

UNIVERSIDAD CARLOS III DE MADRID

ESCUELA POLITÉCNICA SUPERIOR

Departamento Tecnología Electrónica

Grado en Ingeniería de Sistemas Audiovisuales



TRABAJO FIN DE GRADO

**ANÁLISIS DEL ALGORITMO DTW PARA
RECONOCIMIENTO BIOMÉTRICO DE PERSONAS
MEDIANTE FIRMA MANUSCRITA ON-LINE**

Autor: Patricia de los Ángeles Alonso Fuertes

Tutor: Luís Mengíbar Pozo

Septiembre 2012

Agradecimientos

"Aprendí que no se puede dar marcha atrás, que la esencia de la vida es ir hacia adelante. La vida, en realidad, es una calle de sentido único".

Agatha Christie

Con este proyecto finaliza una etapa de mi vida muy importante, que me ha marcado y definirá en gran parte la persona que seré en unos años. Por este punto y seguido quiero hacer una parada para dar las gracias a todas esas personas que me han ayudado a forjarme como persona y como profesional de las telecomunicaciones.

En primer lugar quiero agradecer a mi tutor, Luís Mengíbar, la oportunidad que me brindó al asignarme este proyecto y el dirigirme y asesorarme en la realización del mismo. Ha sido un gran apoyo y le estoy muy agradecida por el seguimiento que ha hecho de mi proyecto, por los consejos que me ha dado y por atenderme siempre que lo he necesitado.

Siguiendo en la línea estudiantil, quiero resaltar la satisfacción que siento ahora de haber estudiado la carrera en la Universidad Carlos III, por su prestigio y dureza. Aunque haya habido momentos duros, siento que he salido fortalecida de ellos.

En los años universitarios se conoce mucha gente y entre ella yo he podido sacar dos grandes amistades, mis compañeras y amigas Ana y Marian. Hemos compartido todos estos años, hemos estudiado juntas y hemos desesperado ante las prácticas, pero siempre lo hemos sacado a delante. Gracias chicas por vuestro apoyo, paciencia y dedicación.

En cuanto a mi familia, quiero agradecer en primer lugar los valores que me han inculcado y el haberme dado la oportunidad de realizar mis estudios. Ellos me han hecho como soy y sin ellos nunca habría terminado esta carrera. Me han apoyado y ayudado cada uno como ha podido y yo eso lo valoro de verdad. Gracias por toda la comprensión, el apoyo y la confianza que tenéis en mí.

Por último, quiero dar las gracias a la persona que más cree en mí, la persona que más me conoce y me entiende, mi novio Víctor. Él ha sido mi pilar todos estos años, ambos nos hemos ayudado mucho y somos lo que somos gracias al apoyo mutuo. Él hace que

quiera ser mejor, que quiera llegar más lejos, que pueda levantarme cuando caigo. No hay palabras para expresar toda mi gratitud, pero él lo sabe.

En fin, me siento orgullosa de haber llegado hasta aquí y espero que en estos tiempos difíciles haya un hueco para mí en el mundo de las telecomunicaciones.

Índice

Capítulo 1: Introducción y objetivos	11
1.1.- Introducción	12
1.2.- Objetivos.....	12
Capítulo 2: Introducción a la biometría	14
2.1.- Definición de la biometría	15
2.2.- Historia de la biometría	16
2.3.- Sistemas biométricos: Identificación y verificación	17
2.3.1.- Sistema de identificación.....	17
2.3.2.- Sistema de verificación	18
2.4.- Indicadores biométricos actuales.....	19
2.5.- Principales sistemas.....	19
2.5.1.- Huella dactilar	19
2.5.2.- Patrones oculares	21
2.5.3.- Voz	22
2.5.4.- Rostro	23
2.5.5.- Geometría de la mano	23
2.5.6.- Andadura	24
2.5.7.- Dinámica de teclado	24
2.5.8.- Firma manuscrita	25
2.6.- Comparativa de los principales sistemas	25
Capítulo 3: Introducción a la firma manuscrita	28
3.1.- Historia de la firma	29
3.2.- Firma: definición, características y elementos	29
3.3.- Sistemas de reconocimiento de firmas	30
3.4.- Adquisición de la firma y dispositivos de captura	31
3.4.1.- Adquisición Off-line	31
3.4.2.- Adquisición On-line	31

