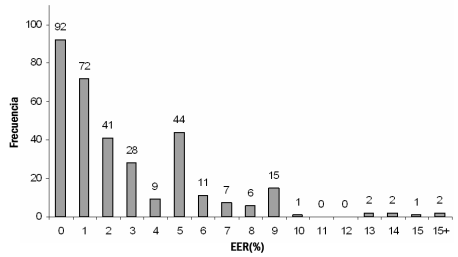
(a) Promedio de errores por nº de gaussianas



(b) Hist. de N_G óptimas. Media = 2.43



(c) Hist. de N_S óptimos. Media = 59.89



(d) Hist. de errores por usuario

FIGURA 4.9: Resultados con HMM-DU en escenario casual

TABLA 4.3: Resultados con HMM-DU por Universidad participante en la adquisición de MCYT (pruebas con impostor casual y HMM-DU)

Universidad	**Duración (sg)**	**% EER**
UVA	3.87	1.62
EUPMT	3.09	1.74
UPM	3.37	2.50
EHU	7.83	3.01
Media	4.45	2.33

numerosas que en el anterior escenario, debido al menor número de estados de los modelos. La explicación a este fenómeno puede deberse a que la inclusión de parámetros tales como la presión y las componentes angulares aportan inestabilidad al reconocimiento de la firma propia, pero ayudan a discriminar las firmas imitadas. Dicha capacidad de discriminación parece que se concentra en el dominio espacial (gaussianas) de manera más pronunciada de lo que ocurría en el escenario casual.

Respecto a la distribución de los errores cometidos entre los distintos usuarios lo más reseñable es el alto número de modelos que no cometieron ningún error en la verificación (un 69 % frente al 28 % del caso de impostor casual). En este caso el 80 % del error se acumula en tan solo el 21 % de los usuarios.

En el siguiente apartado veremos los resultados diseccionados por origen de las firmas.

4.6.4. Resultados por subcorpus

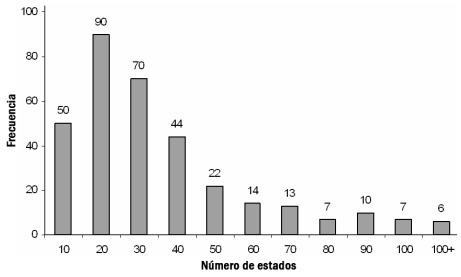
La tabla 4.3 recoge los resultados de las pruebas teniendo en cuenta el origen de los datos, que en el caso del corpus MCYT son las cuatro Universidades que participaron en su adquisición.

Es de destacar la diferencia entre el menor error obtenido con las firmas recogidas en UVA y el mayor error correspondiente a EHU que casi duplica el error anterior. Examinando las firmas de este último observamos que en promedio son considerablemente más largas en duración que las del resto de Universidades. Ello nos indujo a representar los errores en función de la duración temporal de la firma resultando la gráfica de la figura 4.11[1]. Se observa que las firmas de menor duración son las que dan lugar a menos errores. Creemos puede deberse a que si la firma se realiza de forma suficientemente rápida, normalmente presentará mayor consistencia y estabilidad que aquellas de mayor duración temporal en la que la variabilidad intrausuario puede ser mayor.

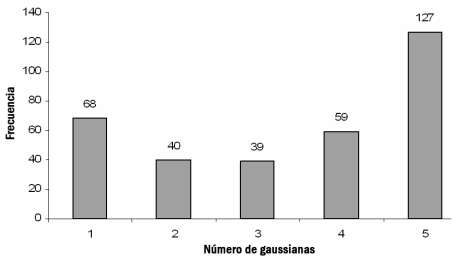
[1]Cada punto de la gráfica representa la media del error en grupos de 30 usuarios previa ordenación de todos los usuarios del corpus por duración de su firma.

N_G	**% EER**
1	3.29
2	3.29
3	3.42
4	3.71
5	3.59
$N_{Gopt.}$	**2.06**

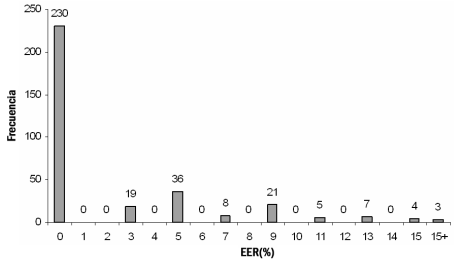
(a) Promedio de errores por nº de gaussianas



(b) Hist. de N_S óptimo. Media = 30.83



(c) Hist. de N_G óptimo. Media = 3.41



(d) Hist. de errores por usuario

FIGURA 4.10: Resultados con HMM-DU en escenario seguro

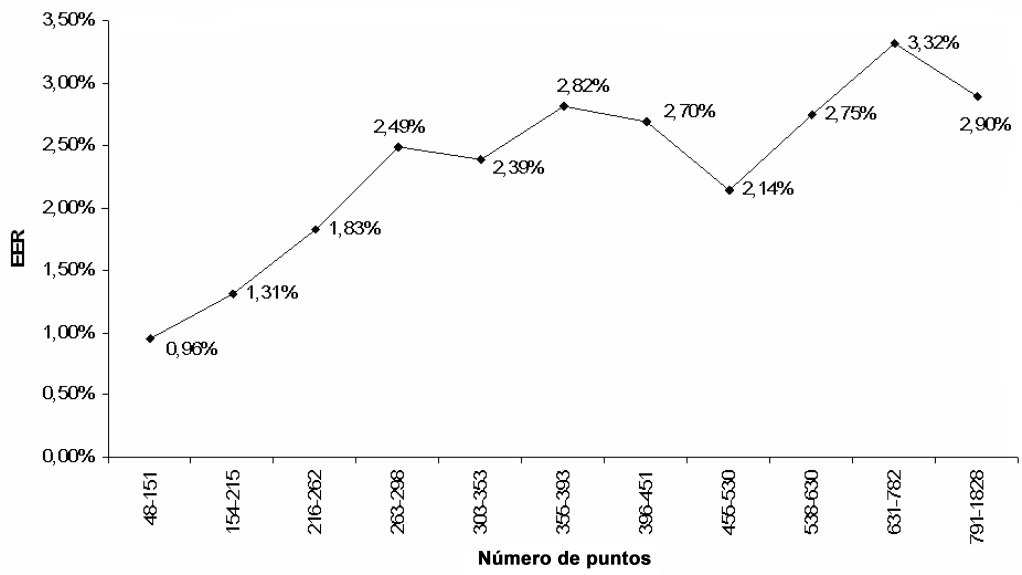


FIGURA 4.11: Error en función de la duración temporal (número de puntos) de la firma

4.7 Comparativa con otros trabajos

Antes de pasar al sistema propuesto comentaremos algunos de los resultados obtenidos por otros sistemas de VAFD basados en HMM-IU.

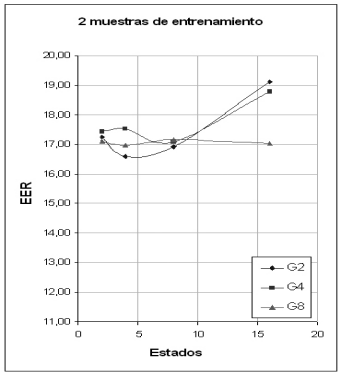
Moneo-Agapito (2005) usó las mismas características y normalización geométrica que las empleadas en nuestro estudio con HMM-DU. También estudió la influencia del número de firmas de entrenamiento en el rendimiento del sistema (con 2, 4 y 6 firmas -R2, R4 y R6 en la tabla 4.4-) así como la influencia de añadir a las características originales sus primeras y segundas derivadas temporales. La tabla 4.4 y su representación gráfica (fig. 4.12) muestra sus resultados obtenidos para las distintas combinaciones de estados y gaussianas.

Estos resultados muestran que al añadir las primeras derivadas temporales se obtuvo una mejora en el rendimiento: un 10 % con seis muestras de entrenamiento, 5 % con cuatro y 2 % con dos. Ello corrobora el hecho de que el rendimiento de nuestro sistema basado en los HMM-DU podía ser mejorado seleccionando un conjunto de características más adecuado.

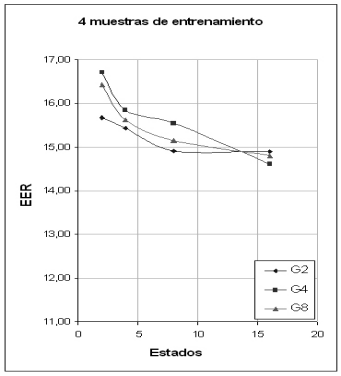
En la figura 4.12 vemos también la dependencia entre el rendimiento de los modelos y el número de estados. Este hecho es válido para "casi" cualquier número de muestras de entrenamiento (únicamente cuando se utilizaron sólo dos muestras de entrenamiento, el sistema no mejoró y sólo en los casos con 5 y 15 características). En dicha figura se ve claramente la tendencia que muestra que si el autor hubiera continuado aumentando el número de estados del modelo habría obtenido mejores resultados.

Respecto al sistema de VAFD descrito por Fierrez-Aguilar (2006), hay que decir que también hace uso de HMM-IU como elemento base del clasificador, con la diferencia de que en este trabajo se realiza un estudio previo del sistema en función de las características utilizadas.

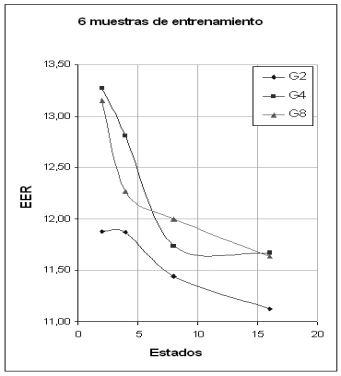
En concreto, con la misma configuración inicial formada por los cinco parámetros



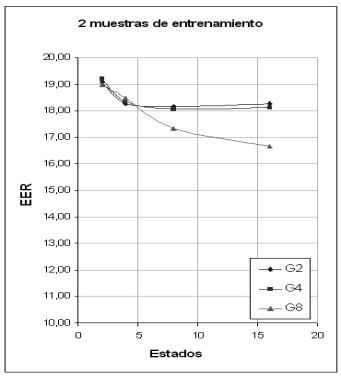
(a) 5 características



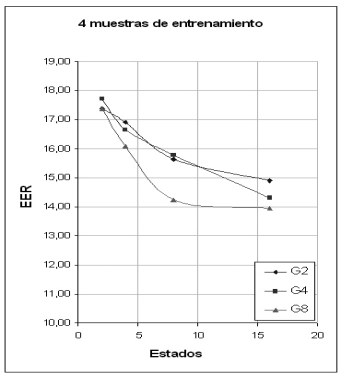
(b) 5 características



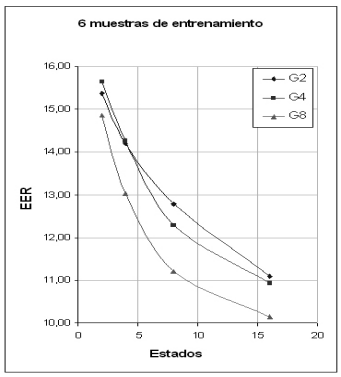
(c) 5 características



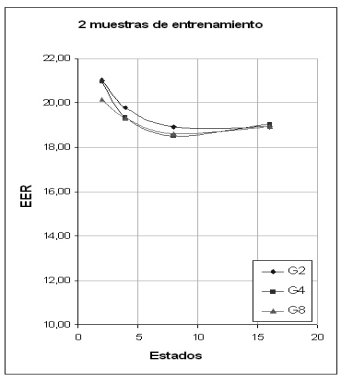
(d) 10 características



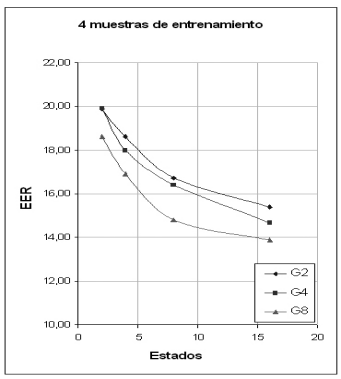
(e) 10 características



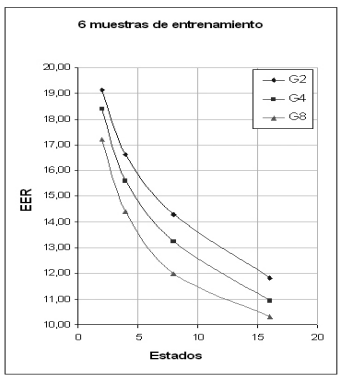
(f) 10 características



(g) 15 características



(h) 15 características



(i) 15 características

FIGURA 4.12: Representación gráfica de los datos de la tabla 4.4

TABLA 4.4: Resultados obtenidos con HMM-IU (Moneo-Agapito, 2005)

		5 parámetros			10 parámetros			15 parámetros		
N_S	N_G	R2	R4	R6	R2	R4	R6	R2	R4	R6
2	2	17.24	15.68	11.88	19.07	17.38	15.37	21.03	19.89	19.13
	4	17.44	16.71	13.27	19.20	17.72	15.64	20.97	19.92	18.39
	8	17.10	16.43	13.15	19.00	17.39	14.85	20.17	18.63	17.21
4	2	16.60	15.43	11.87	18.27	16.91	14.20	19.78	18.63	16.61
	4	17.54	15.85	12.81	18.33	16.65	14.27	19.37	17.99	15.59
	8	16.97	15.62	12.27	18.48	16.09	13.02	19.33	16.92	14.40
8	2	16.93	14.92	11.44	18.15	15.63	12.79	18.93	16.73	14.30
	4	17.09	15.54	11.74	18.05	15.78	12.29	18.51	16.40	13.24
	8	17.15	15.15	12.00	17.33	14.23	11.20	18.61	14.79	11.99
16	2	19.11	14.90	11.13	18.27	14.91	11.10	18.92	15.41	11.82
	4	18.79	14.61	11.67	18.12	14.32	10.93	19.04	14.67	10.93
	8	17.03	14.80	11.64	16.67	13.95	10.14	18.95	13.88	10.31

(x, y, p, az, in) utilizada en nuestro trabajo (más las primera derivada temporal de cada una de ellas), el sistema obtiene un EER de 6.25 % con cinco muestras de entrenamiento. Al eliminar las componentes angulares az, in y añadir cuatro nuevas características locales (junto a sus primeras derivadas temporales), el error se redujo casi a una décima parte del anterior (hasta llegar a un 0.68 %). Esto demuestra que la selección de las características utilizadas junto a su preprocesamiento es crucial en el rendimiento del sistema.

4.8 Sistema propuesto para entornos prácticos

Una vez demostrada la eficacia de la personalización de la estructura de los HMM al usuario, nuestro siguiente objetivo era la construcción de un sistema basado en este nuevo enfoque. En su diseño sacaríamos partido de lo aprendido en los experimentos ya realizados y plantearíamos otros nuevos para mejorar el rendimiento.

Hasta ahora hemos demostrado que:

- La selección de los parámetros estructurales mejora notablemente el rendimiento de los HMM, aunque el sistema debe ser perfeccionado mediante una más cuidada selección de características.
- Los mejores resultados se obtuvieron con un número de gaussianas bajo, mientras que el número de estados límite quedó demostrado que podría haberse aumentado para mejorar el rendimiento.

Además, el sistema debe seguir siendo fiel a nuestros objetivos de aplicabilidad práctica, por lo que el número de firmas de entrenamiento se seguirá manteniendo en tres por usuario.

Se plantearon una nueva serie de experimentos con el fin de reducir la tasa de error obtenida con el sistema de referencia anterior. A continuación describimos estos nuevos experimentos.

4.8.1. Optimización preliminar del sistema

Es bien sabido que una cuidadosa selección de las características del rasgo biométrico junto a una adecuada normalización de las mismas puede mejorar drásticamente el rendimiento de un sistema biométrico sin modificar el módulo de clasificación.

Para la selección y procesamiento de características de nuestro sistema nos basaremos en los trabajos de otros autores (Fierrez-Aguilar *et al.*, 2005c; Kholmatov & Yanikoglu, 2005), que en el caso de la firma dinámica ya habían apuntado a que las características angulares del lápiz (acimut e inclinación) empeoraban el rendimiento por lo que las descartamos.

Kholmatov & Yanikoglu (2005) demostraron que era posible obtener excelentes resultados utilizando únicamente las características geométricas básicas (x, y) lo cual estaba en la línea de nuestros objetivos de aplicabilidad práctica, al ser un conjunto de características muy sencillo y con pocas necesidades de almacenamiento. Además, se trata de un conjunto de características que puede considerarse universal, ya que todo dispositivo hardware utilizado para la captura de firma dinámica debería al menos proporcionar las coordenadas geométricas de la trayectoria en función del tiempo. Por todo ello decidimos evaluar nuestro sistema utilizando únicamente las características geométricas (x, y) de la firma. A dicho conjunto de características geométricas se añadió la primera derivada temporal.

Selección del número de gaussianas. Establecido el nuevo conjunto de características evaluamos el impacto en el rendimiento de los dos parámetros estructurales en estudio aumentando los límites experimentales anteriores. El objetivo era ver cuál de los dos parámetros estructurales era más relevante en el caso de la firma con el nuevo conjunto de características. Para ello, seleccionamos un subconjunto de cohorte de MCYT formado por 50 usuarios (MCYT-50) y realizamos una evaluación en escenario casual variando el número de gaussianas entre 1 y 50 y el número de estados entre 1 y 350. El procedimiento seguido fue el mismo que para el sistema de referencia. La figura 4.13 muestra los resultados obtenidos en estos experimentos. Se observa que con este nuevo conjunto de características el error crece a medida que el número de gaussianas aumenta, siendo ahora el modelo formado por una única gaussiana por estado el que proporcionaba los mejores resultados. Decidimos entonces fijar el número de gaussianas a una por estado y de este modo reducir el proceso de personalización del modelo a la búsqueda del número de estados apropiado para cada usuario.

Estos experimentos de optimización del número de gaussianas fueron realizados con la normalización del sistema de referencia (ec. 4.6).

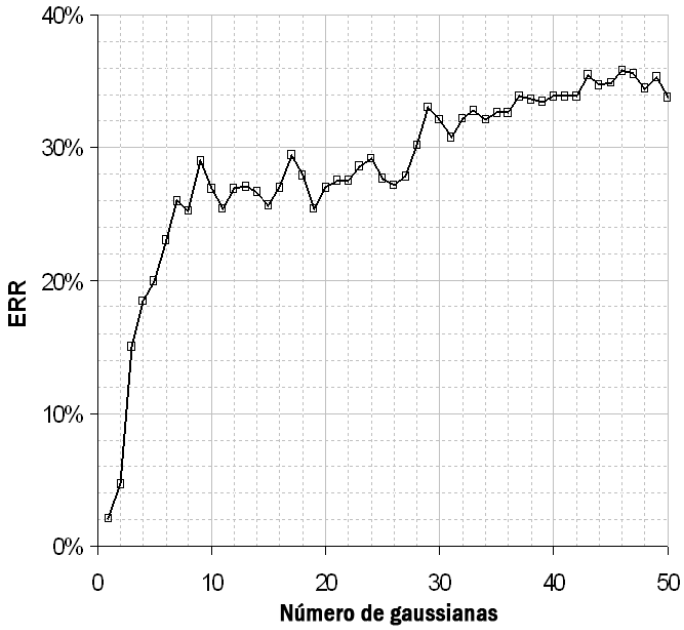


FIGURA 4.13: Evolución del error (EER) en función del número de gaussianas por estado con MCYT-50

$$\overrightarrow{F}_{N1} = \begin{cases} x'_t = (x_t - x_{min}) \\ y'_t = (y_t - y_{min}) \\ \Delta x_t = x_{t+1} - x_{t-1} \\ \Delta y_t = y_{t+1} - y_{t-1} \end{cases} \tag{4.6}$$

Normalización de características. Una vez decidido el conjunto de características y el número de gaussianas por estado se analizaron en más detalle varias normalizaciones geométricas para intentar mejorar el rendimiento del sistema. Según puede verse en la tabla 4.5 los mejores resultados se obtuvieron con la normalización XY_{ob} por ajuste al punto inicial de la firma con un valor de 1.65 % de EER. Tras añadir la primera derivada temporal el resultado mejoró hasta obtener un EER de 0.63 %. La ecuación 4.7 muestra la normalización geométrica adoptada finalmente en donde S_w y S_h representan el ancho y alto de la firma respectivamente.

$$\overrightarrow{F}_{N2} = \begin{cases} x'_t = (x_t - x_0)/S_w \\ y'_t = (y_t - y_0)/S_h \\ \Delta x_t = x_{t+1} - x_{t-1} \\ \Delta y_t = y_{t+1} - y_{t-1} \end{cases} \tag{4.7}$$

Con todas estas optimizaciones preliminares redujimos la personalización del modelo al número de estados. En los siguientes apartados describiremos los resultados obtenidos

TABLA 4.5: Normalizaciones geométricas evaluadas en la construcción del sistema basado en HMM-DU

Norm.	**Descripción**	**%EER**
$XsYs$	Normalización en traslación mediante substracción de mínimos	2.91
$XgYg$	Normalización en traslación al centro geométrico	8.27
$XoYo$	Normalización en traslación al punto inicial	2.15
$XobYob$	$XoYo$ + Normalización en magnitud del ancho y alto	1.65
$XobrYobr$	$XobYob$ + normalización en rotación	2.23
$XobYob\_D$	$XobYob$ + primeras derivadas temporales	0.63

TABLA 4.6: Errores del sistema HMM-DU con selección de número de estados *a posteriori*

	EER(%)	
Resultados	**Escenario casual**	**Escenario seguro**
Deciles 1 a 6	0.00	0.00
Decil 7	0.00	2.00
Decil 8	0.24	6.27
Decil 9	4.53	12.82
% de usuarios sin error	74.77	65.47
EER máximo	9.97	29.91
EER medio	**0.72**	**3.07**

con selección del número de estados tanto *a posteriori* como *a priori*.

4.8.2. Selección a posteriori del número de estados

Para determinar el número de estados del modelo óptimo del usuario evaluamos un total de 500 modelos por usuario, todos ellos con una única gaussiana por estado, variando el número de estados entre 1 y 500. El límite superior de estados elegido (500) se puede decir que es un valor considerablemente alto hablando de HMM, y de hecho para algunas firmas ese límite no pudo ser alcanzado al no disponer de suficiente cantidad de datos. El valor medio de este límite superior de estados fue de 321.4 estados.

Cada modelo fue entrenado con tres firmas del usuario obteniéndose los errores tanto en escenario casual como escenario seguro por separado. Para medir el error se siguió la misma metodología que en los experimentos anteriores. La tabla 4.6 muestra los errores en cada escenario.

Mostramos la distribución de errores en forma por usuario de deciles para dar más información del comportamiento del sistema. La primera fila representa el error de verificación del primer al sexto decil, y como puede observarse en todas ellas no existe error

TABLA 4.7: Comparación entre el sistema HMM de referencia y el optimizado

Escenario	%EER Base	%EER Optimizado	% Reducción del error
Casual	3.46	0.72	**79.19**
Seguro	3.29	3.07	**6.69**

en ninguno de los dos escenarios. Esto quiere decir que para más del 60 % de los usuarios probados no hubo ningún error de verificación. En el caso de firmas casuales este valor es incluso mayor llegando al 74.77 % mientras que con firmas imitadas fue del 65.47 %. Las filas segunda a cuarta muestran como se distribuye el error entre el resto de usuarios. Los valores de la fila sexta son los errores máximos del usuario con peor rendimiento, siendo destacable el caso de firmas casuales en el que se obtuvo un error máximo de menos del 10 %.

Con las optimizaciones realizadas hemos reducido el error del sistema respecto a nuestro sistema de referencia en una cantidad notable (79.19 %) en escenario casual y en 6.69 % en escenario seguro (tabla 4.7). Achacamos estas diferencias principalmente a las características seleccionadas las cuales parece que mejoran la capacidad de detectar la singularidad de la firma, aunque no evita en tanta medida su 'imitabilidad'. Este resultado sugiere que para desplegar un sistema como el nuestro en escenarios seguros podrían añadirse al conjunto de características base otras características adicionales como la presión y su primera derivada temporal, aunque ello mermara en algo la capacidad de reconocimiento de firmas genuinas.

4.8.3. Selección a priori del número de estados

En este apartado mostraremos un método de selección *a priori* del número de estados del modelo, y cual es el rendimiento del mismo. Como se ha expuesto en el apartado 4.2 otros autores también realizan selección del número de estados más o menos dependiente del usuario, aunque en ninguno de ellos se da una explicación demasiado detallada del proceso de selección ni los justifican de forma sistemática. En nuestro caso, al disponer de la configuración óptima de los modelos obtenida en la prueba anterior, representamos el número de estados óptimo en función de la duración media de las firmas de entrenamiento del usuario, obteniendo la gráfica de la figura 4.14.

Puede verse una fuerte dependencia entre las dos variables representadas ($r^2 = 0{,}989$) lo cual indica que, dada la duración media de las firmas registradas del usuario, podría aplicarse una simple regresión lineal para obtener el número de estados cercano a su valor óptimo. Utilizando esta técnica de estimación *a priori*, los valores medios de EER cometidos por el sistema son de 2.09 % y 6.14 % para los escenarios casual y seguro respectivamente, lo que supone aproximadamente tres veces el error *a priori*. Este resultado sugiere que aunque la correlación entre estados y duración de la firma es alta, pequeñas variaciones alrededor del óptimo producen importantes pérdidas de precisión. Sin embargo los resultados finales consideramos que siguen siendo interesantes a nivel del estado del arte sobre todo teniendo en cuenta el número reducido de firmas de entrenamiento y las pocas características utilizadas (sólo cuatro). Todos los resultados *a priori* se muestran de

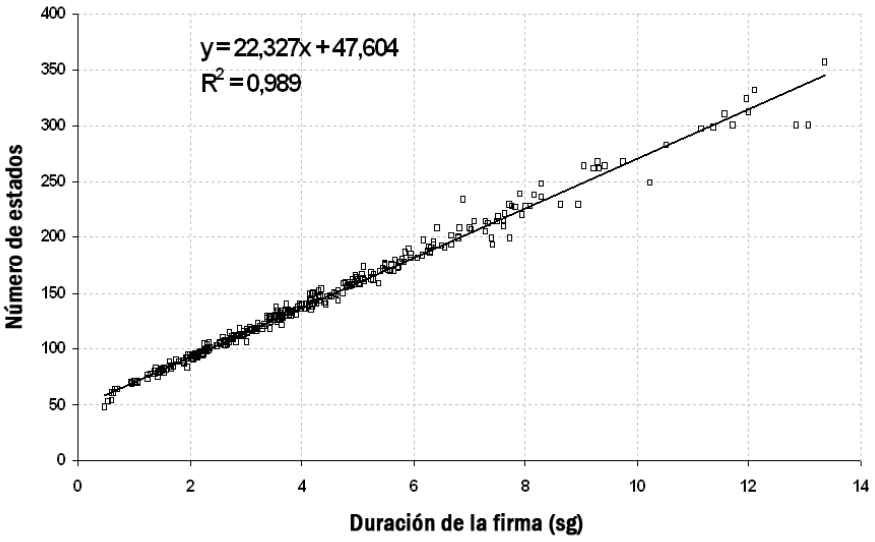


FIGURA 4.14: Valores *a posteriori* del número de estados óptimo para cada usuario de MCYT en función de la duración media de su firma

forma más detallada en la tabla 4.8 la cual tiene la misma interpretación que la tabla 4.6 ya comentada anteriormente.