3.5.- Parámetros de estudio en una firma	32
3.6- Marco legal y tecnológico	33
Capítulo 4: Algoritmos de reconocimiento de firmas: DTW	35
4.1.- Alineamiento Temporal Dinámico – DTW	36
Capítulo 5: Diseño y desarrollo de la aplicación	42
5.1- Introducción	43
5.2- Base de datos MCyT	43
5.3- Primera parte: Algoritmo DTW básico	44
5.4- Segunda parte: Algoritmo DTW modificado	51
Capítulo 6: Análisis del algoritmo y pruebas	52
6.1- Introducción	53
6.2- Escenario de los experimentos	54
6.3- Sistemas de evaluación	54
6.4- Experimentos y pruebas	57
6.4.1- 1er Conjunto de experimentos: Tipos de restricciones locales.....	57
6.4.2- 2º Conjunto de experimentos: Tipos de distancias	68
6.4.3- 3er Conjunto de experimentos: Tipos de normalizaciones	75
Capítulo 7: Conclusiones y trabajos futuros	79
7.1- Introducción	80
7.2- Conclusiones	80
7.3- Trabajos futuros	82
Anexos	84
Anexo 1: Planificación	85
Anexo 1.1: Diagrama EDT	85
Anexo 1.2: Diagrama PERT	86
Anexo 1.3: Panificación de tareas y diagrama GANTT	88
Anexo 2: Presupuesto	90
Anexo 2.1: Personal	90
Anexo 2.2: Equipos	90
Anexo 2.3: Subcontratación de tareas	90

Anexo 2.4: Otros costes directos	90
Anexo 2.5: Costes indirectos	90
Anexo 2.6: Presupuesto desglosado del proyecto	91
Bibliografía	94

Índice de figuras

Capítulo 2: Introducción a la biometría	14
Figura 2.1: Imagen ilustrativa de rasgos biométricos	15
Figura 2.2: Esquema de un sistema biométrico genérico.....	17
Figura 2.3: Gráfica FAR-FRR	18
Figura 2.4: Esquema características biométricas	19
Figura 2.5: Características de huellas dactilares	20
Figura 2.6: Proceso de escaneo de la huella dactilar	20
Figura 2.7: Partes de la retina	21
Figura 2.8: Reconocimiento del iris	22
Figura 2.9: Reconocimiento de voz	22
Figura 2.10: Rastreo de rostro	23
Figura 2.11: Posicionamiento geometría de la mano	24
Figura 2.12: Parámetros de la forma y del movimiento	24
Figura 2.13: Dinámica de teclado	24
Figura 2.14: Captura de firma	25
Figura 2.15: Grado de implantación sistemas biométricos	27
Capítulo 3: Introducción a la firma manuscrita	28
Figura 3.1: Adquisición off-line	31
Figura 3.2: Tableta digital.....	32
Capítulo 4: Algoritmos de reconocimiento de firmas: DTW	35
Figura 4.1: Banda de Sakoe-Chiba y Paralelogramo de Itakura.....	37
Figura 4.2: Algoritmo DTW	38
Figura 4.3: P(0) Simétrico	41
Capítulo 5: Diseño y desarrollo de la aplicación	42
Figura 5.2.1: Izda.: Firma verdadera usuario 55 muestra 24; dcha.: Firma falsificada usuario 55 muestra 24	44
Figura 5.3.1: Distancia acumulada y camino óptimo	46
Figura 5.3.2: Señales originales y alineadas	46

Figura 5.3.3: Variables de estudio: coordenada x e y, y velocidades	47
Figura 5.3.4: Diagrama de flujos funciones DTW básico (parte 1, parte 2 y parte 3) .	48
Capítulo 6: Análisis del algoritmo y pruebas	52
Figura 6.2: Escenario de los experimentos	54
Figura 6.3.1: Gráfica FAR vs FRR	55
Figura 6.3.2: Curva ROC	56
Figura 6.3.3: Curva DET	56
Figura 6.4.1.1: Curva DET y ROC para P(0) Simétrico	59
Figura 6.4.1.2: Curva DET y ROC para P(0) Asimétrico	60
Figura 6.4.1.3: Curva DET y ROC para P(1/2) Simétrico	61
Figura 6.4.1.4: Curva DET y ROC para P(1/2) Asimétrico	63
Figura 6.4.1.5: Curva DET y ROC para P(1) Simétrico	64
Figura 6.4.1.6: Curva DET y ROC para P(1) Asimétrico	65
Figura 6.4.1.7: Curva DET y ROC para P(2) Simétrico	66
Figura 6.4.1.8: Curva DET y ROC para P(2) Asimétrico	68
Figura 6.4.2.1: Curva DET y ROC para P(0) Simétrico y L_1	71
Figura 6.4.2.2: Curva DET y ROC para P(1/2) Asimétrico y L_1	72
Figura 6.4.2.3: Curva DET y ROC para P(0) simétrico y $L_{1/2}$	73
Figura 6.4.2.4: Curva DET y ROC para P(1/2) Asimétrico y $L_{1/2}$	74
Figura 6.4.3.1: Curva DET y ROC para P(0) simétrico, L_1 y normalización de máximo y mínimo	77
Figura 6.4.3.2: Curva DET y ROC para P(1/2) Asimétrico, $L_{1/2}$ y normalización de máximo y mínimo	78
Anexos	84
Figura A.1.1.1: Diagrama EDT	85
Figura A.1.2.1: Diagrama PERT	87
Figura A.1.3.1: Diagrama GANTT	89

Índice de tablas

Capítulo 2: Introducción a la biometría	14
Tabla 2.1: Tabla comparativa sistemas biométricos	26
Tabla 2.2: Ventajas e inconvenientes de los sistemas biométricos	26
Capítulo 4: Algoritmos de reconocimiento de firmas: DTW	35
Tabla 4.1: Restricciones locales (conjuntos de producciones P)	40
Capítulo 6: Análisis del algoritmo y pruebas	52
Tabla 6.4.1.1: Resultados experimento P(0) Simétrico	58
Tabla 6.4.1.2: Resultados experimento P(0) Asimétrico	59
Tabla 6.4.1.3: Resultados experimento P(1/2) Simétrico	61
Tabla 6.4.1.4: Resultados experimento P(1/2) Asimétrico	62
Tabla 6.4.1.5: Resultados experimento P(1) Simétrico	63
Tabla 6.4.1.6: Resultados experimento P(1) Asimétrico	64
Tabla 6.4.1.7: Resultados experimento P(2) Simétrico	66
Tabla 6.4.1.8: Resultados experimento P(2) Asimétrico	67
Tabla 6.4.2.1: Resultados experimento P(0) Simétrico y L_1	70
Tabla 6.4.2.2: Resultados experimento P(1/2) Asimétrico y L_1	71
Tabla 6.4.2.3: Resultados experimento P(0) Simétrico y $L_{1/2}$	72
Tabla 6.4.2.4: Resultados experimento P(1/2) Asimétrico y $L_{1/2}$	73
Tabla 6.4.3.1: Resultados experimento P(0) Simétrico, L_1 y normalización de máximo y mínimo	76
Tabla 6.4.3.2: Resultados experimento P(1/2) Asimétrico, $L_{1/2}$ y normalización de máximo y mínimo	77
Capítulo 7: Conclusiones y trabajos futuros	79
Tabla 7.2.1: Resumen del experimento con los mejores resultados en cuanto a tasa de equierror y coste computacional	81
Tabla 7.2.2: Resumen del experimento con los peores resultados en cuanto a tasa de equierror	82
Tabla 7.2.3: Resumen del experimento con los peores resultados en cuanto a coste computacional	82

Anexos	84
Tabla A.1.2.1: Datos diagrama PERT	86
Tabla A.1.3.1: Planificación tareas	88
Tabla A.2.6.1: Presupuesto	91

Capítulo 1:

Introducción y objetivos

1.1.- Introducción

Todos somos conscientes del gran avance tecnológico que nuestra sociedad ha experimentado y la tendencia cada vez mayor a un mundo totalmente informatizado.

La informatización de la sociedad es el proceso de utilización de las tecnologías de la información y la comunicación en todos los ámbitos de la vida. Siendo el objetivo lograr una eficacia y eficiencia cada vez mayor en todos los procesos.

En este contexto, se ve la importancia de la identificación de cada persona tanto por privacidad como por seguridad. Por ello la identificación biométrica está en pleno desarrollo.

Tradicionalmente, el uso de una llave o una contraseña como herramienta de seguridad era suficiente, pero en los tiempos que corren y por las tendencias futuras, se requieren métodos más seguros, eficientes, eficaces y robustos. La biometría proporciona estos aspectos en gran medida, permite identificar al usuario por características intrínsecas propias de él, además los rasgos biométricos no pueden perderse ni robarse.