Una representación más visual de la dependencia entre estados y la duración de la firma puede verse en la figura 4.15. En ella algunas firmas con diferentes complejidades son representadas junto a su número de estados óptimo (*a posteriori*).

4.9 Resumen

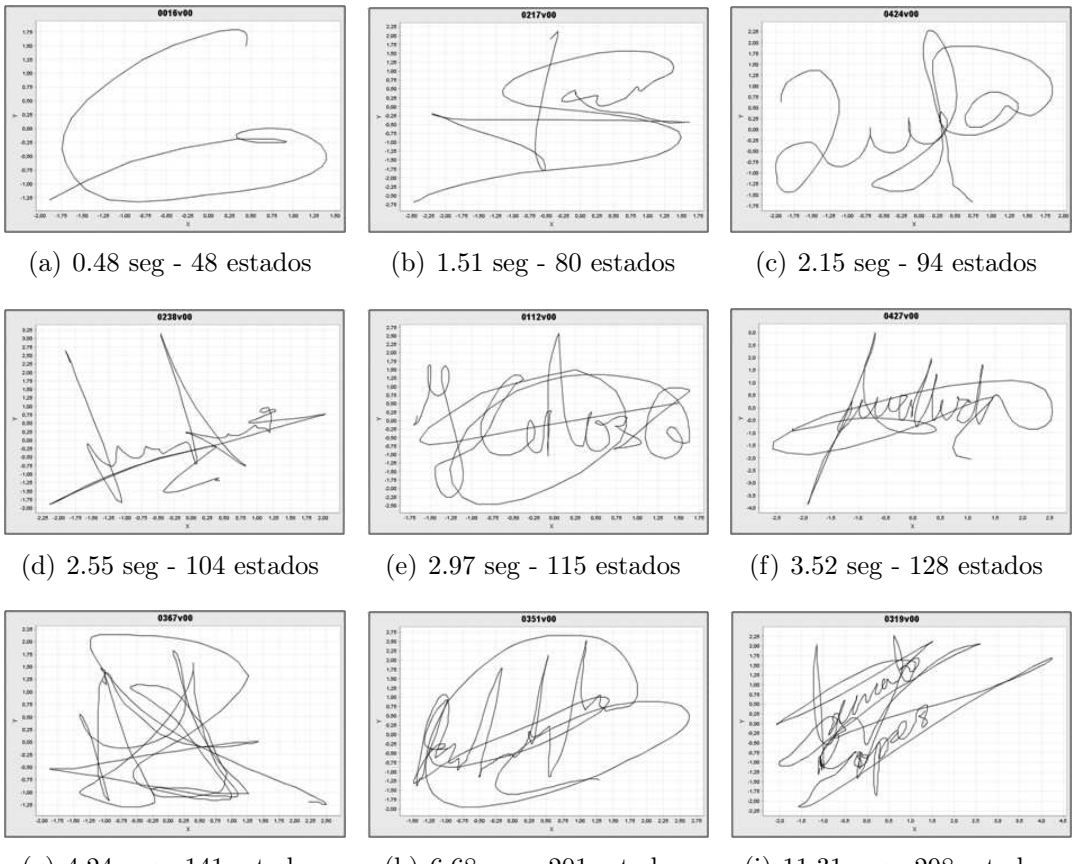
En este trabajo hemos proporcionado evidencias experimentales de que una optimización de la estructura de los modelos ocultos de Markov dependiente del usuario puede dar lugar a mejoras significativas de rendimiento en los sistemas de VAFD basados en HMM. Se ha estudiado la influencia de dos parámetros estructurales relevantes: el número de estados del modelo y el número de gaussianas por estado. La optimización del primero de ellos fue el que nos proporcionó las mayores mejoras.

Se ha introducido por primera vez de forma explícita el concepto de HMM dependiente de usuario -a nivel estructural- (HMM-DU). Estos modelos han proporcionado más rendimiento que los modelos independiente del usuario (HMM-IU) usando un número reducido de firmas de entrenamiento. Nuestro enfoque al problema rompe además con la tradición de otros autores en los que para obtener tasas de error competitivas con sistemas basados en HMM se necesita un conjunto de características amplio, normalmente formado por características locales y globales. Todos estos factores suponen una ventaja del sistema que proponemos de cara a aplicarlo a soluciones comerciales.

Nuestro primer sistema de referencia generó errores de 2.33 % en escenario casual y 2.06 % en escenario seguro. Estos valores representan una reducción de error en un factor

TABLA 4.8: Resultados obtenidos con selección de modelo *a priori*

Resultados	**EER-RAN(%)**	**EER-SKI(%)**
Deciles 1 a 3	0.00	0.00
Decil 4	0.00	2.00
Decil 5	0.00	4.27
Decil 6	0.30	4.27
Decil 7	1.51	8.55
Decil 8	4.53	10.55
Decil 9	5.41	17.09
% de usuarios con EER = 0 %	50.51	39.04
EER máximo	27.94	44.73
EER medio	**2.09**	**6.14**



(a) 0.48 seg - 48 estados (b) 1.51 seg - 80 estados (c) 2.15 seg - 94 estados

(d) 2.55 seg - 104 estados (e) 2.97 seg - 115 estados (f) 3.52 seg - 128 estados

(g) 4.24 seg - 141 estados (h) 6.68 seg - 201 estados (i) 11.31 seg - 298 estados

FIGURA 4.15: Muestras de firmas con diferentes complejidades junto a su número de estados óptimo

de 6 sobre el mejor resultado con HMM-IU en las mismas condiciones experimentales. Demostrada entonces la conveniencia de adaptar los HMM al usuario hemos optimizado el sistema de referencia para mejorar su rendimiento. El sistema final que hemos propuesto necesita menos muestras de entrenamiento que otros sistemas de la literatura y un menor número de características. Sólo utiliza cuatro características básicas, las coordenadas geométricas y sus primeras derivadas temporales $x, y, \Delta x, \Delta y$, las cuales tienen la ventaja añadida de poder ser obtenidas desde virtualmente cualquier dispositivo de captura de firma.

Una desventaja inicial de los HMM-DU es la falta de un método de selección de la estructura del modelo *a priori*, es decir, antes de su puesta en funcionamiento. Hemos aportado una posible solución a este problema en dos pasos. Primero mediante la restricción del número de gaussianas del modelo a una por estado y en segundo lugar seleccionando el número de estados mediante regresión lineal a partir de la duración temporal de la firma. Aunque ya existían publicaciones en las que se proponía la estimación del número de estados a partir de la duración de la firma (Fuentes *et al.*, 2002; Shafiei & Rabiee, 2003) en ellas la estimación es de tipo heurística. En este trabajo hemos realizado un estudio bajo condiciones experimentales realistas.

Con el método propuesto para seleccionar el número de estados se produjo un incremento del error desde 0.72 % a 2.09 % para firmas casuales y desde 3.07 % a 6.14 % con firmas imitadas. A pesar de ello consideramos el aumento aceptable ya que los errores siguen estando a nivel del estado del arte sobre todo teniendo en cuenta los mínimos requerimientos del sistema. Los resultados principales con el sistema final optimizado a nivel de selección y preprocesamiento de características se muestran en la tabla 4.9.

En el siguiente capítulo presentaremos el segundo sistema desarrollado en esta tesis. Dicho sistema está basado en la aproximación metodológica consistente en el alineamiento de características. En concreto hemos utilizado el algoritmo DTW, que puede considerarse el algoritmo de referencia del statu quo actual en reconocimiento de firma.

TABLA 4.9: Resumen de resultados del sistema de VAFD basado en HMM-DU expuesto en este capítulo

Sistema	EER (esc. casual)	EER (esc. seguro)
A posteriori	0.72 %	3.07 %
A priori	2.09 %	6.14 %

5 Alineamiento Temporal Dinámico

El Alineamiento Temporal Dinámico (DTW, Dynamic Time Warping) puede considerarse actualmente como el método de referencia en el campo del reconocimiento biométrico basado en firma dinámica. Los resultados de las dos competiciones internacionales realizadas sobre firma dinámica en el año 2004 (Yeung *et al.*, 2004) y 2009 (Dorizzi *et al.*, 2009) así lo indican.

En este capítulo presentamos un sistema de verificación de firma dinámica basado en DTW. Las principales aportaciones del sistema se encuentran a nivel de tratamiento de las características ya que no se ha modificado el algoritmo DTW respecto a su versión clásica. Se ha realizado un estudio de cómo afectan al rendimiento las características utilizadas, pero no sólo a nivel individual, sino (sobre todo) al combinarlas entre sí.

En primer lugar, describimos los distintos módulos que componen el sistema propuesto. Tras ello, analizamos el rendimiento del sistema con distintas combinaciones de características obtenidas con métodos clásicos de reconocimiento de patrones. Proporcionamos resultados con combinaciones de características óptimas tanto universales o independientes del usuario, como particulares o dependientes del usuario. Finalmente, el estudio se finaliza con una evaluación del sistema propuesto sobre cuatro bases de datos de firmas dinámicas y una comparativa con otros sistemas publicados que fueron evaluados con dichas bases de datos.

Parte de los resultados que se verán en este capítulo han sido publicados (Pascual-Gaspar *et al.*, 2008; 2009).

5.1 Introducción

Aunque el algoritmo DTW básico es bien conocido desde hace décadas (Rabiner *et al.*, 1978) el desarrollo de un sistema de VAFD basado en él no es por ello trivial.

Con el fin de maximizar el rendimiento obtenido con el algoritmo DTW manteniendo las restricciones de uso práctico, hemos desarrollado un sistema de acuerdo al siguiente plan:

1. Utilizaremos sólo cinco firmas de referencia por cliente, por ser éste un valor que sigue siendo poco gravoso durante el registro del usuario, pero manteniendo un buen rendimiento de reconocimiento (Fierrez & Ortega-Garcia, 2007).

2. Las características de la firma a utilizar serían sencillas y simples de calcular por lo que partiríamos de las obtenidas directamente de la tableta gráfica.

3. Para reducir las necesidades de almacenamiento a las características anteriores añadiríamos únicamente su primera y segunda derivada temporal.

4. Cada tipo de característica se ponderaría por igual dentro del vector al que pertenece, por lo que se realizaría una normalización estadística inicial, además de una normalización geométrica sobre las características de tipo posicional.

5. Buscaríamos un método para combinar las características de partida de forma eficaz con lo que obtener combinaciones óptimas tanto independiente como de forma dependiente del usuario.

6. Para adaptar el sistema a las condiciones de seguridad (escenario casual o escenario adverso) seleccionaríamos también combinaciones de características óptimas en función de dichas condiciones.

7. Desarrollaríamos un método de normalización de puntuaciones con el que minimizar la pérdida de rendimiento al evaluar el error con umbral universal en vez de con umbral individual. Este método será desarrollado para el sistema enviado a la competición BSEC 2009, que será descrito en la sección 6.2.5.

5.2 Diseño del sistema

Esta sección describe los cuatro módulos que forman el sistema de acuerdo al esquema de Jain (Jain *et al.*, 2004b).

5.2.1. Sensor

Existen muchos dispositivos (lápices digitales, asistentes personales digitales (PDAs), Tablet-PCs, ...) que pueden ser usados para capturar la firma dinámica de un usuario. Sin embargo, las tabletas gráficas proporcionan la mejor resolución espacial y temporal a un precio asequible. Por ello, muchas bases de datos de firma dinámica han sido adquiridas mediante tabletas gráficas y, en particular, así ha sido con las empleadas en este trabajo.

No obstante, tal y como se verá más adelante, los resultados obtenidos son fácilmente trasladables a otros dispositivos debido a la sencillez del sistema propuesto.

Centrándonos en las tabletas gráficas, éstas registran información gestural del firmante en una secuencia de vectores a frecuencia típica de 100 Hz. Las características registradas por ellas pueden ser clasificadas en dos categorías (fig. 3.1 del capítulo 3):

- **Características posicionales**: los puntos 2D seguidos por el lápiz a lo largo de la trayectoria de la firma.

$$f_k^p \epsilon \{x_k, y_k\} \tag{5.1}$$

- **Características ergonómicas**: resultantes de la interacción entre mano y lápiz. En nuestro caso son la *presión* p ejercida por el lápiz sobre la tableta y los ángulos de inclinación y acimut (*acimut* a e *inclinación* i).

$$f_k^e = \{p_k, a_k, i_k\} \tag{5.2}$$

5.2.2. Extracción de características

Una firma dinámica S es representada mediante una secuencia temporal finita de N vectores de características, donde N depende de la duración temporal real de la firma y de la frecuencia de muestreo del dispositivo de captura (ec. 5.3):

$$S = \{(x_t, y_t, p_t, a_t, i_t)\}_{t=1..N} = \{f_{k,t}\}_{t=1..N}^{k=1.,5} \tag{5.3}$$

A la hora de seleccionar las componentes del vector de características podemos distinguir tres posibles estrategias:

1. Usar directamente las características capturadas por el sensor (Faundez-Zanuy, 2007; Vivaracho-Pascual *et al.*, 2009). Aunque este enfoque no excluye algunas tareas de preprocesamiento, el mayor énfasis del sistema recae típicamente en el módulo clasificador.

2. Generar características extendidas a partir de las básicas y seleccionar las más efectivas apoyándose en los resultados ya publicados por otros autores y en la propia experimentación. Esta solución puede dar lugar a sistemas de alto rendimiento pero normalmente tendrán un coste computacional más elevado, al tener que generar conjuntos de características extra (de mayor o menor complejidad). Además, tienen el riesgo de dar lugar a implementaciones dependientes de la base de datos utilizadas (Jain *et al.*, 2005).

3. Seleccionar una combinación de las características básicas y sus derivadas temporales (Kholmatov & Yanikoglu, 2005). Esta alternativa intermedia entre las dos anteriores es idónea para sistemas que operen en tiempo real y da lugar a soluciones más independientes de las características concretas de la base de datos (y por ende implementables en un mayor número de dispositivos).

En nuestro sistema el conjunto de características inicial (ec. 5.4) fue extendido para incluir la primera y segunda derivada temporal, formando un vector de características de partida de 15 componentes (ec. 5.5).

$$F = (x, y, p, a, i) \tag{5.4}$$

$$\widehat{F} = (F, \Delta F, \Delta\Delta F) = (f_k, df_k, ddf_k) \tag{5.5}$$

$$df_k = (f_{k,t+1} - f_{k,t})/\Delta t$$

$$ddf_k = (df_{k,t+1} - df_{k,t})/\Delta t$$

Normalización de las características. Se aplicaron dos métodos de normalización sobre el conjunto final de características basados en (Ortega-Garcia *et al.*, 2003a):

- Una *normalización geométrica* sobre las características posicionales (x, y) para situar el origen del sistema de coordenadas en el centro geométrico de la firma

$$f_k^{N_1} = f_k^p - \mu_k^p \tag{5.6}$$

- Una *normalización estadística* sobre todas las características para que tuvieran media cero y varianza uno. Con ella se pretendía que su ponderación relativa dentro del vector de características fuera la misma para todas ellas.

$$\begin{aligned} f_k^{N_2} &= (f_k - \mu_k)/\sigma_k \\ \mu_k &= (\sum_{t=1}^{N} f_{k,t})/N \\ \sigma_k &= \sqrt{(\sum_{t=1}^{N} (f_t - \mu_k)^2)/(N-1)} \end{aligned} \tag{5.7}$$

5.2.3. Módulo comparador

Las dos alternativas más empleadas para determinar la semejanza entre series temporales correspondientes a firmas dinámicas son las *basadas en plantillas* y las *basadas en modelos* (Fierrez & Ortega-Garcia, 2007).

Los sistemas basados en plantillas requieren el almacenamiento de varias instancias de firma del cliente para tener en cuenta la variabilidad intraclase (fig.5.1(e)). Sin embargo, los sistemas basados en modelos no necesitan almacenar firmas del usuario, sino una representación compacta de algunos de sus parámetros que las caractericen (el modelo). Ambas alternativas han sido usadas con éxito y rendimiento similar en sistemas a nivel del estado del arte (Yeung *et al.*, 2004).

Para el sistema que presentamos en este capítulo hemos elegido el enfoque basado en plantillas utilizando el algoritmo DTW para calcular las distancias entre pares de firmas. Este enfoque combina una gran precisión (Martens & Claesen, 1996; Griess, 2000; Jain *et al.*, 2002; Kholmatov & Yanikoglu, 2005) con una implementación eficiente y sencilla, válida para un amplio espectro de escenarios prácticos.

Cálculo de la distancia. El algoritmo DTW proporciona un alineamiento óptimo no lineal entre dos secuencias de vectores de distinta longitud. Para ello minimiza la distancia total acumulada a lo largo del camino de alineamiento entre las dos secuencias temporales. Así pues, la distancia entre una firma de referencia S_R, formada por N vectores (ec. 5.8), y una firma de prueba S_T, formada por M vectores (ec. 5.9), se calcula completando la matriz DTW según la ecuación (5.11).

$$S_R = \{r_i\}_{i=1..N} \tag{5.8}$$
$$S_T = \{t_j\}_{j=1..M} \tag{5.9}$$

$$DTW_{N+1\times M+1} = DTW[i,j] = \overbrace{dist(i,j)}^{\text{coste actual}} + \overbrace{\begin{cases} dist(i-1,j) \\ dist(i,j-1) \\ dist(i-1,j-1) \end{cases}}^{\text{coste acumulado}}$$
$$dist(f_i^R, f_j^T) = \sqrt{\sum_{k=1}^{K}(f_{i,k}^R - f_{j,k}^T)^2} \qquad \text{K: dim. vec. caracts.} \tag{5.10}$$
$$DTW[0,0] = 0$$
$$DTW[i,0] = DTW[0,j] = \infty$$

La distancia final entre la firma de referencia y la de prueba será almacenada en la celda superior izquierda de la matriz DTW (ec. 5.11).

$$Dist(S_R, S_T) = DTW[N,M] \tag{5.11}$$

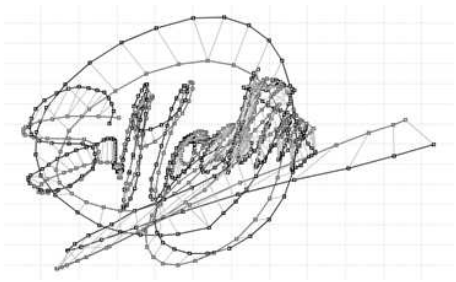
Para tener en cuenta la variabilidad intraclase almacenamos cinco firmas del usuario, cantidad suficiente para proporcionar buenos resultados (Fierrez & Ortega-Garcia, 2007) y no supone demasiadas molestias en el registro del usuario en un entorno práctico.

Para obtener una única medida de distancia a partir de las cinco distancias entre la firma de entrada (desconocida) y las cinco firmas de referencia se usó la media aritmética.

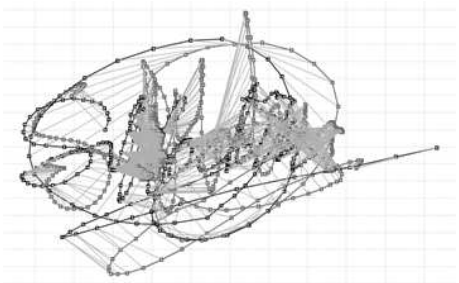
Medición del error. Para calcular la tasa de equierror (EER) se realizaron 10 pruebas distintas seleccionando aleatoriamente cinco firmas de usuario en cada una.

Se realizaron pruebas separadas en dos escenarios con restricciones de seguridad distintas. Una prueba en escenario casual (de baja seguridad) y otra en escenario seguro (de ataque). Para calcular la tasa de falsa aceptación (FAR) en el primer escenario se tomó una firma auténtica del resto de los usuarios de la base de datos. En el segundo caso, se emplearon todas las imitaciones disponibles de cada usuario. Para determinar la tasa de falso rechazo (FRR) se utilizaron en ambos escenarios todas las firmas auténticas disponibles y que no se habían empleado en el entrenamiento.

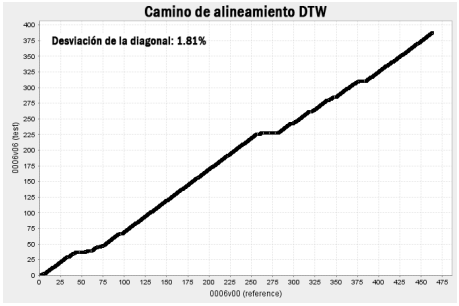
El error individual por usuario fue medido usando umbral individual ajustando dicho umbral para que los valores de FAR y FRR coincidieran. El error final del sistema se obtuvo como la media de los errores individual para cada uno de los usuarios. Al usar umbral individual no fue necesario normalizar las distancias por lo que se tomaron los valores absolutos.



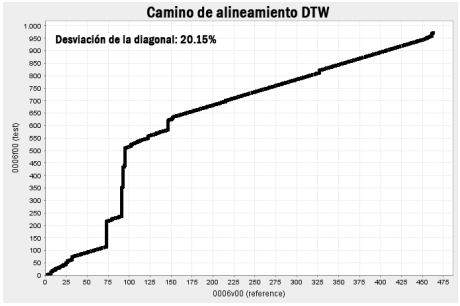
(a) Auténtica-auténtica



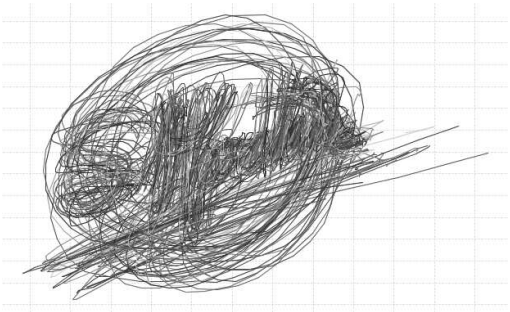
(b) Auténtica-imitación



(c) Camino de alineamiento auténtica-auténtica



(d) Camino de alineamiento auténtica-imitación



(e) Variabilidad intraclase

FIGURA 5.1: Las figuras a) y b) ilustran el alineamiento entre pares de firmas auténticas y auténtica-imitación. Los respectivos caminos de alineamiento pueden verse en c) y d). El camino de alineamiento entre genuinas es más cercano a la línea diagonal que entre genuina e imitación. La figura e) muestra un ejemplo de variabilidad intraclase

5.2.4. Bases de datos

En este trabajo hemos utilizado firmas de cuatro bases de datos de firma dinámica adquiridas por diferentes grupos de investigación (tabla 5.1). Todas fueron creadas usando modelos similares de tableta gráfica, de modo que todas las firmas comparten un conjunto común de características. Además, todas poseen imitaciones de las firmas de los usuarios, lo cual era condición indispensable para poder probar nuestro sistema en escenarios susceptibles de ser atacados [1]. Otro criterio por el que se seleccionaron estas bases de datos es la existencia de trabajos con los que poder comparar resultados.

El conjunto formado por todas las firmas fue dividido en dos conjuntos:

- Un *conjunto de desarrollo* (C_D) el cual estaba formado por firmas de los 50 primeros usuarios del corpus MCYT-100 (Ortega-Garcia *et al.*, 2003b) (de aquí en adelante MCYT-A). Este conjunto fue usado para obtener las combinaciones de características óptimas.
- Un *conjunto de prueba* (C_P) formado por el resto de las firmas disponibles, esto es, las firmas de los 50 usuarios restantes de MCYT-100 y las firmas de las otras bases de datos utilizadas: SVC 2004[2] (Yeung *et al.*, 2004), BIOMET (Garcia-Salicetti *et al.*, 2003) y MyIDEA (Dumas *et al.*, 2005). Ninguna de las firmas de este subconjunto fueron utilizadas para entrenar u optimizar el sistema.

Se pueden consultar más detalles de las bases de datos utilizadas en el capítulo 3.

TABLA 5.1: Algunos datos sobre la composición de las bases de datos empleadas

Conjunto de datos	Base de datos	Núm. usuarios	Firmas auténticas	Imitaciones	Total
C_D	MCYT-A	50	25	25	**2500**
C_P	MCYT-B	50	25	25	**2500**
	SVC 2004	40	20	20	**1600**
	BIOMET	84	15	17	**2688**
	MYIDEA	69	18	36	**3726**
Total		**293**	**5802**	**7212**	**13014**

5.3 Selección de combinaciones óptimas

Las tabletas gráficas utilizadas capturaban tanto la parte visible (con tinta) como los levantamientos (penups) del lápiz durante la firma. Para dotar de mayor universalidad a nuestro sistema optamos por eliminar los segmentos no visibles.

El primer reto para mejorar el rendimiento del sistema era seleccionar un conjunto de características óptimo. Para ello, realizamos evaluaciones de características tanto a nivel individual como combinadas. Aunque ya existía un estudio de la influencia en el rendimiento

[1] Los usuarios para los que no existían imitaciones fueron eliminados.

[2] La parte pública de esta base de datos.

al combinar parámetros globales de firma (Fierrez-Aguilar *et al.*, 2005b), pensamos que éste es el primer estudio sistemático sobre combinación de características locales. El análisis de características fue realizado en las dos siguientes etapas.

5.3.1. Evaluación individual

En una primera etapa las 15 características iniciales (ec. 5.5) fueron evaluadas individualmente sobre el conjunto de desarrollo C_D para ver si era posible reducir la dimensión del espacio de búsqueda. Esta primera etapa es importante ya que aunque sólo se partía de 15 características, las posibles combinaciones que pueden formarse con ellas ascendían a 32767 ($2^{15} - 1$). La evaluación individual de las 15 características iniciales en los dos tipos de escenario proporcionó interesantes resultados (fig. 5.2). Como ya lo apuntaban otros trabajos (Fierrez-Aguilar *et al.*, 2005c; Lei & Govindaraju, 2005) las características angulares del lápiz (acimut a e inclinación i) tienen un rendimiento más pobre que el resto de características, en nuestro caso en los tres dominios de la señal y en los dos tipos de escenario. Estos primeros resultados nos permitieron descartar estas características reduciendo la dimensión del espacio de búsqueda de 15 a 9 variables. Esta reducción de tan sólo 6 variables conlleva sin embargo una reducción del 98 % del espacio de búsqueda de combinaciones de características (desde $2^{15}-1 = 32767$ a $2^9-1 = 511$). Con ello, podemos abordar un proceso de selección de forma computacionalmente más simple, pero sin pérdida relevante en rendimiento. El resto de características que nos quedaron disponibles (las coordenadas x, y y la presión p, en los tres dominios de la señal) son un conjunto mínimo de características que está usualmente disponible en la mayoría de dispositivos de captura de firma dinámica.