Este gran avance permite evitar la suplantación de identidad, que es tan dañina para temas legales como para la privacidad.

Dentro de los métodos de verificación de identidad encontramos la firma manuscrita, que es el objeto de estudio de este proyecto. Se podría decir que es uno de los rasgos biométricos más importantes, puesto que la firma en nuestra sociedad es el principal símbolo identificativo que permite certificar el reconocimiento, la conformidad o el acuerdo de voluntades, lo cual es de vital importancia desde el punto de vista legal.

1.2.- Objetivos

El principal objetivo de este proyecto es analizar los distintos parámetros del algoritmo DTW (Dynamic Time Warping) para evaluar su funcionamiento en una aplicación de reconocimiento biométrico de firma manuscrita on-line.

Con este proyecto se quiere analizar bajo qué condiciones su funcionamiento es mejor, atendiendo sobre todo a dos características: el coste computacional y la tasa de error. De este modo se quiere conseguir un DTW lo más óptimo posible basándose en las características anteriores.

Se establecen dos escenarios de experimentación, uno **casual** y otro **seguro**, para determinar en qué situaciones es mejor un coste computacional bajo y en cuales una tasa de error baja, cuándo se priorizará uno en detrimento del otro. Se podrá ver si estos elementos son directamente proporcionales, inversamente proporcionales o si son independientes el uno del otro. Se asumirá un compromiso entre ambos valores y se obtendrán unas conclusiones, siendo estas el aporte novedoso en el estudio del DTW en el mundo de reconocimiento de la firma manuscrita.

Para conseguir este objetivo, el proyecto se puede dividir en dos subobjetivos:

- ◆ Subobjetivo I: Estudio e implementación del algoritmo DTW. En esta primera parte se estudia el algoritmo e implementa una primera aproximación a modo de prueba para llegar a entender en profundidad el algoritmo, para poder ser capaces de modificar sus parámetros más significativos.
- ◆ Subobjetivo II: Modificación de ciertas características para optimizar el reconocimiento de firmas manuscritas on-line, en función del coste computacional y la tasa de error. Se realizarán diversos experimentos y se extraerán conclusiones. El estudio se centrará en 3 conjuntos de experimentos que consistirán en variar ciertos parámetros del algoritmo, se asumirán unas características como constantes, siendo una decisión de diseño, para acotar el número de experimentos y poder obtener conclusiones.

Capítulo 2:

Introducción a la biometría

2.1.- Definición de biometría

La biometría se define como una rama de la matemática estadística que se ocupa del análisis de datos biológicos. Dicho concepto proviene de las palabras *bios* (vida) y *metron* (medida), la propia palabra indica que todo equipo biométrico mide e identifica características propias de un sujeto.

Se concluye que la biometría es un método automatizado empleado para la identificación o autenticación de individuos mediante el estudio de ciertas características físicas y/o del comportamiento de las personas. [4]

En este contexto surge otro concepto de gran importancia, el indicador biométrico, siendo éste alguna característica con la cual se puede realizar biometría.

Un indicador biométrico es alguna característica con la cual se puede realizar biometría. Cualquier indicador, debe cumplir los siguientes requerimientos y el cumplimiento o no de los mismos es determinante [10]:

- **Universalidad:** cualquier persona posee esa característica.
- **Permanencia:** la característica no varía con el tiempo.
- **Unicidad:** la probabilidad de la existencia de dos personas con una característica idéntica es muy pequeña.
- **Cuantificación:** la característica se puede medir de forma cuantitativa.

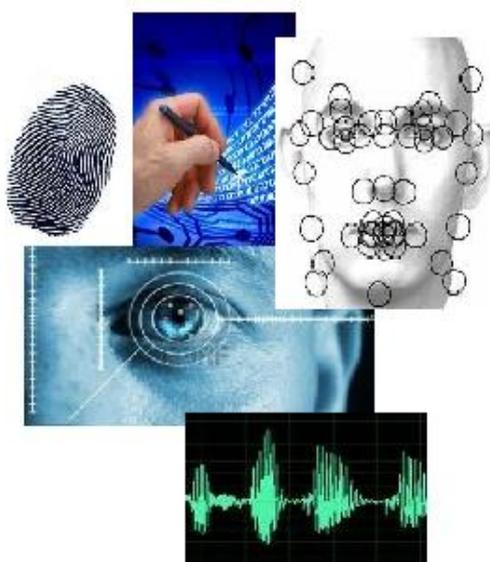


Figura 2.1: Imagen ilustrativa de rasgos biométricos.

2.2.- Historia de la biometría

La primera fecha en la que se tiene constancia del inicio de la biometría es el siglo XIV. Según un escritor y explorador, llamado Joao de Barros, fue en este siglo cuando los chinos estampaban las palmas de las manos de los niños en papel con tinta para distinguir unos de otros [12].

Pero en Occidente, la identificación consistía simplemente en la "memoria fotográfica" hasta que en 1883 el jefe del departamento fotográfico de la Policía de París (Alphonse Bertillon), desarrolló el sistema antropométrico. Éste fue el primer sistema preciso, ampliamente utilizado para identificar a criminales y convirtió a la biométrica en un campo de estudio. El método consistía en ir midiendo de forma precisa ciertas longitudes y anchuras de la cabeza y del cuerpo, así como ir registrando marcas individuales como tatuajes y cicatrices. Este sistema fue muy utilizado hasta que aparecieron defectos, problemas con métodos distintos de medidas y cambios de medida. A raíz de los trabajos de Francis Galton [33] se comenzó a usar la huella dactilar .

La identificación basada en la huella dactilar se viene utilizando desde hace más de cien años. El desarrollo de los sistemas de identificación de la huella dactilar se dio en las décadas de los sesenta y setenta cuando varias empresas desarrollaron productos para automatizar la identificación de las huellas dactilares para fines legales.

Los sistemas automáticos para la medida de otras características biométricas se desarrollaron similarmente al de la huella dactilar. La idea para usar patrones de iris como método de identificación fue propuesta en 1936 por el oftalmólogo Frank Burch pero el primer sistema para la medición de la retina fue introducido en los años ochenta. El trabajo del doctor John Daughman de la Universidad de Cambridge condujo a la primera tecnología para la medición de la retina. La identificación basada en la firma y en el rostro es relativamente nueva [14].

Un profesor suizo, Gunnar Fant, publicó en 1960 un modelo describiendo los componentes fisiológicos de la producción del discurso acústico. Durante 1964 y 1965 Woodrow Wilson Bledsoe, Helen Chan Wolf y Charles Bisson trabajaron en el uso del computador para el reconocimiento facial humano. Desarrollaron el primer sistema semi-automático de reconocimiento. La Aviación Norteamericana desarrolló el primer sistema de reconocimiento de firma en 1965 [13].

Los sistemas biométricos modernos comienzan a surgir en la segunda mitad del siglo XX junto con el desarrollo de los sistemas de computación. En los años noventa se

produce una gran explosión en este campo, lo que introduce a los sistemas biométricos en un sinfín de aplicaciones que se utilizan día a día.

2.3.- Sistemas biométricos: Identificación y Verificación

Un sistema biométrico es un sistema automatizado que realiza labores de biometría. Estos sistemas poseen tres componentes básicos [13]:

- Adquisición de algún indicador biométrico.
- Compresión, procesamiento, almacenamiento y comparación de los datos adquiridos con la base de datos.
- Interfaz.

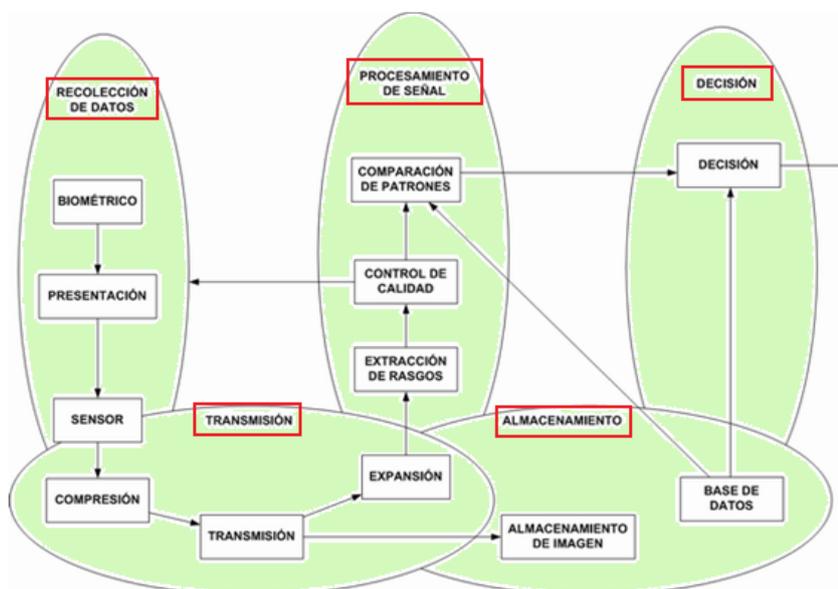


Figura 2.2: Esquema de un sistema biométrico genérico [13].