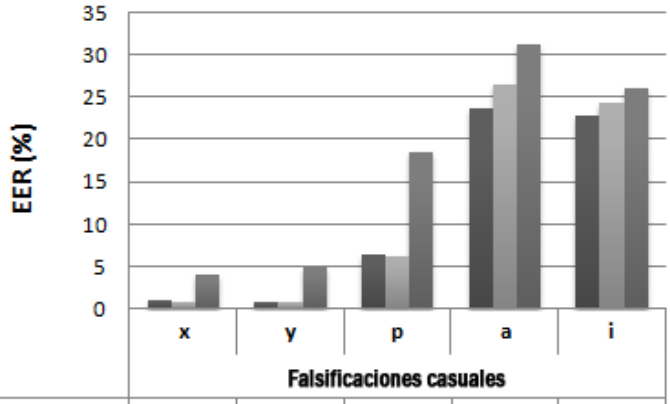
Al igual que en (Plamondon & Parizeau, 1988) los mejores resultados a nivel individual son los obtenidos en el dominio de la velocidad.

5.3.2. Evaluación conjunta

Después de determinar el conjunto de características de partida, el siguiente paso fue encontrar una forma de combinarlas de forma óptima, ya que se sabe que no siempre la combinación de las mejores características a nivel individual es la mejor solución conjunta (Cover, 1974). Utilizamos tres técnicas clásicas de selección de características (Jain *et al.*, 2000):

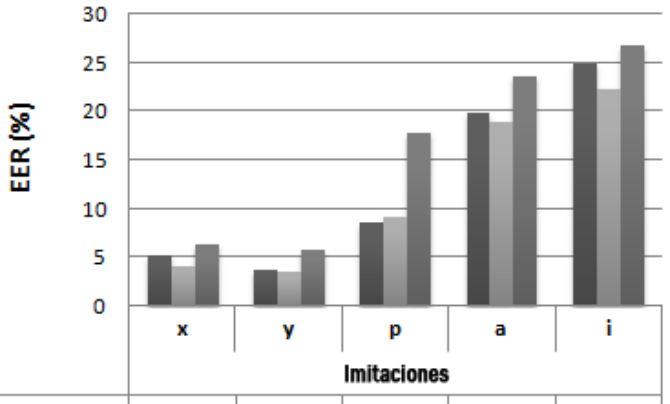
- *Selección Secuencial hacia Delante -en inglés Sequential Forward Selection o SFS-*. Con este método se incorporan progresivamente las características más prometedoras en los conjuntos mayores de forma que una vez que una característica ha sido añadida no puede ser descartada. En términos de coste computacional esta es la solución más eficiente puesto que el tamaño de los conjuntos evaluados se mantiene pequeño hasta los pasos finales del procedimiento.

- *Selección Secuencial hacia Atrás -en inglés Sequential Backward Selection o SBS-*. En cierto modo opuesto a SFS, con SBS se elimina una característica cada vez, de forma que ésta no puede ser recuperada para el conjunto final. Este método es



	x	y	p	a	i
Dominio de la posición	1,04	0,73	6,37	23,72	22,84
Dominio de la velocidad	0,80	0,87	6,26	26,56	24,40
Dominio de la aceleración	4,11	4,90	18,51	31,37	25,99

(a) Escenario casual



	x	y	p	a	i
Dominio de la posición	5,27	3,67	8,67	19,92	24,95
Dominio de la velocidad	4,07	3,46	9,22	18,93	22,20
Dominio de la aceleración	6,36	5,70	17,69	23,66	26,84

(b) Escenario seguro

FIGURA 5.2: Resultados de la evaluación individual de características de firma con el conjunto de datos de desarrollo sobre los dos tipos de escenario evaluados

computacionalmente más costoso que el anterior pero sus defensores argumentan que tiene mejor en cuenta las dependencias entre las características (Guyon & Elisseeff, 2003).

- *Plus-l take away-r o PTA(l,r).* Es una técnica intermedia entre las dos anteriores. Fue evaluada con $l = r = 1$. Este método pretende conseguir un balance entre el coste computacional y un adecuado tratamiento de las dependencias entre las características.

Los resultados de las pruebas con los tres métodos de búsqueda están representados en las gráficas de la figura 5.3(a). Las tres curvas de cada gráfica muestran el error obtenido en cada paso del proceso de selección para cada uno de los tres métodos analizados cuyos valores absolutos se encuentran en la tabla intermedia. La tabla inferior de cada gráfica muestra el orden de incorporación/descarte de características durante el proceso de selección. Obsérvese que el error obtenido con las combinaciones óptimas es significativamente inferior al obtenido con combinaciones no óptimas (desde 0.73 % a 0.20 % para falsificaciones casuales y desde 3.46 % a 1.23 % para imitaciones).

Los tres métodos de selección condujeron a la misma combinación de características óptima (x, y, dx, dy) en escenario casual lo que indica la robustez de esta combinación cuando se busca una baja tasa de falso rechazo más que unas medidas de seguridad muy estrictas. En este escenario la característica de la presión podría ser ignorada lo cual es idóneo para implementar el sistema en más dispositivos.

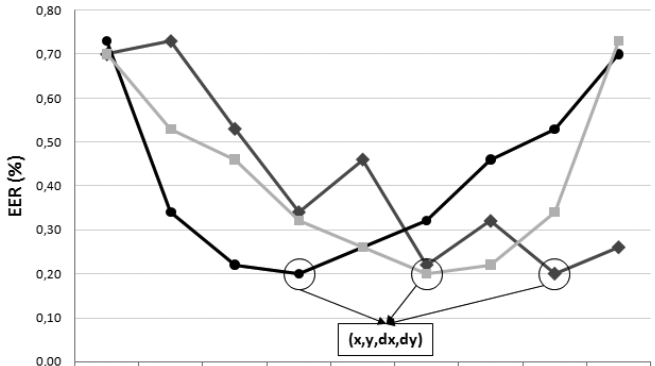
Para el escenario seguro los mejores resultados los obtuvimos con el método SBS. La combinación óptima obtenida (y, dx, dy, p) sí incluye la presión. SFS y PTA encuentran ambos la misma solución, que aunque tiene peor rendimiento que la anterior se consigue con un menor coste computacional. Con este resultado podemos sugerir que para entornos seguros la adición de la presión es muy recomendable, incluso a coste de un incremento controlado de la tasa de falsos rechazos.

5.3.3. Combinación óptima dependiente del usuario

Los resultados previos fueron obtenidos usando el mismo conjunto de características a evaluar sobre todos los usuarios, es decir usando un *conjunto de características universal* o independiente del usuario.

Como complemento a este estudio, pensamos que era interesante evaluar el rendimiento del sistema usando un *conjunto de características dependiente del usuario*. Para ello, seleccionamos el mejor conjunto de características *a posteriori* de forma individualizada para cada uno de los usuarios del conjunto de desarrollo C_D. Aunque en este trabajo no vamos a proponer un método *a priori* de selección de características dependientes del usuario, realizamos dicho estudio para establecer la cota de error obtenida al utilizar esta nueva estrategia de selección de características.

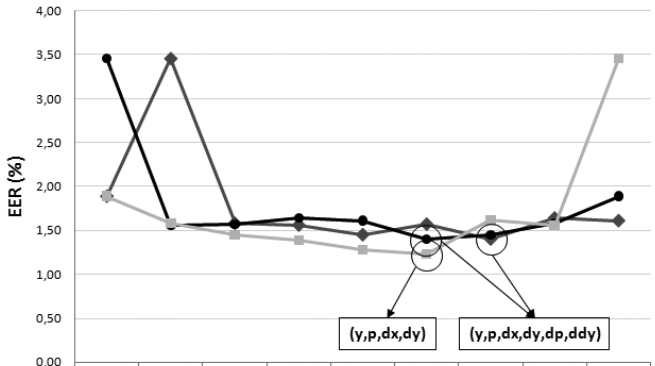
Tras todas las pruebas realizadas con los distintos métodos de selección de características sobre el conjunto de desarrollo a las que les añadimos las combinaciones no óptimas que se verán en la siguiente sección (5.4) disponíamos de 106 combinaciones de características para cada uno de los 50 usuarios de C_D. Si de esas 106 combinaciones asignamos a cada usuario aquella con la que obtuvo menos error, esto es su *combinación óptima personal*, los



	1	2	3	4	5	6	7	8	9
PTA	0,70	0,73	0,53	0,34	0,46	0,22	0,32	0,20	0,26
SFS	0,73	0,34	0,22	0,20	0,26	0,32	0,46	0,53	0,70
SBS	0,70	0,53	0,46	0,32	0,26	0,20	0,22	0,34	0,73

	Paso									Óptimas	
	1	2	3	4	5	6	7	8	9	EER	Combinación
SFS	+y	+dx	+x	+dy	+p	+dp	+ddx	+ddy	+ddp	0,20	(x,y,dx,dy)
SBS	-ddp	-ddy	-ddx	-dp	-p	-dy	-x	-dx	-y	0,20	(x,y,dx,dy)
PTA	+y	-ddp	+dx	-ddy	+x	-ddx	+dy	-dp	+p	0,20	(x,y,dx,dy)
	Característica añadida (+) / Descartada (-)										

(a) Escenario casual



	1	2	3	4	5	6	7	8	9
PTA	1,89	3,46	1,58	1,56	1,45	1,57	1,40	1,64	1,61
SFS	3,46	1,56	1,57	1,64	1,61	1,40	1,45	1,58	1,89
SBS	1,89	1,58	1,45	1,39	1,28	1,23	1,62	1,56	3,46

	Paso									Óptimas	
	1	2	3	4	5	6	7	8	9	EER	Combinación
SFS	+dy	+dx	+ddy	+dp	+y	+p	+x	+ddx	+ddp	1,40	(y,p,dx,dy,dp,ddy)
SBS	-ddp	-ddx	-ddy	-dp	-x	-y	-p	-dx	-dy	1,23	(y,p,dx,dy)
PTA	+dy	-ddp	+dx	-ddx	+ddy	-x	+dp	-p	+y	1,40	(y,p,dx,dy,dp,ddy)
	Característica añadida (+) / Descartada (-)										

(b) Escenario seguro

FIGURA 5.3: Resultados con combinaciones de características en escenarios casual y seguro con el conjunto de datos de desarrollo

resultados obtenidos para cada uno de los dos escenarios de prueba son los que se muestran en la tabla 5.2.

TABLA 5.2: Errores obtenidos con el sistema DTW con combinaciones óptimas universal y personal

Combinación	**% EER Escenario casual**	**% EER Escenario seguro**
Independiente de usuario	0.20	1.23
Dependiente de usuario	**0.06**	**0.49**
Reducción del error	**70 %**	**60 %**

Estos resultados son bastante prometedores, ya que suponen una gran reducción porcentual del error, no sólo respecto al obtenido con combinación óptima universal, sino que también sitúan el error en un orden de magnitud por debajo de los errores actuales a nivel del estado del arte. Por ello, vemos interesante analizarlos con un poco más de detalle, para facilitar la selección *a priori* de combinaciones óptimas dependientes de usuario de cara a futuros trabajos.

Reducción del espacio de búsqueda de combinaciones. Como hemos indicado, disponíamos los errores de verificación de un conjunto de búsqueda formado por 106 combinaciones de características por usuario, de las cuales seleccionando las mejores de forma particular a cada uno de ellos obtuvimos los resultados de la tabla 5.2. Con el fin de reducir este conjunto de búsqueda, cabía preguntarse cuántas de ellas eran realmente necesarias para obtener los resultados dependientes de usuario. Observamos que muchas de las combinaciones no servían a ninguno de los usuarios para obtener su mejor rendimiento. Así pues, denominamos *combinaciones óptimas* a aquellas combinaciones con las que uno o varios usuarios obtuvieron su menor tasa de error. De las 106 combinaciones 99 resultaron ser óptimas, por lo que seguía siendo necesario reducir este conjunto de búsqueda.

Comprobamos que muchos usuarios obtenían su menor tasa de error con muchas combinaciones simultáneamente, debido a una gran consistencia de su firma y/o ser difíciles de imitar. Por ello trazamos una matriz M_E con los errores para cada par (usuario,combinación). A partir de dicha matriz construimos una tabla de equivalencias entre combinaciones T_E marcando aquellas celdas en las que el error era mínimo para el usuario[3]. Esta tabla de equivalencias nos permite obtener lo que denominamos *combinaciones primas*. Las combinaciones primas son aquellas que no están 'dominadas' por otras. Una combinación C_d está dominada por otra combinación C_D cuando esta última es óptima para todos y cada uno de los usuarios para los que C_d es óptima. En el caso en el que ambas combinaciones sean óptimas para los mismos usuarios las denominamos *combinaciones equivalentes*. Ante dos combinaciones equivalentes seleccionamos la que esté formada por un menor número de características.

Veámoslo con un ejemplo con 4 combinaciones y 5 usuarios. La tabla 5.3 contiene la matriz de errores M_E por usuario para cada combinación de características. En ella

[3]Este método de simplificación está inspirado en el método tabular de Quine-McCluskey de minimizacíon de funciones lógicas.

TABLA 5.3: Matriz de errores (M_E) y tabla de equivalencias (T_E) del ejemplo

M_E	Usuario 1	Usuario 2	Usuario 3	Usuario 4	Usuario 5
Combinación 1	3.5 %	1.0 %	**0.0 %**	**0.5 %**	**0.0 %**
Combinación 2	5.0 %	1.0 %	**0.0 %**	2.0 %	**0.0 %**
Combinación 3	**0.0 %**	**0.0 %**	**0.0 %**	10.0 %	1.5 %
Combinación 4	**0.0 %**	1.5 %	**0.0 %**	7.0 %	1.0 %

T_E	Usuario 1	Usuario 2	Usuario 3	Usuario 4	Usuario 5	
Combinación 1			×	×	×	D
Combinación 2			×		×	
Combinación 3	×	×	×			D
Combinación 4	×		×			

marcamos en negrita los valores mínimos del error para cada usuario. Con ellos creamos la tabla de equivalencias T_E marcando sólo aquellas celdas en las que el error era mínimo. Dicha tabla la utilizamos para descartar las combinaciones dominadas. En este caso las combinaciones 2 y 4 están dominadas por las combinaciones 1 y 3 respectivamente que serían las combinaciones primas del ejemplo.

Siguiendo el procedimiento anterior sobre los escenarios casual y seguro conseguimos obtener un conjunto de combinaciones primas mostradas en la tabla 5.4. De este *conjunto canónico de combinaciones* sacamos las siguientes conclusiones:

- Se ha podido reducir el conjunto combinaciones óptimas de 99 a 46. Este resultado supone que escogiendo para cada usuario su combinación óptima personal de entre estas 46 el resultado obtenido es el mismo que si se busca entre las 99. Esta reducción supone una reducción considerable del coste computacional de búsqueda aún si se usa una estrategia de búsqueda por fuerza bruta.

- Se necesitan muchas menos combinaciones para obtener el resultado óptimo en el escenario casual que en el seguro. Para el escenario casual bastan con 20 combinaciones, mientras que para el seguro son necesarias 44. Este resultado puede ser interesante a la hora de aplicar la estrategia de selección de características dependientes del usuario en determinados escenarios de aplicación. Además, 18 de las 20 combinaciones primas del escenario casual, son también combinaciones primas del escenario seguro. Esto podría aplicarse para construir un sistema a partir del conjunto de las 20 características del escenario casual e incorporar progresivamente combinaciones de características según surjan necesidades desde el punto de vista de la seguridad.

- La columna 'veces óptima' muestra el número de usuarios para los que una combinación es óptima. Este valor da una idea de la fortaleza de la combinación.

TABLA 5.4: Lista de combinaciones primas del proceso de selección de combinaciones óptimas dependientes del usuario

	Núm. caracts.	**Combinación**	**Escen. casual**	**Escen. seguro**	**Nº de veces óptima**
1	1	dp		×	1
2	1	dy		×	31
3	1	y	×	×	35
4	2	dy_ddp	×		20
5	2	dy_ddx	×	×	36
6	2	dy_ddy	×	×	26
7	2	dy_dx	×	×	43
8	2	dy_p		×	23
9	2	dy_x		×	40
10	2	y_ddx		×	37
11	2	y_p		×	28
12	2	y_x	×	×	42
13	3	dy_dx_ddp		×	39
14	3	dy_dx_ddx	×	×	41
15	3	dy_dx_ddy		×	38
16	3	dy_dx_dp	×	×	39
17	3	dy_dx_x	×	×	43
18	3	y_x_ddy		×	37
19	3	y_x_ddp	×		40
20	3	y_x_dp		×	40
21	3	y_x_dx		×	43
22	3	y_x_p	×	×	37
23	4	dx_dy_dp_ddx	×	×	39
24	4	dx_dy_dp_ddy	×	×	37
25	4	x_dx_dy_ddx		×	40
26	4	x_dx_dy_dp	×	×	43
27	4	x_p_dx_dy		×	40
28	4	x_y_dx_dp	×	×	42
29	4	x_y_dx_dy		×	42
30	4	x_y_p_dx		×	41
31	4	y_dx_dy_ddx		×	40
32	4	y_p_dx_dy		×	41
33	5	dx_dy_dp_ddx_ddy	×	×	35
34	5	x_p_dx_dy_ddx	×	×	40
35	5	x_y_dx_dy_ddx		×	42
36	5	x_y_p_dx_ddx		×	40
37	5	x_y_p_dx_dy	×	×	43
38	5	y_p_dx_dy_ddx		×	41
39	6	x_dx_dy_dp_ddx_ddy		×	38
40	6	x_p_dx_dy_dp_ddy		×	36
41	6	y_dx_dy_dp_ddx_ddy		×	39
42	6	y_dx_dy_dp_ddy_ddp		×	36
43	7	x_y_dx_dy_dp_ddx_ddy		×	38
44	7	x_y_p_dx_dy_ddx_ddy		×	38
45	8	x_y_dx_dy_dp_ddx_ddy_ddp	×	×	36
46	8	y_p_dx_dy_dp_ddx_ddy_ddp	×	×	34

Evolución del error reduciendo las combinaciones primas. Para finalizar el estudio sobre combinaciones óptimas dependientes de usuario, medimos el error al descartar progresivamente combinaciones primas para ver si era posible acotar aún más el espacio de búsqueda sin que la pérdida del error fuera significativa. Para ello, realizamos pruebas en cada escenario según el siguiente procedimiento:

Paso 1: Partimos del conjunto formado por las N combinaciones primas del escenario en cuestión.

Paso 2: Eliminamos una combinación C_i y evaluamos el error seleccionando la combinación óptima personal entre las N-1 combinaciones restantes.

Paso 3: Repetimos el paso 2 para cada una de C_i $\forall i = [1..N]$

Paso 4: Descartamos la combinación cuya ausencia menos perjudica al resultado final. Esto es, aquella combinación que al extraerla del conjunto de combinaciones óptimas del escenario produce menos incremento de error en el sistema.

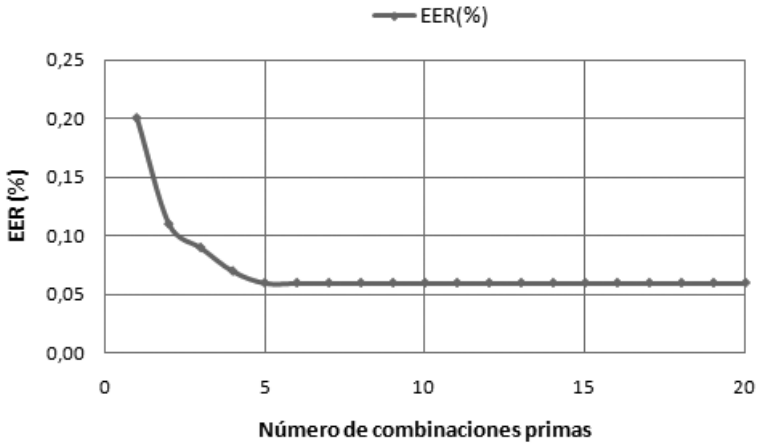
Paso 5: Hacemos N = N - 1 y repetimos el paso 1 hasta que N = 1. En este último caso nos encontramos con el sistema con combinación independiente del usuario.

Los resultados del procedimiento anterior sobre los dos escenarios se muestran en las gráficas de la figura 5.4. Estas gráficas muestran un resultado muy interesante de cara al objetivo inicial de este estudio. Ambas curvas de error muestran un comportamiento exponencial que converge rápidamente. Esto quiere decir que, usando muy pocas combinaciones, podríamos obtener un error muy cercano al error mínimo posible de obtener con todo el espacio canónico de combinaciones.

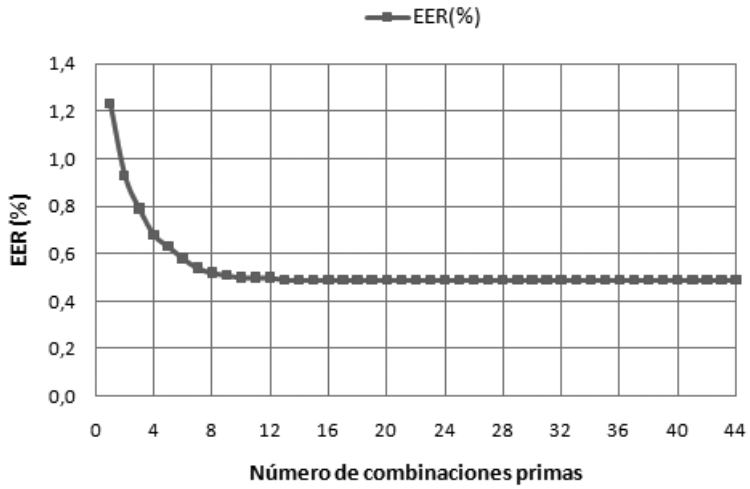
En el caso del escenario casual tenemos que con sólo cinco combinaciones somos capaces de obtener el mismo error que utilizando las 20 combinaciones de partida. El valor del error objetivo para este escenario es de 0.06 % (error en negrita en la tabla 5.2). Para el caso del escenario seguro el error objetivo era de 0.49 % y con 10 combinaciones el sistema obtiene un error de sólo 0.50 %. Si el número de combinaciones se reduce hasta sólo cinco, el error se incrementa también sólo hasta 0.63 %, aun bastante lejos del error de 1.23 % conseguido con la combinación óptima independiente del usuario (tabla 5.2).

Como se indicó inicialmente en este trabajo no hemos buscado un método de selección de la combinación óptima personal para cada usuario. Sin embargo los resultados de este estudio facilitarán futuros trabajos en dicha dirección ya que el espacio de búsqueda ha sido reducido drásticamente. En el caso de escenario casual la reducción ha sido desde 20 combinaciones a sólo 5 sin que ello suponga ninguna pérdida de rendimiento. Para el escenario seguro reduciendo el número de combinaciones desde 44 a 10 el error sólo aumenta en un 2 % (de 0.49 % a 0.50 %). Si la reducción llega hasta sólo cinco combinaciones, el incremento del error es del 29 %, de 0.49 % a 0.63 %.

En el siguiente apartado mostraremos los resultados de nuestro sistema usando combinación óptima de características independiente del usuario sobre las firmas del conjunto de evaluación C_E.



(a) Escenario casual



(b) Escenario seguro

FIGURA 5.4: Evolución del error al reducir las combinaciones primas por escenario

5.4 Resultados sobre el conjunto de prueba

Se realizaron evaluaciones con las firmas del conjunto de prueba (C_P) en los dos escenarios. Para comparar el rendimiento del sistema con combinaciones de características estándar versus las combinaciones óptimas que hemos encontrado realizamos pruebas con las siguientes combinaciones de características.

Combinaciones estándar:

- P: Combinación de características básicas de la tableta (ec. 5.4) en el dominio de la posición.
- $P + V$: Combinación anterior añadiendo las primeras derivadas temporales.
- $P + V + A$: Combinación anterior añadiendo las segundas derivadas temporales.

Combinaciones óptimas:

- F^o_{rd}: Combinación de características independiente del usuario óptima para el escenario casual.
- F^o_{sk}: Combinación de características independiente del usuario óptima para el escenario seguro.

Las tablas 5.5 muestran los errores de nuestro sistema utilizando las combinaciones de características anteriores. En ella los valores mínimos para cada tipo de combinación se muestran en cursiva. La última columna muestra el error medio sobre todas las bases de datos. Las celdas sombreadas en gris más claro muestran la mejor combinación de características. La última fila de cada tabla muestra la reducción porcentual del error de la mejor combinación estándar respecto a la mejor combinación óptima. Las celdas sombreadas en gris más oscuro muestran la reducción media del error de la mejor combinación estándar versus la combinación óptima de cada escenario.

Observamos que empleando las combinaciones óptimas se mejora drásticamente el rendimiento del sistema especialmente en el escenario casual. Otro resultado interesante es que, exceptuando el caso de la base de datos BIOMET, las combinaciones de características óptimas de cada escenario dan siempre mejores resultados que los obtenidos con la combinación de características óptima para el otro escenario. Esto demuestra la consistencia de las combinaciones óptimas que hemos encontrado con el conjunto de desarrollo y su aplicabilidad a sistemas con distintos requisitos de seguridad.

5.5 Comparativa con otros sistemas

Para finalizar el trabajo expuesto en este capítulo realizamos una comparativa (tabla 5.6) del rendimiento de nuestro sistema con el de otros sistemas publicados que usaron las mismas bases de datos para su evaluación.

Vemos que en general nuestro sistema proporciona un mejor rendimiento en casi todos los casos (excepto el sistema 8 que tiene mejores resultados en escenario casual con MCYT),

TABLA 5.5: Errores de las combinaciones de características estándar (Est.) versus combinaciones óptimas (Opt.) para cada escenario

		Escenario casual				
	Combin.	**mcyt-b**	**svc04**	**biomet**	**myidea**	**Media**
Est.	P	5.64	0.78	7.58	*2.45*	**4.11**
	$P+V$	*3.25*	*0.40*	*4.47*	3.52	**2.91**
	$P+V+A$	4.82	1.01	6.58	6.14	**4.64**
Ópt.	F^o_{rd}	*0.38*	*0.00*	*0.33*	*0.92*	**0.41**
	F^o_{sk}	0.46	0.32	0.96	2.39	**1.03**
	EER reducc.	**88.3 %**	**100.0 %**	**92.6 %**	**62.4 %**	**86.0 %**

		Escenario seguro				
	Combin.	**mcyt-b**	**svc04**	**biomet**	**myidea**	**Media**
Est.	P	6.53	4.70	5.41	*2.89*	**4.88**
	$P+V$	*4.21*	*4.15*	*3.69*	3.25	**3.83**
	$P+V+A$	4.23	6.14	4.43	4.10	**4.73**
Ópt.	F^o_{rd}	1.16	3.70	*1.25*	2.94	**2.26**
	F^o_{sk}	*1.06*	*3.38*	1.48	*2.72*	**2.16**
	EER reducc.	**74.8 %**	**18.6 %**	**59.9 %**	**5.9 %**	**40.8 %**

usando un número similar de firmas de entrenamiento. Hay que decir que dado que no existen protocolos de evaluación estándar de sistemas de firma dinámica es difícil comparar de forma objetiva distintos sistemas, incluso aunque éstos utilicen las mismas bases de datos.

TABLA 5.6: Tabla comparativa de algunos sistemas de referencia que usan las mismas bases de datos que las utilizadas en nuestros experimentos

	Autor	MCYT rd [sk]	SVC 2004 rd [sk]	Biomet rd [sk]	MyIDea rd [sk]	Comentarios
1	Hennebert *et al.* (2007)				2.7 [7.3]	- 6 firmas de entrenamiento - Algoritmo basado en GMM - Resultados del experto con variabilidad temporal.
2	Humm *et al.* (2007)				2.6 [7.3]	- 6 firmas de entrenamiento - Algoritmo basado en HMM - Resultados con esquema basado en variabilidad temporal.
3	Garcia-Salicetti *et al.* (2007)	1.22 [3.40]				- 5 firmas de entrenamiento - Algoritmo basado en HMM + distancia - Mejores resultados mediante combinación de los mejores sistemas individuales - La prueba aleatoria también incluye imitaciones
4	Ly *et al.* (2007)	[3.37]	[4.83]	[2.33]		- 5 firmas de entrenamiento - Algoritmo basado en HMM - Algoritmo de fusión entre el camino de Viterbi y verosimilitudes
5	Pascual-Gaspar & Cardenoso-Payo (2007)	2.09 [6.14]				- 3 firmas de entrenamiento - Algoritmo basado en HMM - HMM con estructura dependiente de usuario
6	SVC 2004 (Yeung *et al.*, 2004)		3.02 [6.90]			- 5 firmas de entrenamiento - Mejor sistema sobre imitaciones: DTW; Mejor sistemas sobre fals. casuales: HMM - Resultados sobre el conjunto de desarrollo (40 usuarios) para la tarea 2
7	Fierrez-Aguilar *et al.* (2005a)		0.15 [6.91]			- 5 firmas de entrenamiento - Fusión de algoritmos Local (DTW) y Regional (HMM) - Mejores resultados sobre el conjunto de desarrollo (40 usuarios) para la tarea 2
8	Fierrez-Aguilar *et al.* (2005b)	**0.24** [2.12]				- 5 firmas de entrenamiento - Fusión de expertos basados en parámetros globales (Parzen WC) y locales (HMM)
9	Fierrez *et al.* (2007)	0.05 [0.74]				- 10 firmas de referencia* - Algoritmo basado en HMM
10	Fábregas & Faundez-Zanuy (2008)	1.83 [5.17]				- 5 firmas de referencia - Biometric Dispersion Matcher
11	Nuestro sistema	0.29 [**1.23**]	**0.00** [**3.38**]	**0.33** [**1.48**]	**0.92** [**2.72**]	- 5 firmas de referencia - Algoritmo basado en DTW - Resultados con características optimizadas para cada escenario - Resultados sobre MCYT-100

[1] En negrita los mejores valores de EER para cada base de datos y escenario (rd: escenario casual; sk: escenario seguro).

5.6 Resumen

En este capítulo hemos descrito un sistema de verificación de firma dinámica basado en DTW y combinaciones óptimas de características especialmente diseñado para ser usado en diversos tipos de escenarios desde el punto de vista de la seguridad. El sistema es muy versátil ya que no necesita de características especiales del hardware de adquisición, únicamente las coordenadas geométricas y opcionalmente la presión. De este modo puede ser implementado en una gran variedad de dispositivos abarcando desde tabletas gráficas a PDAs. El sistema requiere de un reducido número de firmas del usuario con las que hemos obtenido un alto rendimiento.

Hemos hecho un estudio detallado del rendimiento con distintas combinaciones de características y se han encontrado las combinaciones óptimas de nuestro sistema para dos escenarios opuestos desde el punto de vista de la seguridad ante ataques. Además, de estas dos combinaciones independientes del usuario, hemos evaluado el rendimiento de nuestro sistema al utilizar combinaciones de características dependientes del usuario y hemos observado que el rendimiento con ellas aumenta significativamente.

El sistema ha sido finalmente evaluado sobre cuatro bases de datos de firma dinámica muy populares (MCYT , SVC 2004, BIOMET y MyIDea) demostrando que nuestro sistema proporciona excelentes resultados especialmente en el caso de escenario seguro.

Se ha dejado abierta una línea de investigación orientada a encontrar un método de selección *a priori* de la combinación óptima personal de cada usuario. Para facilitar esta tarea hemos realizado una reducción del espacio de búsqueda drásticamente sin perdidas significativas de rendimiento.

En el siguiente capítulo se muestran dos escenarios de carácter práctico en los que se ha implementado el sistema descrito en este capítulo: una competición de imitación de firmas celebrada en septiembre de 2008, y en segundo lugar, la participación en la última competición de reconocimiento de firma cuyos resultados fueron publicados en junio de 2009.

6 Escenarios prácticos

En este capítulo expondremos dos escenarios de carácter práctico llevados a cabo a lo largo del trabajo de tesis en las que hemos aplicado el sistema DTW del capítulo 5. En primer lugar, veremos el rendimiento de un sistema de verificación de firma dinámica basado en web ante ataques no controlados de imitadores organizado a modo de concurso. Tras ello describiremos la competición BSEC 2009, sus reglas y resultados oficiales así como el sistema con el que la Universidad de Valladolid participó.

6.1 Concurso de imitaciones

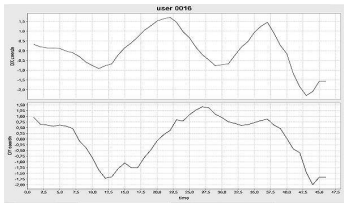
Con motivo de las IV Jornadas de Reconocimiento Biométrico de Personas [JRBP (2008)] celebradas en Valladolid los días 11 y 12 de septiembre de 2008 el grupo ECA-SIMM de la Universidad de Valladolid aprovechó la circunstancia para organizar una competición de imitaciones de firmas entre los asistentes a las Jornadas.

6.1.1. Objetivo

El objetivo era evaluar 'in vivo' la robustez ante ataques de falsificadores de la firma dinámica como rasgo biométrico. Para ello, se propuso a los asistentes imitar cinco firmas de diferente complejidad extraídas de la base de datos MCYT. Las cinco firmas (fig. 6.1) fueron clasificadas por complejidad desde el grado 'muy fácil' al grado 'muy difícil'. La complejidad de las firmas fue establecida por los organizadores de forma subjetiva en base a la apariencia visual y al perfil de velocidades en los ejes de coordenadas (Vx,Vy).

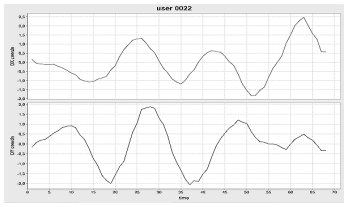
La tabla 6.1 muestra además los valores estadísticos de los parámetros globales más relevantes de las firmas utilizadas como plantillas por el sistema de verificación del concurso. Contiene la media y la desviación estándar del número de puntos de las firmas (proporcional a la duración temporal), la longitud total de las firmas y su velocidad media. Los valores de estas dos últimas vienen expresadas en las magnitudes de la tableta, que no mostramos por no ser necesarias a efectos comparativos entre ellas. De dichos valores y los perfiles de velocidad intentaremos dar una explicación a los resultados obtenidos en el concurso.

(a) Dif.: Muy Fácil



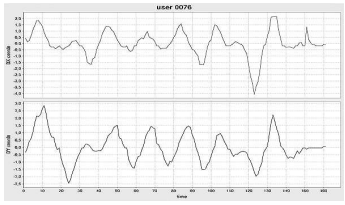
(b) Perfil de velocidades

(c) Dif.: Fácil

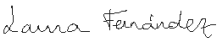


(d) Perfil de velocidades

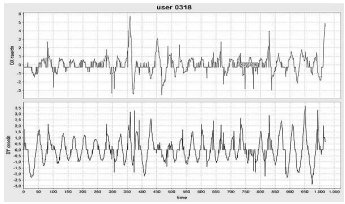
(e) Dif.: Media



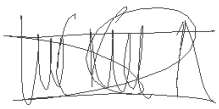
(f) Perfil de velocidades



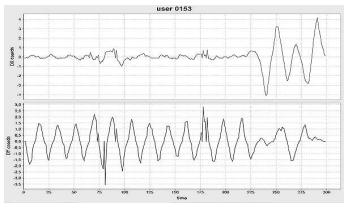
(g) Dif.: Difícil



(h) Perfil de velocidades



(i) Dif.: Muy Difícil



(j) Perfil de velocidades

FIGURA 6.1: Firmas utilizadas en el concurso de imitaciones durante las IV JRBP: grados de dificultad y perfiles de velocidad en los ejes X,Y

TABLA 6.1: Parámetros globales básicos de las firmas del concurso de imitaciones

	Media			Desv. estándar (%)		
Dificultad	**Puntos**	**Longitud**	**Velocidad**	**Puntos**	**Longitud**	**Velocidad**
MF	45.00	10728.51	238.20	2.81 %	7.82 %	5.84 %
F	70.00	13474.83	192.71	3.50 %	6.33 %	7.39 %
M	160.50	11387.95	71.07	4.49 %	6.90 %	8.34 %
D	989.83	15576.16	15.77	4.17 %	7.51 %	8.89 %
MD	299.33	35843.03	119.73	0.81 %	3.32 %	2.67 %

6.1.2. FirmWeb

La plataforma bajo la que se llevó a cabo el concurso fue la aplicación web FirmWeb[1]. FirmWeb (figuras 6.2 y 6.3) se trata de una aplicación web desarrollada por el profesor de la Universidad de Valladolid, Carlos Enrique Vivaracho Pascual, y el autor de esta tesis, Juan Manuel Pascual Gaspar, con fines de investigación en el campo del reconocimiento de firma. La aplicación consta de un módulo de registro en el que a través de un applet Java se facilita el registro y posterior acceso a la zona restringida de la aplicación. Una singularidad de FirmWeb es que permite el registro tanto con tableta gráfica (modo tableta) como sin ella (modo puntero). El modo puntero puede utilizarse por ejemplo si se accede a la aplicación desde un dispositivo móvil (PDA) o Tablet PC. En dicho caso se registran las coordenadas geométricas recogidas por el applet. Si se dispone de tableta gráfica WACOM (junto al driver apropiado) se recogen las coordenadas suministradas por ella. La aplicación permite interacción en ambos modos (modo tableta y/o modo puntero) adaptándose a cada sensor de forma transparente al usuario.

Para participar en el concurso cada usuario debía registrarse previamente en FirmWeb. El procedimiento de registro consta de dos fases, separadas en 24 horas. En cada una de ellas el usuario debe proporcionar tres firmas al sistema. El objetivo de este registro multisesión es capturar la variabilidad temporal de la firma del usuario. En el caso del concurso se permitió completar el registro sin que transcurrieran las 24 horas habituales entre las dos fases para dar más tiempo a los participantes a practicar.

[1] `http://www.greidi.uva.es/JRBP08_firmas/inicio/`

FIGURA 6.2: Página principal FirmWeb

FIGURA 6.3: Acceso a FirmWeb desde PDA

6.1.3. Reglas

Los participantes podían visualizar la firma a imitar tantas veces como quisieran, tanto el aspecto final como la dinámica de realización. Para ello disponían de un módulo software para reproducir la dinámica de ejecución a distintas velocidades: lenta, media y a velocidad real. Para practicar las imitaciones los participantes disponían de un puntero con tinta y plantillas en papel con la imagen de las firmas. Es obvio que se trata de un escenario de ataque irrealmente adverso, pero nuestro objetivo era precisamente el verificar la robustez en un caso límite de seguridad.

Una vez el usuario se había registrado en el sistema podía realizar tantas imitaciones como quisiera, sin límite en el número de intentos ni de aciertos. Tras cada intento el usuario recibía la puntuación obtenida en una pantalla (fig. 6.4) de forma que podía utilizarla para

ir mejorando su técnica de imitación hasta conseguir 'engañar' al sistema. La correcta imitación de una firma otorgaba un número de puntos proporcional a la dificultad de la firma imitada. El vencedor sería aquel usuario que consiguiera más puntos a lo largo de los dos días de duración de las Jornadas. Al ganador se le obsequiaría con un premio valorado en 150 euros.

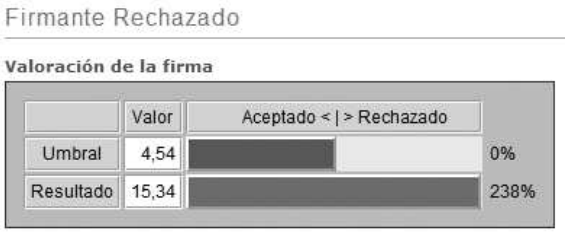


FIGURA 6.4: Pantalla de FirmWeb con puntuación obtenida tras el envío de la imitación

El umbral utilizado para determinar si la firma era aceptada o rechazada se estableció con el valor necesario para que todas las firmas genuinas del cliente no utilizadas como plantilla fueran aceptadas por el sistema (FRR = 0 %). A dicho valor se le añadió una cantidad extra a modo de margen de seguridad del 10 %.

6.1.4. Algoritmo de verificación

El sistema de verificación de firma es básicamente el descrito en el capítulo 5. En este caso, las características utilizadas fueron únicamente las componentes instantáneas de la velocidad en los ejes x-y. Cuando el usuario interactuaba en modo puntero con la aplicación el sistema realiza una normalización temporal de los datos de la firma a la frecuencia de la tableta (100Hz). Como firmas de referencia se utilizaron las seis primeras disponibles.

6.1.5. Resultados

Los resultados del concurso se muestran en la tabla 6.2. Cada celda contiene el ratio (E/I) entre el número de éxitos E de cada usuario por firma imitada y el número de intentos I realizados. Algunos usuarios no realizaron imitaciones de algunas firmas por lo que las celdas correspondientes aparecen sin rellenar.

Los resultados que nos interesan son las tres filas finales de la tabla que totalizan los resultados de todos los usuarios. Como puede verse, la firma que más veces se consiguió imitar fue la de menor dificultad (dificultad MF). El porcentaje de error cometido por el sistema con esta firma fue del 15.33 % (FAR=15.33 % cuando FRR=0 %) cifra que puede considerarse alta para un sistema real si la seguridad es un aspecto crítico. Sin embargo, en un sistema de baja seguridad en el que sólo se busque sustituir la mera observación visual del empleado por un sistema automático (por ejemplo en puntos de venta), consideramos que el error obtenido no es demasiado elevado, teniendo en cuenta las facilidades dadas para realizar las imitaciones del concurso.

Comparando los perfiles de velocidad de las firmas de dificultad Fácil (F) y Media (M), podría inducirse que la primera es más fácil de imitar que la segunda. Sin embargo la firma

TABLA 6.2: Resultados del concurso de imitaciones de las IV JRBP

Participante	**MF**	**F**	**M**	**D**	**MD**
1	1/3	0/1	-	0/1	-
2	1/2	-	0/5	0/25	0/2
3	2/13	2/24	0/34	0/7	0/8
4	2/3	0/21	0/2	0/5	0/10
5	0/13	-	-	-	-
6	2/18	0/1	0/1	-	-
7	20/60	11/92	0/24	0/33	0/30
8	21/25	0/13	-	-	-
9	0/10	-	-	-	-
10	4/17	-	-	-	-
11	9/54	0/35	0/15	0/8	0/3
12	0/29	-	-	-	-
13	3/6	2/30	22/143	0/15	0/61
14	1/2	0/15	0/16	0/1	0/2
15	0/6	-	0/20	0/13	0/1
16	20/105	1/123	0/32	0/1	0/6
17	1/195	0/16	0/3	0/2	-
18	1/2	0/2	-	0/1	-
19	0/2	-	0/9	-	-
20	0/3	-	-	-	-
21	0/6	-	-	-	-
Intentos	574	373	304	111	123
Fallos	88	16	22	0	0
% Error	15.33 %	4.29 %	7.24 %	0.00 %	0.00 %

M ha resultado porcentualmente más veces imitada que la firma F. No obstante, todas las imitaciones de la firma M fueron realizadas por el mismo participante (13), por lo que claramente se trata de un resultado propio del concurso, ya que imitar esta firma puntuaba más que la anterior, por lo que el mismo usuario se dedicó más a la de mayor complejidad. La firma F fue imitada por cuatro participantes, por lo que pensamos que la clasificación inicial era correcta.

Ningún participante del concurso consiguió imitar las firmas de complejidades Difícil (D) y Muy Difícil (MD). En este caso, para determinar cuál de las dos firmas es más compleja de imitar, medimos la distancia de la mejor imitación al umbral en términos relativos. Para el caso de la firma D el margen desde el umbral a la mejor imitación era del 6 %, mientras que en el caso de la firma D este margen es del 18 %. Concluimos que la clasificación realizada de forma no sistemática por los organizadores resultó correcta aunque a vista de los resultados la creación de una medida objetiva de la fortaleza ante ataques resulta un prometedor campo para futuras investigaciones.

En este sentido, datos básicos de carácter global como los mostrados en la tabla 6.1, pueden ayudar a su determinación. Obsérvese que para que una firma sea robusta ante ataques debe tener las siguientes propiedades:

- Debe contener una cantidad suficiente de información. Esta medida será proporcional a la longitud de la firma.
- Debe ser estable a lo largo del tiempo. Puede medirse en función de las desviaciones típicas de los parámetros globales.
- La información contenida debe ser difícilmente reproducible por personas distintas al autor. Ello se conseguiría con una firma veloz y a ser posible no legible, ya que los rasgos legibles son más fácilmente reproducibles que los rasgos tipo 'garabato'.

6.1.6. Comparativa con imitaciones de MCYT

Para comprobar la calidad de las imitaciones generadas en el concurso realizamos una comparativa del error obtenido con ellas en forma de EER respecto a los errores obtenidos con las imitaciones disponibles de dichas firmas en la base de datos MCYT.

En la tabla 6.3 vemos que, a excepción de la firma MF, en todas las demás el error con las imitaciones del concurso es igual o mayor al obtenido con las imitaciones de MCYT. El resultado es el esperado ya que los imitadores de las firmas MCYT no disponían de las facilidades de los participantes del concurso de imitadores de JRBP08. El mayor error obtenido con la firma MF puede deberse al protocolo de asignación de roles de imitador a los participantes de MCYT ya que las 25 imitaciones fueron realizadas por cinco participantes, por lo que si uno de ellos fue capaz de imitar correctamente la firma MF el porcentaje de error debido a dicho usuario podría incrementar 'artificialmente' el error total.

Tabla 6.3: EER obtenido con las imitaciones del concurso JRBP08 versus EER obtenido con las imitaciones de MCYT

	MF	F	M	D	MD
JRBP08	6.64 %	1.47 %	0.49 %	0.00 %	0.00 %
MCYT	13.89 %	0.00 %	0.00 %	0.00 %	0.00 %

6.1.7. Conclusiones y futuros trabajos

Hemos descrito el protocolo y los resultados del concurso de imitaciones de firmas llevado a cabo durante las JRBP08. No existen demasiados trabajos de ataques a sistemas de verificación de firma por lo que los resultados obtenidos pensamos que son interesantes. Se ha comprobado que si el imitador dispone de acceso a la dinámica de la firma la calidad de las imitaciones mejora, aunque quizás no en la proporción que podría pensarse de antemano. Además no todas las firmas poseen la misma dificultad de ser imitadas. A mayor longitud, velocidad y estabilidad, las firmas son más difíciles de imitar. Por el contrario la legibilidad de la firma influye negativamente en la robustez ante ataques. Ballard *et al.* (2006) llegaron a conclusiones similares en este sentido creando un sistema automático generador de imitaciones de frases manuscritas.

Los datos con los que se ha contado para realizar este trabajo son insuficientes para dotar de mayor validez estadística a los resultados ya que provienen de un concurso no

una investigación. Por ello, para el futuro, es necesario ampliar la cantidad de datos del experimento, así como evaluar el ataque a los sistemas de verificación de firmas desde sistemas automáticos.

6.2 BSEC 2009

La 'BioSecure Signature Evaluation Campaign' (BSEC 2009) es la siguiente competición internacional tras SVC 2004 que está centrada exclusivamente en verificación de firma dinámica. A diferencia de su predecesora en la que la base de datos fue adquirida con un único sensor (tableta gráfica), en BSEC 2009 los datos provienen de dos sensores: una tableta gráfica (WACOM INTUOS3 A6) y una PDA (HP iPAQ hx2790). Ambos recogen los datos introducidos con un puntero/lápiz a una frecuencia de 100Hz.

6.2.1. Objetivos

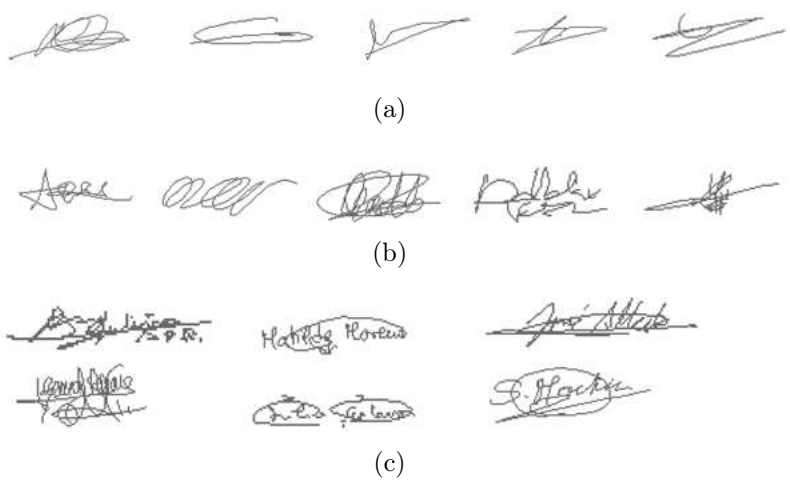
La competición se plantea con tres objetivos para cada uno de los cuales se realiza una evaluación de sistemas:

- **Evaluación 1**: Medir el impacto en el rendimiento bajo condiciones de movilidad. Para ello, se han utilizado las dos mayores bases de datos adquiridas con distintos dispositivos que contienen las mismas personas: la base de datos BioSecure (Ortega-Garcia *et al.*, 2009) que está compuesta por los conjuntos DS2 (adquirida con tableta gráfica) y DS3 (adquirida con PDA).
- **Evaluación 2**: Medir la influencia de la variabilidad temporal (paso del tiempo desde el registro) en los sistemas.
- **Evaluación 3**: Medir el rendimiento en función de la 'cantidad de información' que contiene de la firma. Para ello se ha utilizado la medida de 'entropía del cliente', descrita en los trabajos de Salicetti *et al.* (2008) y Garcia-Salicetti *et al.* (2009). En la figura 6.5 puede verse un ejemplo de firmas con distintas medidas de entropía de cliente.

6.2.2. Bases de datos

La base de datos para el concurso se dividió en dos partes: un conjunto de desarrollo formado por 50 usuarios y un conjunto de evaluación compuesto por 382 usuarios. En ambos casos se dispone de los conjuntos DS2 (tableta) y DS3 (PDA). En el caso de las firmas adquiridas con tableta se dispone de las usuales cinco características (coordenadas x-y, presión y ángulos acimut e inclinación del lápiz respecto a la tableta) tomados cada 10 milisegundos (100Hz). De las firmas adquiridas con PDA sólo se registran las coordenadas geométricas (x,y) a intervalos de tiempo no constantes. El organizador realizó una interpolación temporal a 100Hz suministrando los datos ya interpolados.

La base de datos contiene dos sesiones separadas en dos semanas. En cada sesión el usuario realizó 15 firmas genuinas y 10 imitaciones. Las imitaciones se realizaban en grupos



(a)

(b)

(c)

FIGURA 6.5: Ejemplos de firmas de MCYT-100 con (a) entropía alta, (b) entropía media y (c) baja entropía (Salicetti *et al.*, 2008)

TABLA 6.4: Composición de la base de datos BSEC 2009 para los conjuntos de desarrollo y de evaluación por tipo de firma (genuina, imitación) y totales por usuario y totales globales

Conj.	Gen./sesión	Imit./sesión	Sesiones	Tot./usuario	Usuarios	Total firmas
Desarrollo	15	10	2	50	50	2500
Evaluación	15	10	2	382	50	19100

de cinco repeticiones sobre dos usuarios. Así pues, el total de firmas de cada conjunto era de 2500 en el conjunto de desarrollo y 19100 en el conjunto de evaluación.

La base de datos de evaluación (382 usuarios) fue 'secuestrada' durante la competición y dado que no estaba disponible todavía durante la redacción de esta memoria de tesis (los organizadores comentaron que la liberarían en unos meses), los resultados aquí expuestos estarán referidos a la base de datos de desarrollo[2].

6.2.3. Protocolo

Se definieron dos tareas en función de las características de la firma utilizadas. En la tarea 1, se considerarían las cinco componentes (de la tableta). En la tarea 2, sólo se considerarían las coordenadas geométricas (tableta y PDA).

El entrenamiento se realizaba con cinco firmas genuinas de la primera sesión. Las pruebas con firmas genuinas se realizaban utilizando las 10 firmas de la primera sesión no utilizadas en el entrenamiento y las 15 firmas disponibles en la segunda sesión. Las pruebas con firmas imitadas y con firmas casuales se medirían por separado. En el caso de impostores casuales se realizaron con 30 firmas de usuarios diferentes escogidos aleatoriamente (una firma por usuario). En la prueba con firmas imitadas se utilizaron las 20

[2]Las firmas correspondientes al usuario 11 de la base de datos DS3 (PDA) fueron eliminadas del conjunto de desarrollo por iniciativa de los organizadores, ya que las sesiones 1 y 2 de éste correspondían a firmas de distintos usuarios.