En el campo de la biometría se distinguen dos tipos de sistemas [11]:

- Sistemas de identificación.
- Sistemas de verificación.

2.3.1.- Sistemas de identificación

Un sistema biométrico en el modo de identificación tiene por objeto identificar la identidad del sujeto mediante la captura de ciertas características, para ello se hace una búsqueda exhaustiva en la base de datos que se tenga. Esto conduce a una comparación del tipo “uno-a-muchos” para establecer la

identidad del individuo. De este modo se responde a la pregunta: ¿Quién es usted?

2.3.2.- Sistemas de verificación

Un sistema biométrico en el modo de verificación comprueba la identidad de un sujeto, comparando la característica sólo con parámetros guardados anteriormente del individuo. Esto conduce a una comparación “uno-a-uno” para determinar si la identidad reclamada por el individuo es verdadera o falsa. De este modo se responde a la pregunta: ¿Es usted quién dice ser?

Dentro de un sistema de verificación (autenticación) es muy importante la etapa de Decisión, la cual establece si un usuario es válido o no. Podemos definir las dos características básicas de fiabilidad de un sistema de este tipo [29]:

- **Tasa de falsa aceptación** (False Acceptance Rate-FAR): esta tasa es la proporción esperada de operaciones falsamente reclamadas que son incorrectamente confirmadas. En identificación positiva y verificación se tendrá una falsa aceptación cuando erróneamente le sea asignada una identidad de un cliente a un individuo. Por otro lado, en la identificación negativa cuando sea rechazado un cliente.
- **Tasa de falso rechazo** (False Reject Rate-FRR): esta tasa es la proporción esperada de operaciones correctamente reclamadas que son incorrectamente rechazadas. En identificación positiva y verificación se produce un falso rechazo cuando se rechaza un cliente. Mientras, en identificación negativa cuando es confirmado como un cliente cuando en realidad no lo es.

Estas dos tasas son inversamente proporcionales, por lo que siempre debe llegarse a un compromiso. El punto de intersección entre la tasa de falsa aceptación y la tasa de falso rechazo se conoce como la tasa de error igual o tasa de equierror (Equal Error Rate-EER).

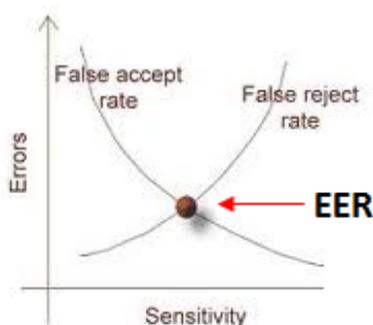


Figura 2.3: Gráfica FAR-FRR [12].

Generalmente es más difícil diseñar un sistema de identificación que uno de verificación, por ello los sistemas biométricos suelen ser sistemas de verificación.

2.4.- Indicadores biométricos actuales

Dentro de los indicadores biométricos se distinguen dos grandes grupos:

- Características físicas: son las que el sujeto tiene físicamente.
- Características del comportamiento: son las que el sujeto ha desarrollado.

En la Figura 2.4, se presenta un esquema que muestra cuáles de las características más comunes pertenecen a cada clase:

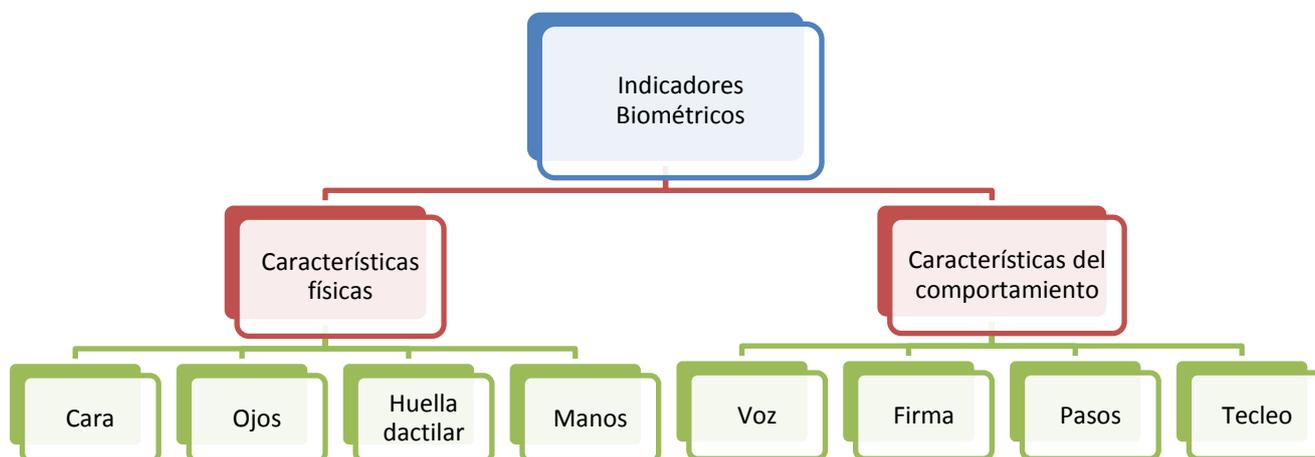


Figura 2.4: Esquema características biométricas.

2.5.- Principales sistemas

A continuación, se expone una breve descripción de las técnicas de reconocimiento más conocidas y utilizadas [7]:

2.5.1.- Huella dactilar

La comparación de la huella dactilar es una de las técnicas más antiguas, más utilizada y aceptada a nivel global.

Una huella dactilar es la impresión que produce el contacto de las crestas papilares de un dedo de la mano sobre una superficie, es decir, es la

representación de la morfología superficial de la epidermis de un dedo. Está formada por una serie de líneas que representan las crestas y unos espacios en blanco que representan los valles. A veces, se producen intersecciones entre las líneas y terminan de forma abrupta, dando lugar a lo que se conoce como minucias. La identificación con huellas digitales está basada principalmente en las minucias [15].

Otros puntos singulares de una huella dactilar son puntos que reciben el nombre de *cores* y *deltas*. La característica más interesante que presentan tanto las minucias como los puntos singulares *cores* y *deltas* es que son únicos para cada individuo y permanecen inalterados a través de su vida [13].

En España para la identificación de un individuo, se buscan como mínimo doce puntos característicos, aunque la obtención de al menos ocho ya tiene validez jurídica.

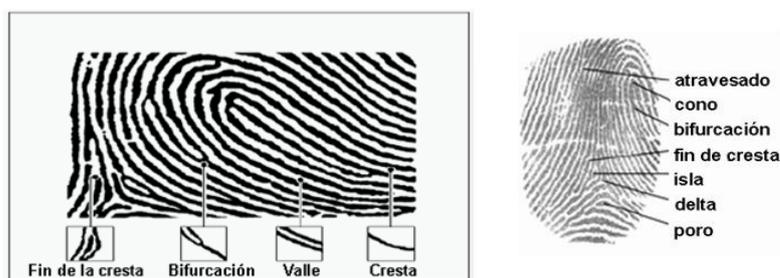


Figura 2.5: Características de huellas dactilares [13].

La forma general de procesar una huella dactilar se muestra en la figura 2.6:

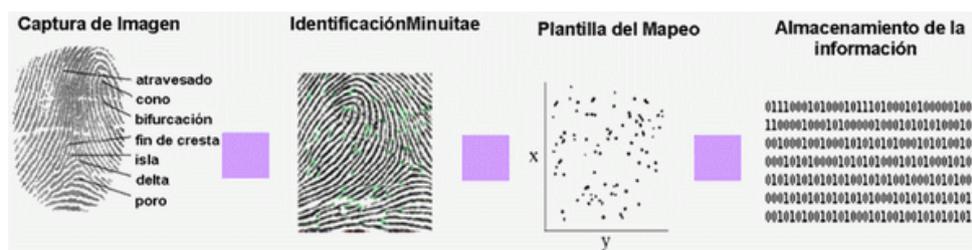


Figura 2.6: Proceso de escaneo de la huella dactilar [13].