TABLA 6.5: Sistemas participantes en BSEC 2009

Sistema	Universidad	Pais	Algoritmo
1	EUP Mataró	España	Biometric DM (Fábregas & Faundez-Zanuy, 2009)
2	EUP Mataró	España	Estadístico
3	U1 Res. Inst.	Hungría	DTW
4	Seikey Univ.	Japón	DTW
5	Ain Shams Univ.	Egipto	DTW
6	Univ. Valladolid	España	DTW (Pascual-Gaspar *et al*, 2009)
7	Sabanci Univ.	Turquía	DTW
8	UAM	España	DTW random
9	UAM	España	DTW skilled
10	UAM	España	HMM
11	UAM	España	Global
12	UAM	España	Fusion
13	Waseda Univ	Japón	DTW
14	Univ. Magdeburg	Alemania	Biometric Hash
Ref	TMSP	Francia	HMM (Van-Bao *et al*, BSEC2009 Reference System, 2009)

imitaciones disponibles por usuario. El error se midió mediante curvas DET y EER con umbral universal.

6.2.4. Resultados

Los resultados presentados a continuación han sido obtenidos de la página web oficial de los resultados del concurso[3].

Un total de 14 sistemas fueron presentados por 9 instituciones diferentes (tabla 6.5). Las tablas 6.6, 6.7 y 6.8 muestran los resultados de la competición para cada una de las evaluaciones. Como se observa en general las menores tasas de error para cada evaluación estuvieron repartidas entre los cinco sistemas siguientes[4]: 6 (Univ. Valladolid), 7 (Univ. Sabanci), y -8, 9 y 12- (UAM).

A diferencia de SVC 2004, esta vez el algoritmo con el que se obtuvieron los mejores resultados fue claramente DTW si bien la mayoría de sistemas estaban construidos en base a él. Reseñar el pobre resultado del algoritmo basado en HMM a pesar de que en SVC 2004 el sistema propuesto por UAM estuvo entre los dos mejores.

La tabla 6.9 muestra los resultados agrupados por dispositivo (Tableta y PDA) y por escenario desde el punto de vista de la seguridad (escenario casual -random- y seguro -skilled-). Aunque no existe un sistema que funcione mejor que el resto en todos los casos, puede decirse que el sistema que en promedio ha obtenido los mejores resultados es el sistema 12 (UAM) que está compuesto por la fusión de los sistemas individuales (8, 9, 10 y 11), también enviados por la UAM.

[3] `http://biometrics.it-sudparis.eu/BSEC2009/downloads/BSEC2009_results.pdf`

[4] Los sistemas 7, 8, 9 y 12 pertenecen a las Universidades vencedoras de la competición SVC 2004.

Tabla 6.6: Resultados evaluación 1 BSEC 2009: impacto en condiciones móviles de adquisición

			Tableta(x,y) Sesión 1		PDA Sesión 1	
Sistema	**Universidad**	**Algoritmo**	**Skilled**	**Random**	**Skilled**	**Random**
1	EUP Mataró	Biometric DM	4.40	1.85	8.18	2.05
2	EUP Mataró	Clasif. estadístico	4.91	2.33	7.38	1.86
3	U1 Res. Inst.	DTW	13.99	8.98	18.32	8.36
4	Seikey Univ.	DTW	2.88	1.58	7.87	1.29
5	Ain Shams Univ.	DTW	3.82	2.67	31.57	30.64
6	Univ. Valladolid	DTW	**2.20**	0.97	6.58	1.65
7	Sabanci Univ.	DTW	2.98	2.23	**4.99**	4.32
8	UAM	Random tuned DTW	4.18	**0.51**	12.20	**0.55**
9	UAM	Skilled tuned DTW	2.88	1.47	5.77	1.54
10	UAM	HMM	19.23	24.14	25.85	21.34
11	UAM	Global	6.71	3.31	13.26	4.70
12	UAM	Fusion	2.23	0.63	5.47	0.66
Ref	MSP	HMM	4.47	1.74	11.27	4.80

Con firmas obtenidas con tableta gráfica las tasas de error con imitaciones siguen en el mismo orden de magnitud que la anterior competición (SVC 2004), ligeramente por debajo del 3 %. Sin embargo la mejor tasa de error en escenario casual se ha reducido en aproximadamente un 1 % en valor absoluto respecto a SVC 2004.[5] Como se observa el estado del arte no ha avanzado significativamente en los cinco años que separan ambas competiciones.

Con firmas obtenidas con PDA, la tasa de error con imitaciones es más del doble que con tableta, aunque hay que decir que el imitador disponía de un software en la propia PDA para visualizar la dinámica de la firma, así como de una imagen de la firma a imitar que podía utilizar a modo de plantilla. Respecto al resultado en escenario casual, los resultados del mejor sistema son parecidos a los obtenidos con tableta gráfica, aunque en general, se produce una cierta degradación debida sin duda a las condiciones de captura del entorno móvil.

Otro resultado interesante es el hecho de que los sistemas que obtienen buenos resultados en un escenario (por ejemplo escenario casual) no suelen obtener buenos resultados en el escenario opuesto (por ejemplo escenario seguro). Esta tendencia se observa para los dos dispositivos de captura. Por ello creemos que el sistema basado en fusión es el que ha obtenido los mejores resultados en promedio para los distintos escenarios y dispositivos. Este resultado pensamos que establece una línea importante a seguir en el campo de reconocimiento de firma dinámica para los próximos años, ya que parece que con un único sistema individual los resultados obtenibles se encuentran estancados.

Por este motivo destacamos el segundo puesto obtenido por el sistema presentado por la Universidad de Valladolid considerando el rendimiento promedio entre tableta y PDA. Dicho sistema es el que mejor resultado ha obtenido en la competición por un sistema

[5] Valores comparados con los resultados de SVC 2004 con el conjunto de evaluación.

TABLA 6.7: Resultados evaluación 2 BSEC 2009: impacto en la variabilidad temporal

	Tableta (x,y)				**PDA**			
	Sesión 1		**Sesión 2**		**Sesión 1**		**Sesión 2**	
Sistema	**Skilled**	**Random**	**Skilled**	**Random**	**Skilled**	**Random**	**Skilled**	**Random**
1			7.15	4.40	8.71	2.22	14.24	3.94
2			6.20	4.02	7.38	1.85	11.25	3.76
3			19.03	12.29	18.32	8.36	24.68	12.40
4			5.99	3.55	6.37	2.00	9.43	3.72
5			6.61	5.23	7.55	4.24	11.51	6.78
6			4.21	2.24	5.69	1.50	8.06	2.90
7			5.13	3.96	4.98	4.31	7.69	7.02
8			7.26	1.80	10.4	**0.70**	14.51	**1.67**
9			**4.08**	2.93	5.24	2.09	**7.42**	2.83
10			20.47	25.62	24.79	27.29	23.52	26.81
11			10.92	5.37	10.49	2.93	15.00	5.01
12			4.18	**1.70**	**4.93**	1.41	**7.42**	1.93
13					5.98	1.44	9.93	3.48
Ref			5.99	3.16	11.27	4.80	14.03	6.06

	Tableta (x,y,p)				**Tableta (x,y,p,a,i)**			
	Sesión 1		**Sesión 2**		**Sesión 1**		**Sesión 2**	
Sistema	**Skilled**	**Random**	**Skilled**	**Random**	**Skilled**	**Random**	**Skilled**	**Random**
1	4.03	1.70	6.88	4.16				
2	4.50	1.96	6.28	3.61	4.52	1.91	5.99	3.53
3	13.69	8.62	18.54	11.81	13.41	8.63	17.91	11.77
4	2.76	1.33	6.07	3.42	3.02	1.49	6.02	3.52
5								
6	2.19	0.97	4.21	2.23	**2.19**	**0.97**	**4.21**	**2.23**
7								
8	3.26	**0.42**	6.21	**1.37**				
9	2.38	1.17	**3.48**	2.46				
10	27.76	20.51	30.13	21.61				
11	5.90	2.02	9.52	3.65				
12	**1.71**	0.65	3.49	1.46				
13	2.84	1.38	5.10	3.19	17.94	24.06	19.34	25.61
14					4.82	1.98	8.73	4.24
Ref	4.07	1.65	5.32	2.96	4.07	2.39	5.72	3.87

Tabla 6.8: Resultados evaluación 3 BSEC 2009: impacto según información contenida

			Tableta(x,y) - Sesión 1			
			Alta entropía		Baja entropía	
Sys	**Universidad**	**Algoritmo**	**Skilled**	**Random**	**Skilled**	**Random**
1	EUP Mataró	Biometric DM	6.50	2.33	3.94	1.50
2	EUP Mataró	Estadístico	6.58	2.83	4.57	1.80
3	U1 Res. Inst.	DTW	14.00	7.22	14.50	9.98
4	Seikey Univ.	DTW	4.08	1.52	2.92	1.38
5	Ain Shams Univ.	DTW	5.67	2.55	3.14	2.47
6	Univ. Valladolid	DTW	**3.75**	0.83	1.68	0.87
7	Sabanci Univ.	DTW	4.00	1.61	2.89	2.27
8	UAM	Random tuned DTW	7.83	**0.80**	2.95	**0.27**
9	UAM	Skilled tuned DTW	4.17	1.19	2.48	1.42
10	UAM	HMM	9.92	11.27	21.18	32.43
11	UAM	Global	9.00	3.83	6.83	3.14
12	UAM	Fusion	4.17	0.91	**1.49**	0.62
Ref	TMSP	HMM	6.00	1.52	3.81	1.62

individual por lo que puede considerarse como el sistema más simple (no fusionado) y versátil (multidispositivo) de entre los presentados. En el capítulo 5 se dan los detalles de la construcción de este sistema.

6.2.5. Sistema presentado

A continuación describiremos en detalle el sistema presentado por la Universidad de Valladolid en la competición BSEC 2009. A pesar de que el sistema es básicamente el expuesto en el capítulo 5, las reglas del concurso y las bases de datos utilizadas nos obligaron a realizar pequeñas variaciones no descritas en dicho capítulo. Se realizaron las dos siguientes modificaciones al sistema base del capítulo 5 para presentarlo a la competición:

- Una nueva selección de características óptimas.
- Normalización de las puntuaciones.

La primera se realizó para obtener los mejores resultados posibles con la base de datos del concurso. La segunda modificación se debió a las propias reglas del concurso, ya que los organizadores decidieron que el error se determinaría con umbral universal. A continuación pasamos a describir en detalle cada una de estas dos modificaciones.

I. Selección de características

Para determinar el conjunto de características óptimo, realizamos pruebas siguiendo el protocolo de evaluación proporcionado por la organización de la competición. La mejor combinación de características fue realizada usando umbral individual. Estudiamos tanto

Tabla 6.9: Promedio de resultados de BSEC 2009

			Tableta			PDA			Tableta & PDA		
Sistema	**Universidad**	**Algoritmo**	**Ski**	**Ran**	**av(rd,sk)**	**Ski**	**Ran**	**av(rd,sk)**	**Ski**	**Ran**	**av(rd,sk)**
12	UAM	Fusion	**2.88**	1.00	**1.94**	5.94	1.33	**3.64**	**4.41**	1.17	**2.79**
6	Univ. Valladolid	DTW	3.04	1.35	2.20	6.78	2.02	4.40	4.91	1.69	**3.30**
9	UAM	Skilled tuned DTW	3.25	1.77	2.51	6.14	2.15	4.15	4.70	1.96	3.33
4	Seikey Univ.	DTW	3.84	2.22	3.03	7.89	2.34	5.12	5.87	2.28	4.07
7	Sabanci Univ.	DTW	3.75	2.52	3.14	**5.89**	5.22	5.56	4.82	3.87	4.35
2	EUP Mataró	Estadístico	5.44	2.75	4.10	7.38	2.49	4.94	6.41	2.62	4.52
8	UAM	Random tuned DTW	5.28	**0.86**	3.07	12.37	**0.97**	6.67	8.83	**0.92**	4.87
1	EUP Mataró	Biometric DM	5.48	2.66	4.07	10.38	2.74	6.56	7.93	2.70	5.32
Ref	TMSP	HMM	4.93	2.36	3.65	12.19	5.22	8.71	8.56	3.79	6.18
11	UAM	Global	8.15	3.55	5.85	12.92	4.21	8.57	10.54	3.88	7.21
13	Waseda Univ	DTW	11.31	13.56	12.44	7.96	2.46	5.21	9.64	8.01	8.82
5	Ain Shams Univ.	DTW	4.81	3.23	4.02	16.88	13.89	15.39	10.85	8.56	9.70
3	U1 Res. Inst.	DTW	15.63	9.91	12.77	20.44	9.71	15.08	18.04	9.81	13.92
10	UAM	HMM	21.45	22.60	22.03	24.72	25.15	24.94	23.09	23.88	23.48
14	Univ. Magdeburg	Biometric Hash	6.78	3.11	4.95						

la combinación de características como el estadístico de fusión de distancias. Respecto a las combinaciones de características, partimos un subconjunto de las combinaciones óptimas obtenidas en el capítulo 5. Para fusionar las distancias entre las cinco plantillas y la firma de prueba evaluamos tres estadísticos: la media aritmética (`avg`), el mínimo (`min`) y el máximo (`max`).

Combinación óptima para tableta gráfica. En la tabla 6.10 vemos los resultados obtenidos para cada combinación de características con la base de datos DS2 (tableta). En ella mostramos los resultados agrupados por:

- **Sesión**: dado que era uno de los objetivos de la competición, medimos el error del sistema al probar con firmas sólo de la primera sesión, sólo de la segunda y el promedio de ambas.
- **Escenario**: los organizadores indicaron que se realizarían pruebas separadas con firmas casuales e imitaciones por lo que no mezclamos los dos tipos de falsificaciones en una misma prueba.

Extraemos una serie de conclusiones interesantes de los resultados contenidos en dicha tabla:

- Fijada la combinación de características el estadístico que funciona peor es el máximo (`max`). Entre los resultados obtenidos con la media (`avg`) y el mínimo (`min`) no hay muchas diferencias siendo la media la que proporciona los mejores resultados para la mayoría de combinaciones.
- La mejor combinación para el escenario casual fue (y_p_dx_dy) con el estadístico `max`.
- La mejor combinación para escenario seguro fue la combinación (dx_dy) con el estadístico `avg`.
- La mejor combinación en promedio para los dos escenarios fue la combinación (dx_dy) con el estadístico `min`.
- Las dos mejores combinaciones en promedio sobre los dos escenarios tras (dx_dy) fueron (y_p_dx_dy_dp_ddy) y (y_p_dx_dy) con el estadístico `min`. Estas combinaciones son las que determinamos como las óptimas para escenarios seguros en el capítulo 5.

Puesto que nuestro objetivo era enviar a la competición el sistema más versátil para los dos escenarios, la combinación seleccionada para tableta gráfica fue (dx_dy) en conjunción con el estadístico `min`.

TABLA 6.10: Resultados de la selección de la combinación óptima de características para BSEC 2009 sobre el conjunto de desarrollo DS2 (50 usuarios y tableta)

Tableta Gráfica		**Sesión 1**		**Sesión 2**		**avg(sesión 1. sesión2)**			
Combinación	**Estad.**	**Skilled**	**Random**	**Skilled**	**Random**	**Skilled**	**Random**	**avg(ski. ran)**	**Observac.**
	avg	3.05	0.67	5.10	1.10	4.08	0.88	2.48	
x	min	4.05	0.80	5.70	1.03	4.88	0.92	2.90	
	max	4.50	2.23	5.95	1.97	5.23	2.10	3.66	
	avg	2.20	0.53	3.65	0.80	2.93	0.67	1.80	
y	min	2.40	0.50	4.20	0.73	3.30	0.62	1.96	
	max	3.25	0.83	5.05	1.63	4.15	1.23	2.69	
	avg	1.40	0.27	2.00	0.37	1.70	0.32	1.01	
x_y	min	1.85	0.30	2.10	0.27	1.98	0.28	1.13	
	max	1.85	0.67	3.60	0.53	2.73	0.60	1.66	
	avg	0.30	0.33	1.05	0.20	**0.68**	0.27	0.47	Mejor ski.
dx_dy	min	0.30	0.13	1.10	0.20	0.70	0.17	**0.43**	Seleccionada
	max	0.50	0.37	1.30	0.27	0.90	0.32	0.61	
	avg	3.55	0.87	4.70	1.10	4.13	0.98	2.55	
x_dx_dy	min	3.85	0.83	5.35	0.93	4.60	0.88	2.74	
	max	4.75	1.77	6.10	1.67	5.43	1.72	3.57	
	avg	1.65	0.53	4.15	1.10	2.90	0.82	1.86	
y_dx_dy	min	2.40	0.33	4.20	0.87	3.30	0.60	1.95	
	max	2.85	0.73	5.15	1.90	4.00	1.32	2.66	
	avg	1.65	0.13	2.15	0.23	1.90	0.18	1.04	
x_y_dx_dy	min	1.90	0.10	2.75	0.23	2.33	0.17	1.25	
	max	2.25	0.30	3.15	0.30	2.70	0.30	1.50	
	avg	0.80	0.07	1.95	0.67	1.38	0.37	0.87	
dx_dy_dp	min	1.15	0.27	3.20	1.40	2.18	0.83	1.50	
	max	1.25	0.10	3.00	0.53	2.13	0.32	1.22	
	avg	1.25	0.07	1.85	0.23	1.55	0.15	0.85	
y_p_dx_dy	min	1.00	0.03	1.65	0.23	1.33	0.13	0.73	
	max	1.80	0.10	3.00	0.13	2.40	**0.12**	1.26	Mejor ran.
	avg	8.15	5.00	12.95	7.90	10.55	6.45	8.50	
x_y_p_dx_dy	min	6.15	3.73	10.75	5.87	8.45	4.80	6.63	
	max	11.35	8.03	15.80	11.00	13.58	9.52	11.55	
	avg	1.10	0.03	1.70	0.27	1.40	0.15	0.78	
y_p_dx_dy_dp_ddy	min	0.85	0.07	1.65	0.23	1.25	0.15	0.70	
	max	1.80	0.10	2.65	0.30	2.23	0.20	1.21	
	avg	1.25	1.00	1.75	1.47	1.50	1.23	1.37	
dx_dy_dp_da_di	min	2.40	2.63	2.80	2.67	2.60	2.65	2.63	
	max	1.60	0.93	2.85	1.83	2.23	1.38	1.80	

TABLA 6.11: Resultados de la selección de la combinación óptima de características para BSEC 2009 sobre el conjunto de desarrollo DS3 (50 usuarios y PDA)

PDA		Sesión 1		Sesión 2		avg(sesión 1. sesión2)			
Combinación	**Estad.**	**Skilled**	**Random**	**Skilled**	**Random**	**Skilled**	**Random**	**avg(ski. ran)**	**Observaciones**
x	avg	14.69	1.50	19.80	1.53	17.24	1.51	9.38	
	min	19.03	1.43	24.80	1.94	21.91	1.68	11.80	
	max	15.20	2.21	19.23	2.52	17.22	2.36	9.79	
y	avg	16.58	1.39	21.68	1.80	19.13	1.60	10.37	
	min	22.65	1.05	26.99	1.50	24.82	1.28	13.05	
	max	18.62	1.97	23.21	3.03	20.92	2.50	11.71	
x_y	avg	12.45	0.61	17.76	1.43	15.10	1.02	8.06	
	min	17.55	0.48	23.72	1.19	20.64	0.83	10.74	
	max	13.47	1.09	16.84	2.18	15.15	1.63	8.39	
dx_dy	avg	4.08	1.05	7.09	1.39	**5.59**	1.22	**3.41**	Mejor skilled
	min	4.29	1.29	7.50	1.63	5.89	1.46	3.68	
	max	4.64	0.95	7.24	1.33	5.94	1.14	3.54	
x_dx_dy	avg	23.01	2.55	30.56	3.95	26.79	3.25	15.02	
	min	27.09	2.31	35.31	3.40	31.20	2.86	17.03	
	max	23.83	4.46	31.38	6.43	27.60	5.44	16.52	
y_dx_dy	avg	20.26	1.19	27.55	2.11	23.90	1.65	12.78	
	min	28.21	1.53	35.51	3.13	31.86	2.33	17.10	
	max	21.94	1.56	28.21	3.06	25.08	2.31	13.69	
x_y_dx_dy	avg	6.43	0.65	11.07	1.09	8.75	0.87	4.81	
	min	8.52	0.65	12.35	1.09	10.43	0.87	5.65	
	max	6.58	0.65	10.56	0.99	8.57	**0.82**	4.69	Mejor random

Combinación óptima para PDA. Para el caso de PDA al disponer únicamente de coordenadas geométricas la búsqueda de combinaciones óptimas fue más simple. En la tabla 6.11 vemos los resultados obtenidos para este caso.

En el caso de la base de datos DS3 (PDA) la mejor combinación sobre los dos escenarios (en promedio) vuelve a ser (dx_dy) con el estadístico `avg`, por lo que los utilizamos en el sistema sobre PDA. Esta combinación fue además la que mejores resultados obtuvo sobre el escenario seguro. La combinación (x_y_dx_dy) fue la combinación óptima para escenario casual, justo la combinación que encontramos óptima para dicho escenario al realizar el estudio del capítulo 5 sobre otras bases de datos.

Análisis de los errores. Es interesante ver la dependencia del error para cada tipo de imitación (escenario) con respecto a la combinación de características utilizada. En nuestro caso, buscábamos un sistema que funcionara razonablemente para cada tipo de escenario, pero tal y como se pudo ver en los resultados del concurso, la adaptación de un sistema a un escenario en concreto puede mejorar notablemente los resultados, normalmente en detrimento del rendimiento en el otro escenario.

Nótese que el error obtenido con umbral individual es muy inferior a los mejores resultados obtenidos por todos los sistemas de la competición (comparando con el conjunto de evaluación). Este error típicamente se incrementa varias veces al realizar la normalización de distancias para evaluar el error con umbral universal. Pensamos, que habría sido interesante, el que los organizadores hubieran publicado los errores de los sistemas también con

umbral dependiente del usuario, al menos en alguna de las tareas, ya que en el ámbito de firma dinámica su uso suele ser bastante habitual.

Para finalizar, apuntemos que con la existencia de un mecanismo de selección de la combinación de características y el estadístico de fusión de distancias de forma dependiente del usuario, los resultados promedio (con umbral individual) podrían haber sido de:

- Del 0.25 % en escenario seguro y 0.03 % en escenario casual con tableta gráfica.
- Del 3.55 % en escenario seguro y 0.10 % en escenario casual con PDA.

Aunque a la fecha actual no disponemos de este mecanismo, mostramos los resultados como cota de error objetivo para futuros trabajos.

II. Normalización de distancias

Los errores mostrados en el apartado anterior fueron calculados usando un umbral dependiente del usuario sin ningún tipo de normalización. Dado que las reglas del concurso indicaban que los sistemas serían evaluados con umbral universal (independiente del usuario) era necesario utilizar una estrategia de normalización de puntuaciones(e.g. distancias) que mantuviera los resultados obtenidos con el umbral dependiente de usuario.

Métodos de normalización. Fierrez-Aguilar *et al.* (2005c) analizan tres tipos de estrategias de normalización de puntuaciones, todas ellas pertenecientes a la categoría de normalización de puntuaciones dependientes del cliente (target dependent):

1. Métodos centrados en el impostor (IC, Impostor centric).
2. Métodos centrados en cliente (TC, Target centric).
3. Métodos centrados en cliente e impostor (TIC, Target-Impostor centric).

En la categoría de métodos IC la puntuación es normalizada sin utilizar para ello las puntuaciones del propio cliente. Sólo se utilizan valores estadísticos (media y varianza) de la distribución de puntuaciones de impostores. Los impostores pueden ser de dos tipos: casuales e imitadores. Nótese que en aplicaciones reales, así como en el concurso BSEC 2009, no es factible utilizar firmas de imitadores en la normalización de puntuaciones. Las ecuaciones 6.1 muestran tres formas de normalización de puntuaciones de este tipo:

$$\begin{aligned} s_{IC1} &= s - \mu_I \\ s_{IC2} &= s - (\mu_I + \sigma_I) \\ s_{IC3} &= (s - \mu_I)/\sigma_I \end{aligned} \tag{6.1}$$

donde μI y σI son respectivamente la media y desviación estándar de la distribución de puntuaciones de impostor.

Con los métodos TC sin embargo sólo se dispone de información de las puntuaciones del propio cliente. En un sistema real estas puntuaciones deben obtenerse de las propias

firmas de entrenamiento para lo que pueden aplicarse diversas técnicas de muestreo tales como rotación, substitución, validación cruzada, etc. Las ecuaciones 6.2 representan las formas análogas a las anteriores pero sobre el conjunto de puntuaciones del cliente.

$$\begin{aligned} s_{TC1} &= s - \mu_C \\ s_{TC2} &= s - (\mu_C - \sigma_C) \\ s_{TC3} &= (s - \mu_C)/\sigma_C \end{aligned} \tag{6.2}$$

Un tipo más general de normalización que las dos anteriores es la que tiene en cuenta la distribución de puntuaciones de cliente e impostor. Dos ejemplos de normalización TIC son:

$$s_{TIC1} = s - s_{EER}(C, I) \tag{6.3}$$

$$s_{TIC2} = s - (\mu_I\sigma_C + \mu_C\sigma_I)/(\sigma_C + \sigma_I) \tag{6.4}$$

donde $s_{EER}(C, I)$ es el valor de la puntuación en la que se obtiene el error en forma de EER (es decir el punto de la curva ROC en el que FAR = FRR).

Normalización empleada. La normalización de puntuaciones que finalmente adoptamos para nuestro sistema se enmarcaría en esta última categoría (TIC), aunque se trata de una versión más general de la ecuación 6.3. El método de normalización empleado fue el siguiente:

$$s_n = \frac{s - \alpha \cdot s_{EER}(C, I)}{\beta \cdot (s_{max}(C) - s_{min}(C)) + (1 - \beta) \cdot (s_{max}(I) - s_{min}(I))} \tag{6.5}$$

donde:

- α y β son coeficientes determinados experimentalmente.
- $s_{max}(C)$ y $s_{min}(C)$ son respectivamente las distancias máxima y mínima del cliente, obtenidas únicamente de las firmas de entrenamiento mediante el método *leave-one-out* (Devijver & Kittler, 1982).
- $s_{max}(I)$ y $s_{min}(I)$ son respectivamente las distancias máxima y mínima de un conjunto de cohorte formado por impostores casuales.

El numerador de la ecuación anterior realiza un alineamiento de las puntuaciones entorno al umbral en el que se obtiene la tasa de equierror (EER). Dado que *a priori* no se puede determinar este umbral lo ponderamos por el coeficiente α para dotar de mayor libertad al ajuste.

El primer término del denominador correspondería a una normalización $min - max$ (Ross *et al.*, 2006) centrada en cliente. El segundo término equivaldría a una normalización $min - max$ centrada en impostor. El término β no permitiría dar mayor importancia a uno u otro término según los resultados experimentales. De aquí en adelante denominaremos a esta normalización como $norma - EER$.

Procedimiento de cálculo de la norma-EER. Para el cálculo de los distintos términos de la ecuación 6.5 es necesario disponer de un conjunto de puntuaciones de cliente (C) y de impostores (I).

Puntuaciones del cliente. Dado que para obtener las distancias entre cliente sólo se dispone de las (cinco) firmas de entrenamiento, obtuvimos las distancias entre ellas a través del método *leave-one-out* (LOO), una variante del método más genérico conocido como *validación cruzada*. Supongamos que disponemos de N firmas de entrenamiento del cliente. Mediante LOO reservamos una de las firmas de entrenamiento como firma de prueba y las $N-1$ restantes como firmas de entrenamiento. De este modo obtenemos $N-1$ distancias con las cuales obtenemos una única distancia final mediante la media aritmética. Este proceso lo repetimos $N-1$ veces eliminando cada vez una de las firmas de entrenamiento con lo que al final obtenemos N distancias de cliente. Con estas distancias obtenemos los valores de $s_{max}(C)$ y $s_{min}(C)$, mayor y menor distancia de entre las obtenidas.

Puntuaciones de impostores. El conjunto de distancias respecto a firmas de impostor debe obtenerse respecto a firmas casuales. Ello es así por las reglas del concurso, y porque en una aplicación real no se disponen de imitaciones profesionales de las firmas de los usuarios como ocurre con las bases de datos biométricas para investigación. Dado que no queríamos que nuestro sistema dependiera del conjunto de datos de desarrollo y así evitar lo más posible la pérdida de rendimiento al ser evaluado sobre el conjunto de evaluación, decidimos utilizar una base de datos externa como fuente de las falsificaciones casuales. Seleccionamos 50 firmas de la base de datos MCYT tomadas de 50 usuarios seleccionados aleatoriamente, una firma por usuario. Con este conjunto de cohorte casual obtuvimos 50 distancias entre las firmas de entrenamiento del cliente y las 50 firmas de cohorte. De todas ellas seleccionamos la menor y mayor para establecer los términos $s_{min}(I)$ y $s_{max}(I)$ de la ecuación 6.5.

Con estos dos conjuntos de distancias s_C y s_I podemos establecer el denominador de la ecuación de la norma-EER a falta del coeficiente experimental β.

Umbral EER. Para determinar el valor *a posteriori* de la puntuación en la cual se obtiene el EER (umbral EER) es necesario realizar una estimación a partir de los valores disponibles *a priori*. Ello es debido a que el umbral EER obtenido con el conjunto de evaluación varía respecto a cuando sólo se dispone de los datos de entrenamiento y el cohorte casual.

La estimación del umbral EER *a posteriori* ($s_{EER}(C,I)$) la realizamos mediante regresión lineal múltiple[6] a partir de las siguientes variables:

1. e_{EER}: Valor del EER *a priori*.
2. t_{EER}: Umbral EER *a priori*.
3. $\mu(s_C)$: Media de puntuaciones del cliente.

[6]Estimación realizada con la herramienta Weka(Garner, 1995; Hall *et al.*, 2009)

4. $\sigma(s_C)$: Desviación estándar de las puntuaciones del cliente.
5. $\sigma_p(s_C)$: Desviación estándar porcentual de las puntuaciones del cliente respecto a la media.
6. $\mu(s_I)$: Media de las puntuaciones de impostores del cohorte casual.
7. $\sigma(s_I)$: Desviación estándar de las puntuaciones de impostores del cohorte casual.
8. $\sigma_p(s_I)$: Desviación estándar porcentual de las puntuaciones de impostores casuales respecto a la media.
9. $\texttt{max}(s_{Ca})$: Peor puntuación del cliente (máxima) de entre las aceptadas (peor firma genuina).
10. $\texttt{min}(s_{Ir})$: Mejor puntuación de impostor casual de entre las rechazadas (mejor falsificación casual).
11. s_{EER-}: Primera puntuación por debajo del umbral EER.
12. s_{EER+}: Primera puntuación por encima del umbral EER.

Sea r el vector formado por las variables anteriores, entonces el valor de s_{EER} es calculado del siguiente modo:

$$s_{EER}(C, I) = (a_1, ..., a_{12}) \cdot r + b \tag{6.6}$$

donde $(a_1, ..., a_{12})$ son los coeficientes obtenidos mediante la regresión lineal múltiple.

Dado que el umbral EER es distinto si la prueba se realiza con firmas imitadas o firmas casuales se realizaron dos estimaciones en base a las variables anteriores. Una para el cálculo del umbral EER con firmas imitadas (s_{EER-sk}) y otra para firmas casuales (s_{EER-rd}). Dado que no queríamos especializar el sistema en ninguno de los dos escenarios (seguro o casual) tomamos como estimación final del umbral EER a la media de los dos umbrales anteriores.

$$s_{EER}(C, I) = \frac{s_{EER-sk}(C, I) + s_{EER-rd}(C, I)}{2} \tag{6.7}$$

Para dotar de mayor validez a los resultados de las regresiones lineales obtuvimos puntuaciones de 30 pruebas distintas variando en cada una de ellas las firmas de entrenamiento de forma aleatoria. Las tabla 6.12 muestra los coeficientes de la ecuación 6.6 para cada escenario (skilled, random) y dispositivo de captura (DS2, DS3), así como el coeficiente de correlación R de cada una de ellas, que como puede verse son bastante elevados lo que confirma la validez del método.

Resultados con el sistema enviado. Seleccionados los valores de la ecuación 6.5 los valores óptimos de α y β de la misma se determinaron empíricamente. Para ello, determinamos el valor del error con la misma técnica de *bootstrapping* con la que realizamos la regresión lineal múltiple para el cálculo del umbral EER ($s_{EER}(C, I)$). Analizamos la dependencia del error en tres escenarios: casual, seguro y mixto. En el escenario casual las

TABLA 6.12: Coeficientes de regresión lineal múltiple para obtener la estimación del umbral EER *a posteriori*

	DS2 (tableta)		**DS3 (PDA)**	
	Skilled	**Random**	**Skilled**	**Random**
a1	1.0154	-0.4789	-0.8849	-0.3119
a2	0.0000	0.4095	0.0000	-0.7189
a3	0.5940	0.2726	0.6989	0.5671
a4	-1.6358	0.0000	-1.9862	-0.3034
a5	0.0000	-0.3371	0.1153	0.0000
a6	0.0000	0.1943	-0.4129	0.1499
a7	0.0000	0.1916	0.0000	-1.2294
a8	-1.5559	-0.7298	-0.3093	2.1875
a9	0.1795	-0.0408	0.0000	0.0000
a10	1.0376	0.0452	0.0000	-0.1536
a11	-0.5807	-0.2705	0.2054	0.9825
a12	-0.2587	0.5752	0.4003	0.3174
b	109.03	54.30	77.10	-9.08
R	0.9454	0.9885	0.9676	0.9851

firmas de impostor fueron obtenidas de los otros usuarios distintos al cliente evaluado. Se utilizaron 2 firmas por cada usuario de modo que se utilizaron un total de $49 \times 2 = 98$ falsificaciones casuales por usuario. El escenario seguro fue evaluado únicamente con las firmas imitadas disponibles, en total 20 firmas por usuario. El escenario mixto fue evaluado con los dos tipos de falsificaciones (casuales e imitaciones), en total $20 + 98 = 118$ firmas por usuario. Como firmas de cliente se utilizaron las firmas no utilizadas como plantillas, 30 - 5 = 25 firmas por usuario.

Las gráficas de las figuras 6.6 y 6.7 reflejan la evolución del error en función de los parámetros α y β para cada dispositivo y escenario. Variamos α entre los valores 0.5 y 1.4 y β entre 0 y 1. Nótese que cuando $\beta = 1$ estamos en el caso de normalización centrada en cliente (TC), mientras que cuando $\beta = 0$ se trata de normalización centrada en impostor (casual). Los valores óptimos se muestran en la tabla 6.13.

De estas gráficas y de los valores de la tabla extraemos varias conclusiones:

- El parámetro α es el que más influencia tiene al obtener el error con umbral universal. La dependencia con β es más pronunciada con valores extremos de α. Nótese que los valores óptimos se encuentran entorno a $\alpha = 0{,}9$.

- En escenario casual la dependencia con α y β es menos importante que en el escenario seguro.

- Con firmas de PDA la dependencia con α y β es más importante que con firmas de tableta.
- La degradación del error es más pronunciada (en valores porcentuales) con firmas de tableta que con firmas de PDA.
- Con ambos tipos de firma los valores óptimos se obtienen con valores similares de los parámetros α y β.
- En el escenario casual, los mejores resultados se obtienen con normalización centrada en cliente (valores de β cercanos a uno). Es escenario seguro sin embargo, los mejores resultados se obtienen con normalización centrada en impostor (valores de β cercanos a cero).
- La norma EER propuesta funciona mejor para escenario casual que para escenario seguro en el que todavía existe un amplio margen de mejora.

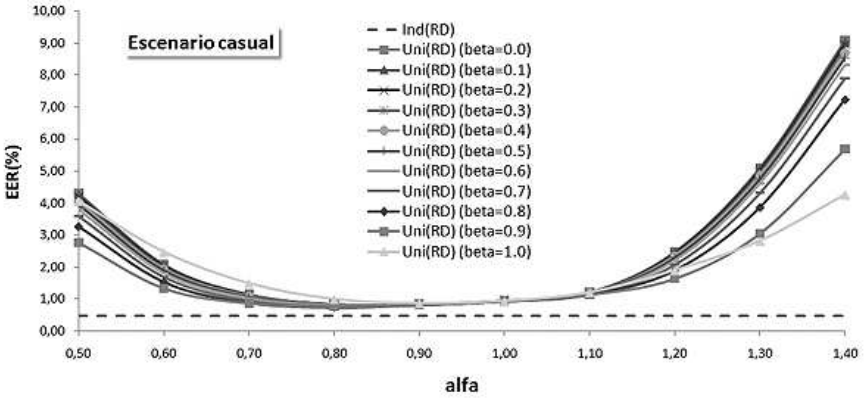
Tabla 6.13: Resultados óptimos del sistema enviado a BSEC 2009 con umbral universal utilizando la norma-EER

	DS2 (tableta)			**DS3 (PDA)**		
	Casual	**Seguro**	**Mixto**	**Casual**	**Seguro**	**Mixto**
% EER Umb. indiv	0.47	0.77	0.60	1.95	5.98	3.04
% EER Umb. univ.	0.75	1.96	1.18	2.94	9.12	4.46
(α, β)	(0.8, 0.8)	(0.9, 0.0-0.1)	(0.8, 1.0)	(0.9, 0.9)	(0.9, 0.0-0.7)	(0.9, 1.0)
% incremento error	60 %	155 %	97 %	51 %	53 %	47 %

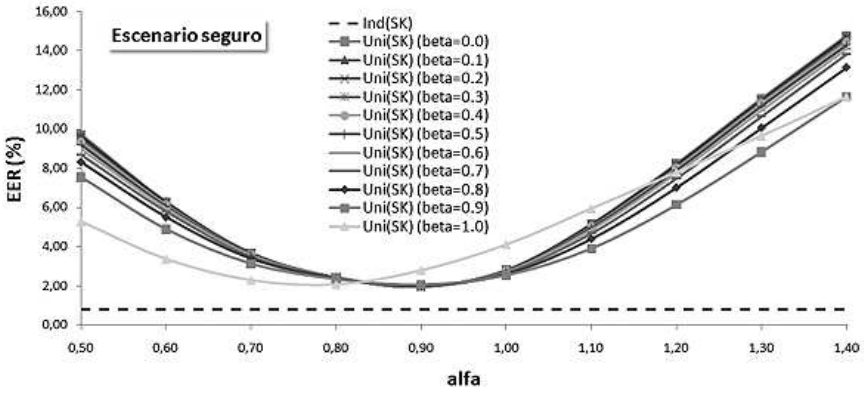
En el caso concreto del concurso el sistema enviado fue configurado con los valores óptimos para el escenario mixto para dotar al sistema de mayor versatilidad.

Influencia del umbral EER. Para comprobar la influencia del valor del umbral EER en el rendimiento del proceso de normalización representaremos a continuación los resultados obtenidos bajo las siguientes circunstancias:

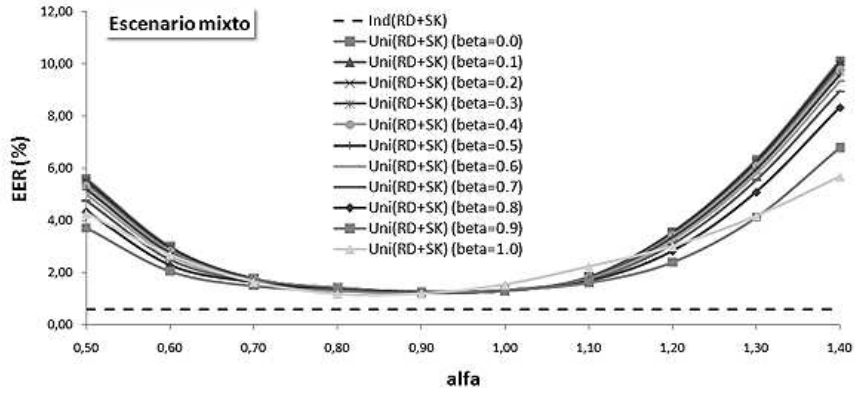
- Resultados con *umbral individual*: indican la capacidad de separación de firmas genuinas de firmas de impostores. Representan la mayor tasa de rendimiento del sistema y el objetivo de la normalización es permitir tasas de error lo más cercanas posible a estas.
- Resultados con *umbral universal*: indican la capacidad de separación de firmas genuinas de firmas de impostores con el mismo umbral para todos los usuarios. Dentro de esta categoría veremos los casos *a priori* y *a posteriori*:
    - *A posteriori*: valores obtenidos a modo de referencia para evaluar la bondad de la estimación realizada de los métodos *a priori*. Evaluamos los casos en los que el umbral se obtiene para los escenarios casual y seguro.



(a) EER(u.indiv.) = 0.47. EER(u.univ.) = 0.75 % con $(\alpha, \beta) = (0.8, 0.8)$

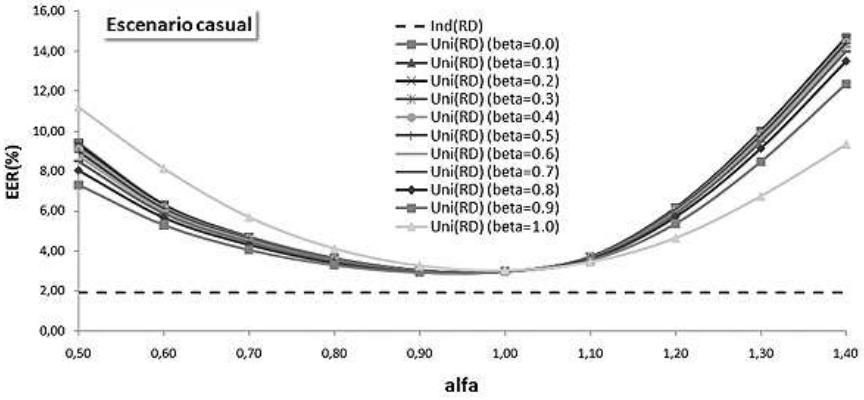


(b) EER(u.indiv.) = 0.77. EER(u.univ.) = 1.96 % con $(\alpha, \beta) = (0.9, 0.0\text{-}0.1)$

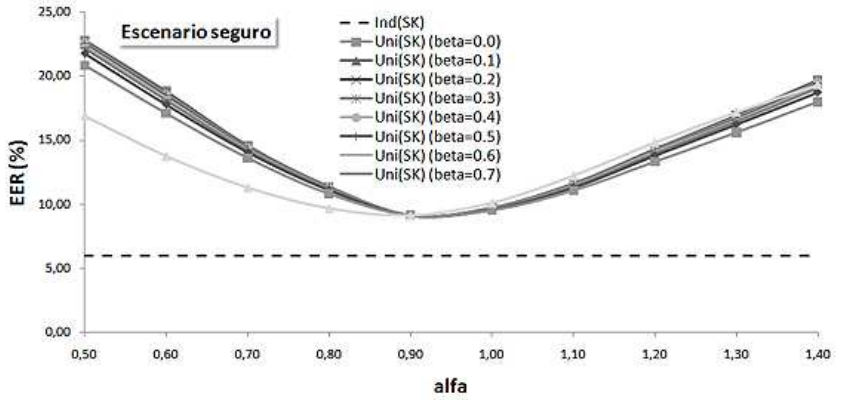


(c) EER(u.indiv.) = 0.60. EER(u.univ.) = 1.18 % con $(\alpha, \beta) = (0.8, 1.0)$

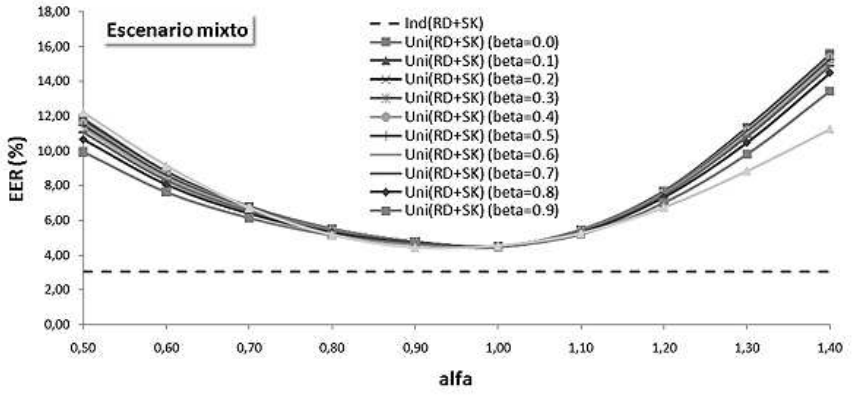
FIGURA 6.6: Evolución del error respecto a los parámetros α y β en escenario casual para DS2 (tableta)



(a) EER(u.indiv.) = 1.95. EER(u.univ.) = 2.94 % con $(\alpha, \beta) = (0.9, 0.9)$



(b) EER(u.indiv.) = 5.98. EER(u.univ.) = 9.12 % con $(\alpha, \beta) = (0.9, 0.0\text{-}0.7)$



(c) EER(u.indiv.) = 3.04. EER(u.univ.) = 4.46 % con $(\alpha, \beta) = (0.9, 1.0)$

FIGURA 6.7: Evolución del error respecto a los parámetros α y β en escenario casual para DS3 (PDA)

- *A priori*: valores sin tener en cuenta los datos de prueba al realizar la normalización. Dentro de este tipo tenemos los siguientes casos.
  - *Sin estimación* del umbral EER: cuando para el valor de s_{EER} de la ecuación 6.5 tomamos directamente el valor del umbral EER de las puntuaciones de cliente (mediante LOO) y las falsificaciones (casuales) del cohorte.
  - *Con estimación* del umbral EER: cuando el valor de s_{EER} es calculado en escenarios seguro, casual y mixto.

Los resultados de todos estos casos se muestran en las tablas 6.14 y 6.15.

TABLA 6.14: EER del sistema enviado a BSEC 2009 medido con umbral individual para los distintos escenarios y tipos de dispositivo sobre el conjunto de desarrollo

DS2 (tableta)			**DS3 (PDA)**		
Casual	**Seguro**	**Mixto**	**Casual**	**Seguro**	**Mixto**
0.47	0.77	0.60	1.95	5.98	3.04

TABLA 6.15: EER del sistema enviado a BSEC 2009 medido con umbral universal para los distintos escenarios y tipos de dispositivo sobre el conjunto de desarrollo

		DS2 (tableta)			**DS3 (PDA)**		
Tipo	**Observaciones**	**Casual**	**Seguro**	**Mixto**	**Casual**	**Seguro**	**Mixto**
Priori	Sin estimación	0.76	2.85	1.42	3.26	10.37	5.07
	Est. para esc. casual	0.75	2.75	1.39	2.97	9.89	4.66
	Est. para esc. seguro	0.88	1.64	1.16	3.01	8.67	4.44
	Est. para esc. mixto	0.75	1.96	1.18	2.94	9.12	4.46
Post.	Optim. para esc. casual	0.52	3.24	1.30	2.19	10.63	4.26
	Optim. para esc. seguro	1.12	0.97	1.19	3.69	6.62	4.75

Centrándonos en los resultados con tableta, en primer lugar de la tabla anterior podemos ver que si fuéramos capaces de conocer exactamente el umbral EER *a posteriori*, es decir el umbral EER tras la prueba del sistema, los resultados obtenidos con umbral universal podrían estar muy cerca de los resultados con umbral individual. En concreto el error con umbral universal en escenario casual podría ser del 0.52 %, muy cercano al 0.47 % con umbral individual. El el caso del escenario seguro el incremento es desde el 0.77 % con umbral individual al 0.97 % con umbral universal. Ambos resultados se obtienen con umbrales optimizados a cada escenario. Obviamente, este umbral *a posteriori*, no puede conocerse de forma exacta hasta que se ha realizado la prueba. Por ello, realizamos un estimación del mismo basándonos en los datos *a priori*.

Para comprobar la conveniencia de la estimación del umbral EER, mostramos el resultado si no se hubiera realizado ninguna estimación. En vez de ello se utilizaría el umbral EER a partir de las distribuciones de puntuaciones de distancias intraplantillas y firmas casuales del conjunto de cohorte. Según vemos en la segunda fila de la tabla, el resultado

para el escenario casual sería prácticamente igual que el obtenido con la estimación, por lo que para dicho escenario podríamos simplificar y no realizar la normalización. Sin embargo, para el caso del escenario seguro, el error obtenido prácticamente triplica el error de referencia (2.85 % versus 0.97 %). Con firmas de PDA la estimación siempre mejora los resultados respecto a los obtenidos sin ella.

Como puede verse, la estimación que menos pérdidas de rendimiento produce al pasar al umbral universal es la del umbral EER optimizado para escenario seguro, mejorando un 1.21 % de EER en valor absoluto, cifra que en esas pequeñas magnitudes de error es difícil de conseguir. Dado que al optimizar para un escenario, los resultados en el otro empeoraban, realizamos la optimización de nuestro sistema para un escenario mixto. La tasa de error para falsificaciones casuales se mantuvo baja, pero el error obtenido con firmas imitadas se incrementó ligeramente. Así pues, nuestros valores de referencia del sistema enviado fueron de 0.75 % en escenario casual y 1.96 % en escenario seguro. Si comparamos estas cifras con las de los resultados finales de la tabla 6.6 observamos un ligero incremento de aproximadamente dos décimas en cada uno de ellos. Pese a todo, es difícil inferir los resultados que se obtendrían si se hubiera utilizado otra configuración, al no disponer de la base de datos de evaluación.

6.3 Resumen

Se han descrito dos escenarios prácticos realizados durante la presente tesis. El primero de ellos, como organizadores de un concurso de imitadores de firmas, ha revelado que las actuales cifras de error obtenidas con las firmas imitadas de las bases de datos públicas, tales como MCYT o SVC, no reflejan del todo el riesgo de ataque por posibles falsificadores profesionales. Pensamos que, en este sentido, una medida de fortaleza de firma ante ataques, al igual que ya existe por ejemplo para las contraseñas (passswords), ayudaría a establecer condiciones más realistas y prácticas para fomentar el empleo de la firma en sistemas biométricos.

La segunda experiencia es la participación en la competición BSEC 2009. Hemos descrito el protocolo y los resultados de dicha competición junto al sistema con el cual hemos participado y con el que obtuvimos resultados destacables.

En el siguiente capítulo se concluirá la presente memoria con las conclusiones del trabajo realizado y los planes de trabajo para el futuro que continúen el iniciado en esta tesis doctoral.

7 Conclusión

La tesis que hemos presentado a lo largo de los capítulos precedentes ha estudiado el uso de la firma manuscrita dinámica como medio de reconocimiento biométrico de personas en escenarios prácticos, demostrándose que es posible obtener tasas de error comparables a las de otras modalidades biométricas.

Es indudable que la firma manuscrita es el mecanismo no automatizado de verificación de la identidad de personas más usado en la actualidad y, según informes de mercado (International Biometric Group, 2007), es la segunda modalidad en importancia en Biometría conductual, justo detrás de la voz.

Con los sistemas desarrollados hemos demostrado que es posible aplicar esta tecnología en condiciones de aplicación realista, por lo que esperamos que esta tesis ayude a que con el tiempo la firma manuscrita sea utilizada en aplicaciones de la vida cotidiana, como ya ocurre con otros rasgos, como por ejemplo, la huella dactilar.

Los sistemas propuestos están basados en las dos técnicas más utilizadas en el campo del reconocimiento de firma y, en cada uno de ellos, hemos realizado aportaciones que mejoran o dan otro punto de vista a la hora de aplicar estos métodos. Siempre se ha tenido en cuenta las condiciones en las que un sistema real de este tipo debería aplicarse, por lo que el número de firmas de entrenamiento ha sido siempre reducido. Se ha buscado crear sistemas simples que sean fácilmente implantables en diversos tipos de soluciones tecnológicas y múltiples dispositivos.

7.1 Conclusiones

Con el trabajo realizado en esta tesis hemos comprobado que la firma manuscrita dinámica es una modalidad biométrica que puede ser usada para reconocer personas eficientemente, a pesar de sus inconvenientes respecto a otras modalidades fisiológicas. Las tasas de error obtenidas por nuestros sistemas se encuentran entorno al 1-2 % en reconocimiento casual y al 2-5 % con imitaciones. Comparando estas cifras con las obtenidas con otras modalidades biométricas (tabla 7.1) comprobamos que la firma se encuentra en los mismos rangos de magnitud.

TABLA 7.1: Tasas de error de la firma dinámica junto a las obtenidas con otras modalidades biométricas

Rasgo	**Fuente (referencia)**		**FRR**	**FAR**
Huella	FVC 2006	(Capelli et al, 2007)	EER: 2.155 %	
Cara	FRVT 2006	(Phillips et al, 2007)	0.01 %	0.001 %
Voz	NIST 2008	(Kajarekar et al, 2009)	EER: 1.954 %	
Iris	ITIRT 2005	(Int. Biometric Group, 2005)	0.99 %	0.94 %
Iris	ICE 2006	(Phillips et al, 2007)	0.09 %	0.001 %
Firma	SVC 2004	(Yeung et al, 2004)	EER: 2.8 % [1]	
Firma	BSEC 2009	(Dorizzi et al, 2009)	EER: 2.2 % [2]	

[1] Mejor resultado con imitaciones sobre el conjunto de evaluación
[2] Mejor resultado de la evaluación 1 con tableta gráfica e imitaciones (obtenido por el sistema de la Univ. de Valladolid presentado en esta tesis)

Los sistemas que hemos desarrollado han sido concebidos para ser empleados en entornos realistas y con esa premisa han sido diseñados, implementados y evaluados. Hemos decidido iniciar el trabajo a partir de los dos métodos que más éxito han tenido en este campo durante las últimas décadas, los modelos ocultos de Markov (HMM) y el alineamiento temporal dinámico (DTW). A partir de las versiones básicas de estos dos métodos hemos desarrollado variaciones que nos han permitido mejorar el rendimiento.

Para su aplicación a entornos prácticos hemos identificado una serie de condiciones que consideramos deben tener los sistemas biométricos fuera del ámbito del laboratorio, de las cuales destacamos las cuatro siguientes: a) el número de muestras de entrenamiento, b) la universalidad de las características, c) la adaptabilidad al firmante y d) la adaptabilidad a los requisitos de seguridad.

El número de firmas usadas para registrar al usuario se ha mantenido entre tres y cinco, lo que puede considerarse un valor de compromiso entre el rendimiento del sistema y el coste de inscripción.

Hemos utilizado un número reducido de características de la firma, buscando que sean fáciles de calcular y con carácter universal para facilitar la implantación del sistema en un mayor rango de dispositivos de captura (tabletas gráficas, PDA, Tablet PC, lápices electrónicos, etc.). El sistema HMM, descrito en el capítulo 4, ha obtenido tasas de error a nivel del estado del arte con sólo tres firmas de entrenamiento, utilizando únicamente las coordenadas geométricas en los dominios de posición y velocidad. Hasta donde nuestro conocimiento alcanza, sólo tenemos constancia de otro trabajo basado en HMM con tres firmas de entrenamiento, destacando que nuestro sistema consigue una mejora significativa del rendimiento respecto a éste en el caso de la prueba con imitaciones (Fierrez-Aguilar *et al.*, 2005a). El sistema DTW, con cinco firmas de entrenamiento, ha obtenido tasas de error que pueden considerarse entre las mejores del estado del arte actual, utilizando sólo dos características geométricas de la firma (en el dominio de la velocidad). Todos estos resultados los hemos obtenido usando una metodología de evaluación objetiva bajo condiciones de uso realista. Además, el sistema DTW ha sido evaluado externamente durante la competición BSEC 2009, certificando la independencia y validez de nuestro procedimiento

de evaluación.

La adaptabilidad al firmante la hemos llevado a cabo de forma práctica en el sistema HMM y de forma teórica en el sistema DTW. En el sistema HMM, la personalización de la estructura del modelo nos ha permitido obtener un rendimiento que otros sistemas basados en HMM sólo habían conseguido utilizando un número mayor de características. Esta personalización estructural tiene el inconveniente de requerir un método para seleccionar la estructura del modelo a partir de las firmas de entrenamiento. Hemos comprobado que existía una fuerte relación lineal entre el número de estados del modelo y el número de vectores de la firma, por lo que con una simple regresión lineal la selección ha sido realizada *a priori*. El rendimiento del sistema DTW ha superado al del sistema HMM sin necesidad de adaptación al firmante, a pesar de utilizar dos firmas de entrenamiento más. No obstante, se ha comprobado que con combinaciones de características dependientes del usuario el sistema puede mejorar su rendimiento sensiblemente. No hemos encontrado un método *a priori* para determinar la combinación óptima para cada usuario, aunque hemos simplificado esta tarea de cara a futuros trabajos reduciendo el espacio de búsqueda a sólo cinco combinaciones, sin pérdida significativa de rendimiento obtenido *a priori*.

La vulnerabilidad ante ataques de potenciales falsificadores tiene gran importancia en los sistemas de firma manuscrita cuando son llevados al banco de pruebas del uso diario. Por ello, los dos sistemas desarrollados han sido evaluados en dos escenarios opuestos, desde el punto de vista de la seguridad: un escenario casual, en el que no se intenta atacar al sistema de forma deliberada, y un escenario adverso (seguro), en el que el sistema es atacado con imitaciones. Como es obvio, los resultados obtenidos en escenario casual superan usualmente a los del escenario adverso, aunque esta tendencia depende del conjunto de características utilizado, pudiéndose incluso invertir la tendencia con determinadas combinaciones. Hemos comprobado que los sistemas se pueden adaptar a los requerimientos de seguridad del escenario a distintos niveles. Con el sistema DTW lo hemos hecho mediante la selección de combinaciones óptimas adaptadas a la seguridad del entorno. Dicho sistema también ha podido adaptarse al nivel de seguridad del entorno mediante la normalización de las puntuaciones, al usar umbral universal en vez de individual. Por ejemplo, el sistema enviado a BSEC 2009 puede obtener un error (EER) variable dependiendo del grado de adaptabilidad al escenario de aplicación, entre 1.64 % y 2.75 % con imitaciones y entre 0.75 % y 0.88 % con firmas casuales, sobre la base de datos DS2 (tableta gráfica).

El sistema DTW ha sido utilizado en la aplicación FirmWeb, disponible en línea en `http://www.greidi.uva.es/JRBP08_firmas/`. Esta aplicación web permite capturar la firma tanto con tableta gráfica (WACOM) como con dispositivos de tipo puntero (incluso ratón). Además, implementa las operaciones habituales de un sistema de reconocimiento de firma dinámica, esto es, el registro en el sistema y la verificación de la identidad del cliente para permitir o denegar el acceso a contenido protegido. Durante las JRBP08 se organizó un concurso de imitaciones para el cual se desarrolló una extensión de FirmWeb con la que se sometió al sistema a ataques no controlados. La evaluación de los resultados del concurso demostró que con un entrenamiento suficiente y mediante la observación de la dinámica de la firma a imitar, el intruso puede llegar a perfeccionar su técnica de imitación y atacar con éxito al sistema. Sin embargo, la facilidad para el ataque depende fuertemente de la complejidad de la firma atacada. Creemos que el desarrollo de una medida de la complejidad/robustez de la firma favorecería la implantación de estos sistemas en

aplicaciones cotidianas. En este sentido, un sistema real debería comprobar si la firma del cliente es lo suficientemente robusta como para permitir su inscripción en el sistema.

A continuación proponemos ésta y otras líneas de investigación para continuar el trabajo comenzado en esta tesis.

7.2 Próximos trabajos

Del trabajo realizado en esta tesis han surgido varias líneas de investigación que describimos a continuación:

Selección de características. Hemos demostrado (sección 5.3.3) que se pueden mejorar significativamente los resultados obtenidos por nuestro sistema basado en DTW mediante la asignación de combinaciones de características personalizadas al usuario. Sin embargo, queda pendiente el desarrollo de un método que permita *a priori* encontrar esta combinación óptima del usuario.

Normalización de puntuaciones. La norma EER, introducida en la sección 6.2.5, ha demostrado ser efectiva para nuestro sistema, al menos con los datos de BSEC 2009. Para dotar de mayor validez a los resultados obtenidos, sería preciso realizar una evaluación más completa con otras bases de datos de firmas, así como compararla con otros métodos de normalización de puntuaciones.

Medidas de calidad. El análisis de resultados de la competición de imitaciones (sección 6.1) sugiere que la creación de una medida de fortaleza de la firma asistiría a los usuarios en la etapa de inscripción. Quizás, para algunos usuarios con firma poco robusta, sería conveniente solicitar un cambio de firma o su sustitución por el nombre manuscrito, aunque habría que analizar la influencia de la legibilidad del nombre manuscrito ante posibles ataques. El establecimiento de umbrales de decisión dependientes de la calidad también puede ser interesante en esta misma dirección. Brault & Plamondon (1993) propusieron una medida de complejidad de curvas manuscritas que podría servir como punto de partida de esta nueva línea de trabajo.

Otra consecuencia extraída de la competición es que no todas las imitaciones de las bases de datos públicamente disponibles tienen la misma calidad. Una medida de la calidad de la imitación, y por ende de la calidad de las imitaciones de la base de datos considerada globalmente, podría ser de interés para comparar de forma más objetiva los resultados obtenidos con imitaciones con distintas bases de datos.

Fusión intramodal. Como quedó demostrado en BSEC 2009, la fusión de sistemas especializados en tareas y/o escenarios determinados, puede conducir a importantes mejoras en el rendimiento. En esta línea, sería interesante comprobar los resultados obtenidos por los sistemas desarrollados en esta tesis bajo diferentes estrategias de fusión.

Vulnerabilidad ante ataques. La continua aparición de dispositivos candidatos a ser empleados en sistemas de reconocimiento de firma dinámica, especialmente en condiciones móviles (PDA, Smart phones, etc.), necesita del estudio de la seguridad de las aplicaciones desarrolladas sobre ellos. La ausencia de características clave para discernir entre firmas auténticas e imitaciones, tales como la presión, hace necesario analizar en detalle la vulnerabilidad ante ataques en estos nuevos escenarios.

Métodos eficientes. La implantación de los sistemas biométricos a larga escala genera problemas que no aparecen en el laboratorio de investigación. En reconocimiento de firma dinámica, uno de los hándicaps más importantes en este aspecto, es el almacenamiento de las firmas. Al almacenar la firma del cliente en su totalidad, disminuye el error del sistema, pero a costa de un mayor coste computacional y además comprometiendo la seguridad del sistema. La propuesta de métodos que generen tasas aceptables de error con datos reducidos o incompletos es una línea de investigación en la que ya hemos dado los primeros pasos (Faundez-Zanuy, 2007; Faundez-Zanuy & Pascual-Gaspar, 2009; Vivaracho-Pascual *et al.*, 2009).

Apéndices

A
Glosario

Acrónimos

BSEC: BioSecure Evaluation Campaign. Campaña de evaluación de firmas manuscritas dinámicas obtenida con dos sensores diferentes sobre el mismo conjunto de personas. Más información en
`http://biometrics.it-sudparis.eu/BSEC2009/`.

DET: Detection Error Tradeoff. Representación gráfica de las tasas de error FAR y FRR en escala logarítmica. Permite la comparación de sistemas de forma más sencilla que con curvas ROC, puesto que con esta nueva escala no lineal las curvas pasan a ser prácticamente lineales, lo que permite una mejor comparación para todo el rango de umbrales de decisión.

DTW: Dynamic Time Warping *(Alineamiento Temporal Dinámico)*. Algoritmo que mide la deformación entre dos secuencias de vectores de distinta longitud mediante técnicas de programación dinámica.

EER: Equal Error Rate (*Tasa de Equierror*). Medida usual en Biometría para determinar el rendimiento de verificación con un único número. Dentro de la una curva ROC o DET es el punto en el que los valores de FAR y FRR coinciden. Por lo general, cuánto más bajo sea, mayor será el rendimiento del sistema.

FAR: Tasa de falsas aceptaciones (*False Acceptance Rate*). Medida usada para determinar el rendimiento biométrico en la tarea de verificación. Es el porcentaje de veces que un sistema produce una falsa aceptación, lo cual ocurre cuando un individuo es erróneamente asociado a la información biométrica existente de otra persona. También denominado error de tipo II.

FTA: Failure to Acquire *(Error de captura)*. Error del sistema biométrico durante la captura o extracción de información de la muestra biométrica.

FTE: Failure to Enroll *(Error de inscripción)*. Error del sistema biométrico en la creación de una referencia durante la fase de inscripción. Es debido usualmente a que los usuarios no están habituados a usar el sistema o a que los sensores no capturan la información con la calidad necesaria para la inscripción.

FERET: FacE REcognition Technology program *(Programa de tecnología para el reconocimiento de cara)*. Programa de desarrollo y evaluación de reconocimiento de cara promovido por el gobierno de los Estados Unidos desde 1993 hasta 1997. Más información en
`http://www.frvt.org/FERET/default.htm`.

FpVTE: Fingerprint Vendor Technology Evaluation (2003) *(Evaluación tecnológica comercial de huellas dactilares)*. Evaluación independiente de soluciones tecnológicas comerciales para reconocimiento de huella dactilar. Más información en `http://fpvte.nist.gov`.

FRGC: Face Recognition Grand Challenge *(Gran desafío en reconocimiento de cara)*. Programa de desarrollo de reconocimiento de cara organizado por el gobierno de los Estados Unidos desde 2003 a 2005.

FRR: Tasa de falsos rechazos (*False Rejection Rate*). Medida utilizada para determinar el rendimiento biométrico en la tarea de verificación. Es el porcentaje de veces que el sistema rechaza de forma errónea a un usuario, esto es, cuando un individuo no es asociado con su plantilla biométrica existente. También denominado error de tipo I.

FRVT: Face Recognition Vendor Test *(Evaluación comercial de sistemas de reconocimiento de cara)*. Serie de evaluaciones de tecnología independientes a gran escala de sistemas para reconocimiento de cara (2000, 2002, y 2005). Más información en `http://www.frvt.org/FRVT2005/`.

GMM: Gaussian Mixture Models *(Modelos de mezclas de gaussianas)*. Técnica estadística empleada para modelar la distribución estadística de una señal mediante una combinación lineal de distribuciones gaussianas.

LTR: Left-To-Right (topología de HMM). Topología de interconexión de estado muy usada en HMM en la que un estado esta conectado únicamente consigo mismo y con el estado siguiente.

HMM: Hidden Markov Model (*Modelo Oculto de Markov*). Técnica estadística para modelar un proceso aleatorio mediante un número finitos de estados interconectados.

NIST: National Institute of Standards and Technology *(Instituto Nacional de Estándares y Tecnología)*. Organismo federal, no regulador, perteneciente a la Cámara de Comercio de los Estados Unidos que desarrolla y promueve medidas, estándares y tecnología para aumentar la productividad, facilitar el comercio y mejorar la calidad de vida. La tarea sobre medidas y estándares que realiza el NIST promueve el bienestar de la nación y ayuda a mejorar, entre otras cosas, la seguridad interna de la nación. Más información en `http://www.nist.gov`.

PDA: Asistente Personal Digital (*Personal Digital Assistant*). Dispositivo móvil que permite la entrada de datos de firma dinámica.

ROC: Receiver Operation Curve. Método gráfico para medir el rendimiento de un sistema biométrico representando las curva de FAR frente a la de FRR.

SVC: Signature Verification Competition (*Competición de Verificación de Firma*). Primera competición de verificación de firma dinámica llevada a cabo en 2004.

VAFD: Verificación Automática de Firma Dinámica. Implícitamente referida a la modalidad dinámica.

Términos

Ataque: Intento deliberado de acceso a un sistema biométrico con muestras biométricas falsas.

Autenticación: En Biometría suele usarse como sinónimo genérico de **verificación**.

Banco de pruebas: Base de datos disponible públicamente sobre la que evaluar un sistema biométrico. También se refiere a competiciones o evaluaciones externas.

Biometría: Término general utilizado para describir un proceso automatizados de reconocimiento de un individuo a partir de sus características fisiológicas y/o de comportamiento.

Benchmark: Proceso mediante el cual se compara el rendimiento medido con un valor de referencia estándar, disponible al público.

Bootstrapping: método de remuestreo utilizado para aproximar la distribución en el muestreo de un estadístico. Se usa para aproximar la varianza de un estadístico, así como para construir intervalos de confianza o realizar contrastes de hipótesis sobre parámetros de interés.

Cliente: Usuario (ver **usuario**) que realmente es quien dice ser.

Conjunto de cohorte: Conjunto pequeño de individuos utilizado para obtener una estimación de un parámetro de interés sin tener que calcularlo con toda la población. Se utiliza para reducir el coste computaciones de los algoritmos.

Conjunto de desarrollo: Subconjunto de una base de datos de rasgos biométricos utilizada para configurar el sistema biométrico. Normalmente es un porcentaje menor al 50 % del tamaño total de la base de datos.

Conjunto de evaluación: Subconjunto de una base de datos de rasgos biométricos utilizada para evaluar el sistema biométrico. Normalmente es un porcentaje superior al 50 % del tamaño total de la base de datos.

Contramedidas: Acciones tomadas por los sistemas biométricos para hacer frente a ataques. Por ejemplo, tomar la temperatura del dedo al capturar la huella dactilar para verificar que pertenece a un ser humano vivo.

Escenario casual: se refiere a la evaluación de un sistema de reconocimiento de firma usando firmas casuales de otros usuarios. También denominado escenario *random*.

Escenario seguro: También denominado escenario de ataque, adverso o adiestrado. Se refiere a la evaluación de un sistema de reconocimiento de firma usando imitaciones de firmas para evaluar la falsa aceptación.

Falsificación casual: Firma generada por un **impostor casual**.

Fase de entrenamiento: Etapa de construcción del sistema biométrico en la que se determinan los parámetros que forman el modelo del usuario. En esta fase sólo deben usarse las firmas de entrenamiento del propio cliente.

Fase de prueba: Etapa de evaluación del sistema biométrico. Para ello se utilizan las **firmas de prueba**.

Firma dinámica: este tipo de firma es capturada usando dispositivos especiales con capacidad de registrar la evolución temporal de varias señales generadas por el lápiz al firmar. En inglés *on-line*.

Firma estática: corresponde a la digitalización de una firma manuscrita a partir de una muestra obtenida en papel. En inglés *off-line*.

Firma de entrenamiento: Firma de referencia. Ver término **referencia**.

Firma de prueba: Firma no utilizada en el entrenamiento del sistema que es utilizada para medir el rendimiento del sistema biométrico de firma. Puede ser de dos tipos, auténtica, en cuyo caso sirve para medir la tasa de falsos rechazos y falsa, para medir la tasa de falsas aceptaciones. En una evaluación honesta no deben utilizarse las firmas de prueba para entrenar el sistema ni las de entrenamiento para medir el error final.

Identificación: Tarea en la cual el sistema biométrico busca en una base de datos una referencia que coincida con la muestra biométrica suministrada y, de encontrarla, devuelve la identidad correspondiente. La identificación es sobre 'conjunto cerrado' si se sabe que la persona forma parte de la base de datos. La identificación es sobre 'conjunto abierto' (también denominada 'lista de vigilancia') si no existe garantía de que la persona sea parte de la base de datos. El sistema debe determinar si la persona es parte de la base de datos y luego devolver la identidad.

Imitador: Impostor (ver **impostor**) que intenta suplantar la identidad de un individuo de forma deliberada mediante una copia ilegítima del rasgo del cliente. En el caso de la firma se trata de un usuario que intenta acceder al sistema imitando la firma del individuo al que intenta suplantar. En inglés *skilled forgery*.

Imitación: Firmas generadas por un usuario **imitador**.

Impostor: Persona que utiliza una muestra biométrica con la intención deliberada o involuntaria de declarar la identidad de otra persona en el sistema biométrico.

Impostor casual: Impostor (ver **impostor**) utilizado para medir el rendimiento del sistema cuando se intenta suplantar la identidad de un individuo con la firma propia, es decir sin intentar imitar la firma del usuario suplantado. En inglés *random forgery*.

Inscripción: Ver *registro*.

Matching: Proceso que incluye la comparación de una muestra biométrica con una plantilla almacenada anteriormente y el cálculo de su grado de semejanza.

Modalidad: Tipo o clase de sistema biométrico. Por ejemplo: reconocimiento de cara, reconocimiento de huellas dactilares, reconocimiento de iris, etc.

Modelo: Representación tipo estadística utilizada para caracterizar a un individuo. Usualmente utilizada por sistemas biométricos basados en características comportamentales.

Muestra biométrica: Característica comportamental o fisiológica obtenida del usuario, ya sea en la fase de inscripción o en la de operación (identificación o verificación).

Multimodalidad: Sistema biométrico en el cual dos o más de los componentes de la misma o distinta modalidad se utilizan al mismo tiempo para llevar a cabo la tarea biométrica.

Normalización de puntuaciones: Proceso de reescalado de las puntuaciones de un sistema biométrico llevado a cabo para medir el error con umbral universal o para fusionar puntuaciones obtenidas por sistemas diferentes en pos de una decisión final conjunta.

Plantilla: Representación digital de las características biométricas distintivas del individuo. Se utilizan durante la autenticación biométrica como base del proceso de comparación. En inglés *template*.

Puntuación: Valor numérico obtenido a través de un algoritmo biométrico, que indica el grado de semejanza o diferencia entre una muestra biométrica y una referencia.

Rasgo biométrico: Característica biométrica del individuo a reconocer. Existen dos tipos: conductual y fisiológica.

Rasgo biométrico conductual: Característica biométrica aprendida y adquirida con el tiempo, y no basada (directamente) con características biológicas. De algún modo, todas las características biométricas dependen tanto de características comportamentales como fisiológicas. Algunos ejemplos de modalidades biométricas en las que dominan las características comportamentales son el reconocimiento de firmas y la dinámica de tecleo (keystroke).

Rasgo biométrico fisiológico: Característica biométrica basada en una característica anatómica en su mayor parte y no en un comportamiento aprendido. Ejemplos de modalidades biométricas en las que pueden dominar las características fisiológicas son la geometría de las huellas dactilares y la geometría de las manos.

Reconocimiento: Término utilizado en la descripción de sistemas biométricos (por ejemplo, reconocimiento de firma o reconocimiento de iris) relativo a su principal función. El término 'reconocimiento' no implica necesariamente verificación ni identificación (cerrada o abierta).

Referencia: Datos biométricos de un individuo almacenados para ser usados en un reconocimiento posterior. Una referencia puede estar formada por una o más plantillas o modelos.

Registro: Proceso de adquisición de la muestra biométrica del cliente, conversión en referencia biométrica y posterior almacenamiento en la base de datos del sistema para su ulterior comparación.

Rendimiento: Término general utilizado para describir la eficacia de reconocimiento de un algoritmo o sistema biométrico. En inglés *performance*.

Score: Ver **puntuación**.

Sensor: Hardware en un dispositivo biométrico que convierte los datos biométricos de entrada en una señal digital y transmite al dispositivo de procesamiento.

Suplantación: Habilidad para engañar a un sensor biométrico y hacer que éste reconozca a un usuario ilegítimo como legítimo (verificación) o que no identifique a una persona que es parte de la base de datos.

Tarea biométrica: Tipos de operación que pueden realizarse con un sistema biométrico: verificación, identificación en conjunto abierto e identificación en conjunto cerrado.

Sistema biométrico: Un sistema biométrico es un sistema automatizado capaz de: 1) capturar una muestra biométrica del usuario, 2) extraer y procesar los datos biométricos de dicha muestra, 3) almacenar la información extraída en una base de datos, 4) comparar los datos biométricos con los existentes en una o más referencias y 5) decidir el grado de similitud e indicar si se ha logrado una identificación o verificación de identidad.

Umbral de decisión: Valor predeterminado asociado a un usuario para las tareas de verificación o identificación de grupo abierto en los sistemas biométricos. La aceptación o el rechazo de los datos biométricos depende de si el resultado de coincidencia se encuentra por encima o por debajo de este umbral.

Umbral universal: Umbral de decisión establecido de forma global para todo el sistema biométrico.

Umbral individual: Umbral de decisión establecido de forma particularizada para cada usuario.

Usuario: Individuo que interacciona con el sistema biométrico, para inscribirse, ser verificado o identificado.

Validación cruzada: Método estadístico consistente en dividir una muestra de datos en subconjuntos de modo que el análisis es inicialmente realizado sobre uno de ellos mientras que los otros subconjuntos se dedican para su validación.

Verificación: Tarea durante la cual el sistema biométrico intenta confirmar la identidad declarada de un individuo, al comparar la muestra suministrada con una o más plantillas almacenadas con anterioridad.

Vulnerabilidad: Riesgo de que el funcionamiento de un sistema biométrico pueda verse comprometido por actividad fraudulenta, defecto de diseño, accidentes, errores de hardware o condiciones ambientales externas.

B

Herramientas desarrolladas

Herramientas de apoyo desarrolladas

Para llevar a cabo el trabajo desarrollado en esta tesis se han desarrollado una serie de herramientas de apoyo a la investigación en el lenguaje de programación las cuales pasamos a enumerar junto a algunas de sus características más relevantes.

1. **Visor de firmas dinámicas**. Permite analizar las firmas dinámicas mediante distintas funcionalidades tales como zoom, visualización individual y en conjunto, a distintas velocidades, visualización de los puntos de la firma, zonas sin tinta, etc. Ver figuras B.1 y B.2.

2. **API Java envoltorio de programas de HTK (Young & et al, 2002)** (Hidden Markov Model Toolkit[1]). Incluye clases para automatizar las fases de creación del modelo, fase de prueba y evaluación del resultado de sistemas de VAFD basados en modelos ocultos de Markov.

3. **Marco para la automatización de experimentos de sistemas basados en firma dinámica**. Se han diseñado una colección de clases de utilidad para realizar experimentos de forma modular que incluye las siguientes funcionalidades:

   - Generador de experimentos: fases de entrenamiento, prueba y evaluación.
   - Implementación del algoritmo DTW.
   - Cálculo del error: EER, FAR, FRR.
   - Procesamiento de ficheros de puntuaciones: para experimentar con normalización de puntuaciones.
   - Analizador de ficheros de firma en múltiples formatos: FPG, ASCII, XML.
   - Extracción de características: parámetros y funciones.
   - Utilidades para normalización geométrica y estadística.
   - Generación dinámica de gráficas: basado en la librería jCharts[2].

[1] Más información en `http://htk.eng.cam.ac.uk/`
[2] `http://jcharts.sourceforge.net/`

4. **Módulos de FirmWeb**: se realizaron distintas utilidades tanto a nivel de cliente como de servidor:

   - Cliente: Applet de captura de firma dinámica adaptable al hardware disponible (figura B.3). Applet de visualización de firmas (figura B.4).
   - Servidor: Módulo de almacenamiento y verificación de firmas. Establecimiento del umbral de decisión dependiente del usuario.

Métricas

Las herramientas anteriores han sido desarrolladas en el lenguaje de programación Java. La tabla B.1 muestra algunas métricas de software que caracterizan el tamaño del software desarrollado.

Tabla B.1: Algunas métricas del software desarrollado

Métrica	Valor
Número de paquetes	44
Número de clases	175
Número de métodos	1391
Nº de líneas de código fuente	19825

Como medida de ejemplo de rendimiento incluimos en la tabla B.2 los tiempos en las fases de extracción de características, creación del modelo y comparación del sistema enviado a BSEC 2009. Los datos corresponden a los tiempos obtenidos con el conjunto de desarrollo (50 usuarios) en promedio sobre todas las tareas. Las pruebas fueron realizadas en un equipo con sistema operativo Linux, con 4 procesadores de 4 núcleos a 2.6 GHz y con 16 GB de memoria RAM.

Tabla B.2: Métricas de rendimiento del sistema participante en BSEC 2009

Fase	TT (sg)	NLP	Valores medios
Extracción de características	290	2500	TPF: 0.12 s
Creación del modelo	157	50	TPU: 3.14 s
Comparación (firmas del cliente)	390	1000	TPF: 0.39 s
Comparación (falsificaciones casuales)	465	1500	TPF: 0.31 s
Comparación (imitaciones)	510	1000	TPF: 0.51 s
Total fase de comparación	1365	3500	TPF: 0.39 s

[1]TT: Tiempo Total
[2]NLP: Número de llamadas al programa
[3]TPF: Tiempo por Firma
[4]TPU: Tiempo por Usuario

Capturas de pantalla

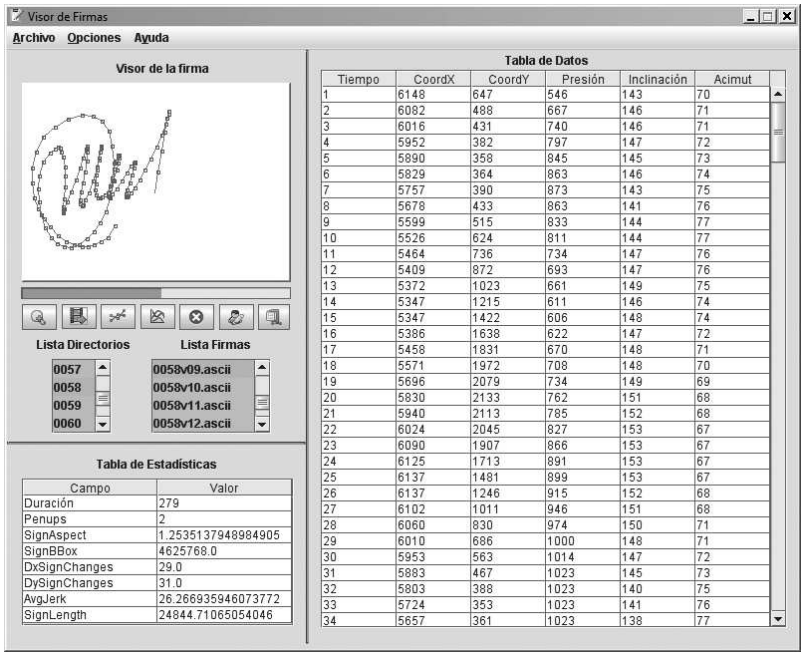


Figura B.1: Visor de firmas: ventana principal

FIGURA B.2: Visor de firmas: ventana de animación conjunta

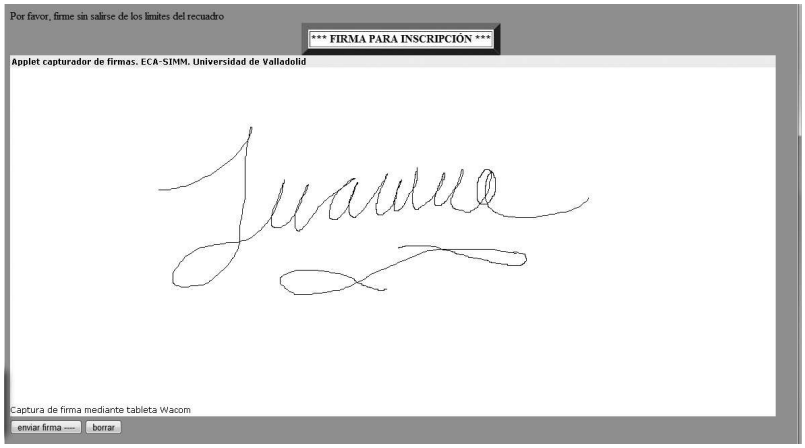


FIGURA B.3: FirmWeb: applet de captura de firma

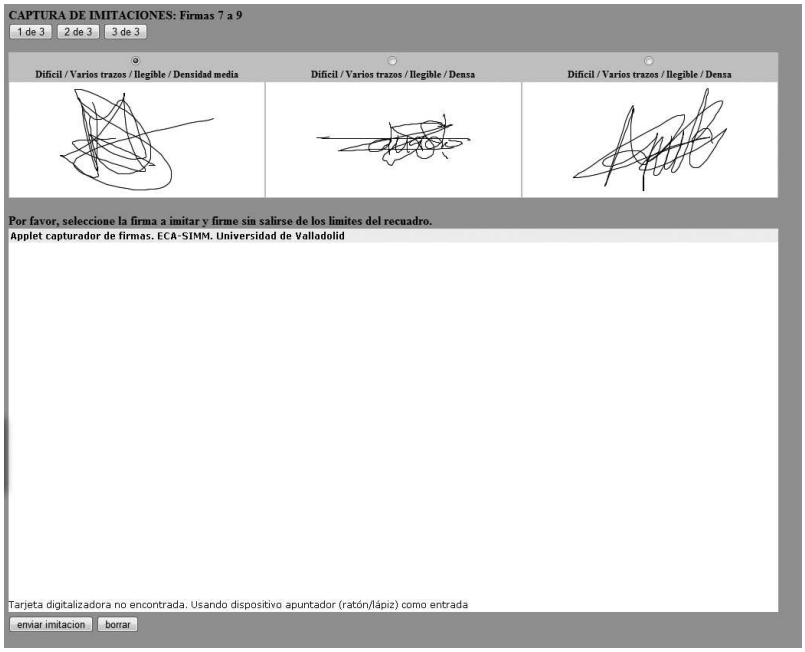


FIGURA B.4: FirmWeb: applet de visualización de firma

Índice alfabético

Referencias

Ballard, L., Lopresti, D. & Monrose, F. (2006). Evaluating the security of handwriting biometrics. In *10th International Workshop on Frontiers in Handwriting Recognition (IWFHR06)*, 461–466, La Baule, France. 85

Brault, J.J. & Plamondon, R. (1993). A complexity measure of handwritten curves: modeling of dynamic signature forgery. *Systems, Man and Cybernetics, IEEE Transactions on*, **23**, 400–413. 110

Cappé, O. (2001). Ten years of hmms. `http://www.tsi.enst.fr/~cappe/docs/hmmbib.html`. 29

Cappelli, M.M., Maio, D., Maltoni, D., Wayman, J.L. & Jain, A.K. (2004). FVC2004: Third fingerprint verification competition. In *in Proceedings of the First International Conference on Biometric Authentication*, 1–7. 24

Cappelli, R., Maio, D., Maltoni, D., L.Wayman, J. & Jain, A.K. (2006). Performance evaluation of fingerprint verification systems. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **28**, 3–18. 2

Cappelli, R., Ferrara, M., Franco, A. & Maltoni, D. (2007). Fingerprint verification competition 2006. *Biometric Technology Today*, **15**, 7–9. 24

Cover, T. (1974). The best two independent measurements are not the two best. *IEEE Trans. Systems, Man, and Cybernetics*, **4**, 116–117. 66

Cyber-SIGN (2007). `http://www.cybersign.com/`. 3

DataVision (2007). `http://www.datavisionimage.com/`. 3

Devijver, P.A. & Kittler, J. (1982). *Pattern recognition: A statistical approach*. Prentice Hall. 97

Dimauro, G., Impedovo, S., Lucchese, M.G., Modugno, R. & Pirlo, G. (2004). Recent advancements in automatic signature verification. In *Proceedings of the Ninth International Workshop on Frontiers in Handwriting Recognition (IWFHR'04)*, 179–184, IEEE Computer Society, Los Alamitos, CA, USA. 3, 13

Dolfing, J., Aarts, E. & van Oosterhout, J. (1998). On-line signature verification with hidden markov models. In *Proceedings of the Fourteenth International Conference on Pattern Recognition, 1998*, vol. 2, 1309–1312. 30

Dorizzi, B., Cappelli, R., Ferrara, M., Maio, D., Maltoni, D., Houmani, N., Garcia-Salicetti, S. & Mayoue, A. (2009). Fingerprint and on-line signature verification competitions at icb 2009. In *ICB*, 725–732. 2, 3, 5, 59

Dumas, B., Pugin, C., Hennebert, J., Humm, D.P.D.A., Evequoz, F., Ingold, R. & von Rotz, D. (2005). MyIDea - Multimodal biometrics database, description of acquisition-protocols. In *Proceedings of Third COST 275 Workshop (COST 275)*, 59–62, Hatfield (UK). 16, 18, 65

Dynalink (2007). `http://www.dynalink.com/`. 3

Electronics, I. (2007). `http://www.interlinkelectronics.com/`. 3

Fábregas, J. & Faundez-Zanuy, M. (2008). Biometric dispersion matcher. *Pattern Recognition*, **41**, 3412–3426. 77

Faundez-Zanuy, M. (2007). On-line signature recognition based on VQ-DTW. *Pattern Recognition*, **40**, 981–992. 9, 41, 61, 111

Faundez-Zanuy, M. & Pascual-Gaspar, J. (2009). Efficient On-line signature recognition based on Multi-section VQ, submitted to Pattern Analysis and Applications (PAA). 111

Fierrez, J. & Ortega-Garcia, J. (2007). *On-line signature verification. Handbook of Biometrics*. 3, 9, 59, 62, 63

Fierrez, J., Ortega-Garcia, J., Ramos, D. & Gonzalez-Rodriguez, J. (2007). HMM-based on-line signature verification: feature extraction and signature modeling. *Pattern Recognition Letters*, **28**, 2325–2334. 5, 32, 77

Fierrez, J., Galbally, J., Ortega-Garcia, J., Freire, M.R., Alonso-Fernandez, F., Ramos, D., Toledano, D.T., Gonzalez-Rodriguez, J., Siguenza, J.A., Garrido-Salas, J., Anguiano, E., de Rivera, G.G., Ribalda, R., Faundez-Zanuy, M., Ortega, J.A., Cardeñoso-Payo, V., Viloria, A., Vivaracho, C.E., Moro, Q.I., Igarza, J.J., Sanchez, J., Hernaez, I., Orrite-Uruñuela, C., Martinez-Contreras, F. & Gracia-Roche, J.J. (2009). Biosecurid: A multimodal biometric database. *Pattern Analysis and Applications*. 28

Fierrez-Aguilar, J. (2006). *Adapted Fusion Schemes for Multimodal Biometric Authentication*. Ph.D. thesis, Escuela Técnica Superior de Ingenieros de Telecomunicación. Dept. de Señales, Sistemas y Radiocomunicaciones. Universidad Politécnica de Madrid, Madrid. 13, 48

Fierrez-Aguilar, J., Krawczyk, S., Ortega-Garcia, J. & Jain, A.K. (2005a). Fusion of local and regional approaches for on-line signature verification. In *Advances in Biometric Person Authentication*, vol. 3781/2005 of *Lecture Notes in Computer Science*, 188–196, Springer Berlin / Heidelberg. 13, 32, 77, 108

Fierrez-Aguilar, J., Nanni, L., Lopez-Peñalba, J., Ortega-Garcia, J. & Maltoni, D. (2005b). An on-line signature verification system based on fusion of local and global information. In *Audio- and Video-Based Biometric Person Authentication*, vol. 3546/2005 of *Lecture Notes in Computer Science*, 523–532, Springer Berlin / Heidelberg. 32, 66, 77

Fierrez-Aguilar, J., Ortega-Garcia, J. & Gonzalez-Rodriguez, J. (2005c). Target dependent score normalization techniques and their application to signature verification. *IEEE Transactions on Systems, Man and Cybernetics, Part C*, **35**, 418–425. 32, 51, 66, 96

Fuentes, M., Garcia-Salicetti, S. & Dorizzi, B. (2002). On-line signature verification: Fusion of a hidden markov model and a neural network via a support vector machine. In *Proceedings of the Eighth International Workshop on Frontiers in Handwriting Recognition (IWFHR'02)*, 253–258, IEEE Computer Society, Washington, DC, USA. 31, 32, 57

Garcia-Salicetti, S., Beumier, C., Chollet, G., Dorizzi, B., les Jardins, J.L., Lunter, J., Ni, Y. & Petrovska-Delacretaz, D. (2003). BIOMET: A multimodal person authentication database including face, voice, fingerprint, hand and signature modalities. In J. Kittler & M. Nixon, eds., *AVBPA*, LNCS 2688, 845–853, Springer-Verlag Berlin Heidelberg. 16, 18, 65

Garcia-Salicetti, S., Fierrez-Aguilar, J., Alonso-Fernandez, F., Vielhauer, C., Guest, R., Allano, L., Trung, T.D., Scheidat, T., Van, B.L., Dittmann, J., Dorizzi, B., Ortega-Garcia, J., Gonzalez-Rodriguez, J., di Castiglione, M.B. & Fairhurst, M. (2007). Biosecure reference systems for on-line signature verification: A study of complementarity. *Annals of Telecommunications, Special Issue on Multimodal Biometrics*, **62**, 36–61. 32, 77

Garcia-Salicetti, S., Houmani, N. & Dorizzi, B. (2009). A novel criterion for writer enrolment based on a time-normalized signature sample entropy measure. *EURASIP Journal on Advances in Signal Processing*, **2009**. 86

Garner, S. (1995). Weka: The waikato environment for knowledge analysis. In *New Zealand Computer Science Research Students Conference*, 57–64. 98

Genius (2007). `http://www.geniusnet.com/`. 3

Griess, F.D. (2000). On-line signature verification (master's project report). Tech. Rep. MSU-CSE-00-15, Department of Computer Science, Michigan State University, East Lansing, Michigan. 62

Gupta, G.K. & McCabe, A. (1997). A review of dynamic handwritten signature verification, very cited. 3

Guyon, I. & Elisseeff, A. (2003). An introduction to variable and feature selection. *Journal of Machine Learning Research*, **3**, 1157–1182. 68

Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P. & Witten, I.H. (2009). The weka data mining software: An update. In *SIGKDD Explorations*, vol. 11. 98

Hennebert, J., Humm, A. & Ingold, R. (2007). Modelling spoken signatures with gaussian mixture model adaptation. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 07)*. 77

Hu, J., Brown, M.K. & Turin, W. (1996). Hmm based on-line handwriting recognition. *IEEE Trans. Pattern Anal. Mach. Intell.*, **18**, 1039–1045. 13

Humm, A., Hennebert, J. & Ingold, R. (2007). Spoken handwriting verification using statistical models. In *Accepted for publication, International Conference on Document Analysis and Recognition (ICDAR 07), Curitiba Brazil*. 77

Igarza, J.J., Goirizelaia, I., Espinosa, K., Hernaez, I., Mendez, R. & Sanchez, J. (2003). Online handwritten signature verification using hidden markov models. In *CIARP*, 391–399. 31

Igarza, J.J., Gómez, L., Hernáez, I. & Goirizelaia, I. (2004). Searching for an optimal reference system for on-line signature verification based on (x, y) alignment. In D. Zhang & A. Jain, eds., *ICBA04*, LNCS 3072, 519–525, Springer-Verlag Berlin Heidelberg. 31

Impedovo, S. & Pirlo, G. (2007). Verification of handwritten signatures: an overview. In *ICIAP '07: Proceedings of the 14th International Conference on Image Analysis and Processing (ICIAP 2007)*, 191–196, IEEE Computer Society, Washington, DC, USA. 3

International Biometric Group (2007). http://www.biometricgroup.com. 3, 107

Jain, A.K., Duin, R.P.W. & Mao, J. (2000). Statistical pattern recognition: A review. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **22**, 4–37. 66

Jain, A.K., Griess, F.D. & Connell, S.D. (2002). On-line signature verification. *Pattern Recognition*, **35**, 2963–2972. 62

Jain, A.K., Pankanti, S., Prabhakar, S., Hong, L., Ross, A. & Wayman, J.L. (2004a). Biometrics: A grand challenge. *Proc. International Conference on Pattern Recognition (ICPR)*. 1

Jain, A.K., Ross, A. & Prabhakar, S. (2004b). An introduction to biometric recognition. *IEEE transactions on circuits and systems for video technology*, **14**. 60

Jain, A.K., Bolle, R.M. & Pankanti, S. (2005). *Biometrics: Personal Identification in Networked Society*. Springer. 61

JRBP (2008). IV Jornadas de Reconocimiento Biométrico de Personas. `http://eca-simm.infor.uva.es/jrbp08/`. 79

Kajarekar, S.S., Scheffer, N., Graciarena, M., Shriberg, E., Stolcke, A., Ferrer, L. & Bocklet, T. (2009). The sri nist 2008 speaker recognition evaluation system. In *Proceedings of the 2009 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 4205–4208, IEEE Computer Society, Washington, DC, USA.

Kalenova, D. (2003). Personal authentication using signature recognition. `http://www.it.lut.fi/kurssit/03-04/010970000/seminars/Kalenova.pdf`, Unpublished. 3

Kashi, R.S., Hu, J., Nelson, W.L. & Turin, W. (1997). On-line handwritten signature verification using hidden markov model features. In *Proceedings of the Fourth International Conference on Document Analysis and Recognition*, vol. 1, 253–257, Proceedings of the Fourth International conference, Ulm. 30

Kholmatov, A. & Yanikoglu, B. (2005). Identity authentication using improved online signature verification method. *Pattern Recognition Letters*, **26**, 2400–2408. 5, 13, 51, 61, 62

Leclerc, F. & Plamondon, R. (1994). Automatic signature verification: The state of the art 1989-1993. *International Journal of Pattern Recognition and Artificial Intelligence*, **8**, 643–660. 3, 12

Lei, H. & Govindaraju, V. (2005). A comparative study on the consistency of features in on-line signature verification. *Pattern Recognition Letters*, **26**, 2483–2489. 66

Ly, V.B., Garcia-Salicetti, S. & Dorizzi, B. (2007). On using the viterbi path along with hmm likelihood information for online signature verification. *IEEE Transactions on Systems, Man, and Cybernetics, Part B*, **37**, 1237–1247. 32, 77

Maio, D., Maltoni, D., Wayman, J.L. & Jain, A.K. (2006). Performance evaluation of fingerprint verification systems. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **28**, 3–18. 24

Martens, R. & Claesen, L. (1996). On-line signature verification by dynamic time-warping. In *Proceedings of the 13th International Conference on Pattern Recognition*, vol. 3, 38–42. 62

Martin, A., Doddington, G., Kamm, T., Ordowski, M. & Przybocki, M. (1997). The det curve in assessment of detection task performance. 1895–1898. 11

Martinez-Diaz, M., Fierrez, J. & Ortega-Garcia, J. (2008). Incorporating signature verification on handheld devices with user-dependent hidden markov models. In *Proc. International Conference on Frontiers in Handwriting Recognition, ICFHR*. 32

Moneo-Agapito, J.D. (2005). *Aplicacion de los Modelos Ocultos de Markov a la verificacion biométrica basada en la firma en linea*. Master's thesis, Universidad de Valladolid, Proyecto Fin de Carrera. Departamento de Informatica. IX, XI, 48, 50

Muramatsu, D. & Matsumoto, T. (2003). An HMM online signature verifier incorporating signature trajectories. In *Proc. of the Seventh International Conference on Document Analysis and Recognition (ICDAR 2003)*, vol. 1, 438, IEEE, Edinburgh. 31

Nalwa, V.S. (1997). Automatic on-line signature verification. *Proceedings of the IEEE*, **85**, 215–239. 3

Nelson, W., Turin, W. & Hastie, T. (1994). Statistical methods for on-line signature verification. *IJPRAI*, **8**, 749–770. 30

Ortega-Garcia, J., Fierrez-Aguilar, J., Martin-Relloand, J. & Gonzalez-Rodriguez, J. (2003a). Complete signal modeling and score normalization for function-based dynamic signature verification. In *Proc. of the Audio- and Video-Based Biometric Person Authentication (AVBPA 2003)*, 658–667. 62

Ortega-Garcia, J., Fierrez-Aguilar, J., Simon, D., Gonzaleza, J., Faundez, M., Espinosa, V., Satue, A., Hernaez, I., Igarza, J.J., Vivaracho, C., Escudero, D. & Moro, Q.I. (2003b). MCYT baseline corpus: A bimodal biometric database. *IEE Proceedings Vision, Image and Signal Processing, Special Issue on Biometrics on the Internet*, **150**, 395–401. 16, 35, 65

Ortega-Garcia, J., Fierrez, J., Alonso-Fernandez, F., Galbally, J., Freire, M., Gonzalez-Rodriguez, J., Garcia-Mateo, C., Alba-Castro, J.L., Gonzalez-Agulla, E., Otero-Muras, E., Garcia-Salicetti, S., Allano, L., Ly-Van, B., Dorizzi, B., Kittler, J., Bourlai, T., Poh, N., Deravi, F., Ng, M., Fairhurst, M., Hennebert, J., Humm, A., Tistarelli, M., Brodo, L., Richiardi, J., Drygajlo, A., Ganster, H., Sukno, F., Pavani, S.K., Frangi, A., Akarun, L. & Savran, A. (2009). The multi-scenario multi-environment biosecure multimodal database (bmdb). *IEEE Trans. on Pattern Analysis and Machine Intelligence*, to appear. 28, 86

P. J. Phillips, P.J.R., H. Moon & Rizvi, S. (2006). The FERET evaluation methodology for face recognition algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **22**, 1090–1104. 2

Pascual-Gaspar, J. & Cardenoso-Payo, V. (2007). On-line signature verification using hidden markov models with number of states estimation from the signature duration. In *Biometrics Symposium, 2007*. 29, 77

Pascual-Gaspar, J.M. & Cardeñoso-Payo, V. (2006). Verificación automática de firma on-line mediante HMMs con estructura dependiente de usuario. In *III Jornadas de Reconocimiento Biométrico de Personas*, 43–54, Sevilla, Spain. 29

Pascual-Gaspar, J.M. & Cardeñoso-Payo, V. (2007a). Automatic online signature verification using HMMs with user-dependent structure. In S.W. Lee & S.Z. Li, eds., *ICB*, vol. 4642 of *Lecture Notes in Computer Science*, 1057–1066, Springer. 29

Pascual-Gaspar, J.M., Cardeñoso-Payo, V. & Vivaracho-Pascual, C.E. (2008). A comparative study of local feature sets in position, velocity and acceleration domains for on-line signature verification. In *IV Jornadas de Reconocimiento Biométrico de Personas*, 35–42, Valladolid, Spain. 59

Pascual-Gaspar, J.M., Cardeñoso-Payo, V. & Vivaracho-Pascual, C.E. (2009). Practical on-line signature verification. In *ICB*, 1180–1189. 59

PenOp (2007). http://www.penop.com/. 3

Phillips, P.J. (2006a). Face and iris evaluations at NIST. CardTech/SecurTech. 2

Phillips, P.J. (2006b). FRGC and ICE workshop. `http://iris.nist.gov/ICE/ICE_2005_Results_30March2006.pdf`. 24

Phillips, P.J., Moon, H., Rizvi, S.A. & Rauss, P.J. (2000). The feret evaluation methodology for face-recognition algorithms. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, **22**, 1090–1104. 24

Phillips, P.J., Scruggs, W.T., O'toole, A.J., Flynn, P.J., Kevin, W., Schott, C.L. & Sharpe, M. (2007). FRVT 2006 and ICE 2006 large-scale results. Tech. rep., National Institute of Standards and Technology. 24

Plamondon, R. & Lorette, G. (1989). Automatic signature verification and writer identification: The state of the art. *Pattern Recognition*, **22**, 107–131. 3, 11

Plamondon, R. & Parizeau, M. (1988). Signature verification from position, velocity and acceleration signals: a comparative study. In *Proc. of the 9th International Conference on Pattern Recognition*, vol. I, 260–265. 66

Plamondon, R. & Srihari, S.N. (2000). On-line and off-line handwriting recognition: A comprehensive survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **22**, 63–84. 3

Przybocki, M. & Martin, A. (2004). NIST speaker recognition evaluation chronicles. In *Odyssey: The Speaker and Language Recognition Workshop*, 12–22, Toledo, Spain. 2, 24

Przybocki, M.A., Martin, A.F. & Le, A.N. (2006). NIST speaker recognition evaluation chronicles - part 2. In *Speaker and Language Recognition Workshop, IEEE Odyssey*, 1–6. 2, 24

Rabiner, L.R. (1989). A tutorial on hidden markov models and selected application in speech recognition. *Proceedings of IEEE*, **77**, 257–286. 13, 30

Rabiner, L.R., Rosenberg, A.E. & Levinson, S.E. (1978). Considerations in dynamic time warping algorithms for discrete word recognition. *IEEE Transactions on Acoustics, Speech and Signal Processing*, **ASSP-26**, 575–582. 14, 59

Rigoll, G. & Kosmala, A. (1998). A systematic comparison between on-line and off-line methods for signature verification with hidden markov models. In *Proceedings of the Fourteenth International Conference on Pattern Recognition*, vol. 2, 1755–1757. 31

Ross, A.A., Nandakumar, K. & Jain, A.K. (2006). *Handbook of Multibiometrics (International Series on Biometrics)*. Springer-Verlag New York, Inc., Secaucus, NJ, USA. 11, 97

Salicetti, S., Houmani, N. & Dorizzi, B. (2008). A client-entropy measure for on-line signatures. In *Biometrics Symposium, 2008. BSYM '08*, 83–88. IX, 86, 87

Shafiei, M.M. & Rabiee, H.R. (2003). A new on-line signature verification algorithm using variable length segmentation and hidden markov models. In *Proceedings of the Seventh International Conference on Document Analysis and Recognition (ICDAR 2003)*, vol. 1, 443, IEEE Computer Society, Los Alamitos, CA, USA, poster. 31, 57

Tapiador Mateos, M. & Sigüenza Pizarro, J.A. (2005). *Tecnologías biométricas aplicadas a la seguridad*. RA-MA, 1st edn. 12

Titterington, D., Smith, A. & Makov, U. (1985). *Statistical Analysis of Finite Mixture Distributions*. John Wiley and Sons. 13

Vivaracho-Pascual, C., Faundez-Zanuy, M. & Pascual, J.M. (2009). An efficient low cost approach for on-line signature recognition based on length normalization and fractional distances. *Pattern Recognition*, **42**, 183–193. 61, 111

Wacom (2009). `http://global.wacom.com/`. 3, 18

Wessels, T. & Omlin, C. (2000). A hybrid system for signature verification. In *Neural Networks, 2000. IJCNN 2000, Proceedings of the IEEE-INNS-ENNS International Joint Conference on*, vol. 5, 509–514vol.5. 31

Yang, L., Widjaja, B.K. & Prasad, R. (1995). Application of hidden markov models for signature verification. *Pattern Recognition*, **28**, 161–170. 13, 30, 31

Yanikoglu, B. & Kholmatov, A. (2003). An improved decision criterion for genuine/forgery classification in on-line signature verification. In *Proc. of the 2003 Int. Conference on Artificial Neural Networks (ICANN 2003)*, ICANN, Istanbul, Turkey, berrin won the first place at the 1st International Signature Verification Competition, SVC 2004. 13

Yeung, D., Chang, H., Xiong, Y., George, S., Kashi, R., Matsumoto, T. & Rigoll, G. (2004). SVC2004: First international signature verification competition. In *Proc. of the First International Conference on Biometrics Authentication (ICBA 2004)*, 16–22. 2, 3, 5, 13, 16, 32, 59, 62, 65, 77

Yoon, H.S., Lee, J. & Yang, H. (2002). An online signature verification system using hidden markov model in polar space. In *Frontiers in Handwriting Recognition, 2002. Proceedings. Eighth International Workshop on*, 329–333. 31

Young, S. & et al, G.E. (2002). *The HTK Book (for HTK Version 3.2.1)*. Cambridge University Engineering Department. 123