Los sistemas basados en huellas dactilares son relativamente económicos comparados con otros sistemas biométricos.

Sin embargo, existe un importante problema ya que es imposible autenticar a un sujeto que haya sufrido un corte o una quemadura en el dedo. Además, este tipo de sistemas también se ve afectado cuando el dedo contiene suciedad o sudor.

2.5.2.- Patrones oculares

En los sistemas biométricos basados en patrones oculares se conocen dos técnicas atendiendo a distintas partes del ojo:

- **Retina:** se basa en la utilización del patrón de los vasos sanguíneos contenidos en la retina. Se trata de una fuente de información biométrica altamente distintiva ya que no existen dos patrones iguales, por lo que es una de las técnicas biométricas idónea para entornos de alta seguridad.

Presenta ciertas limitaciones como que es una técnica intrusiva pues es necesario que el usuario se coloque muy cerca del dispositivo de captura. Además existen una serie de enfermedades degenerativas de la retina, las cataratas y el glaucoma, que pueden alterar la imagen de la misma a lo largo del tiempo [16].

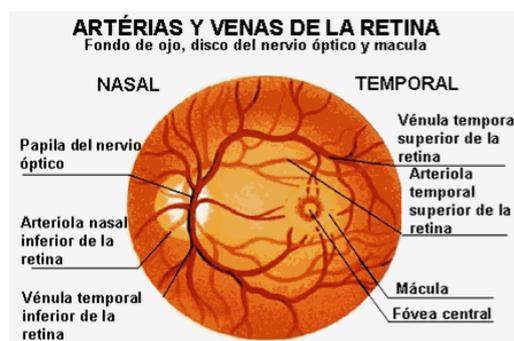


Figura 2.7: Partes de la retina (Fuente: usuarios.discapnet.es).

- **Iris:** el iris es la membrana coloreada y circular del ojo que separa la cámara anterior de la cámara posterior. Se ha demostrado que el iris tiene 266 puntos únicos y se cree que cada ojo es único y permanece estable con el tiempo, estas son las razones por las que es un sistema de autenticación tan bueno y que se encuentra en auge. Otra característica importante es que este tipo de tecnología adquiere los datos necesarios para su funcionamiento de forma no invasiva para el sujeto.

El proceso a seguir es el siguiente [13] [16]:

1. Reconocer el borde (localizar el iris).
2. Estirar zona de interés (pupila).
3. Clasificar el patrón.
4. Distinguir la persona.

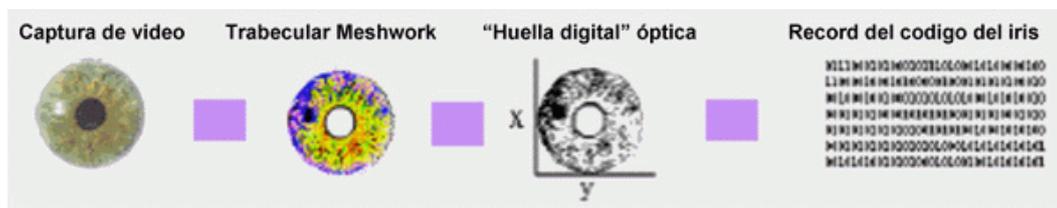


Figura 2.8: Reconocimiento del iris [13].

2.5.3.- Voz

La voz es el sonido que el aire expelido de los pulmones produce al salir de la laringe, haciendo que vibren las cuerdas vocales [17]. Es algo singular de cada ser humano y por ello se utiliza como técnica biométrica. Esta tecnología se soporta sobre la estructura física y características conductuales de cada persona. Existen dos formas de reconocer la voz:

- Dependencia del texto (texto específico o aleatorio).
- Independencia del texto.

Por otro lado, existen dos tipos de sistemas:

- Sistemas de identificación de locutor: para clasificar una señal de VOZ.
- Sistemas de verificación de locutor: para corroborar la voz de un sujeto.

El uso más común de las tecnologías de reconocimiento de voz es en el control por comandos. Otra aplicación de este tipo de tecnología son los sistemas diseñados para personas con problemas auditivos o que tienen dificultad para introducir información a un ordenador a través del teclado.

La identificación mediante el reconocimiento de voz se está introduciendo muy lentamente en el mercado, la principal razón es la mayor tasa de error que se obtiene frente a otros indicadores biométricos.

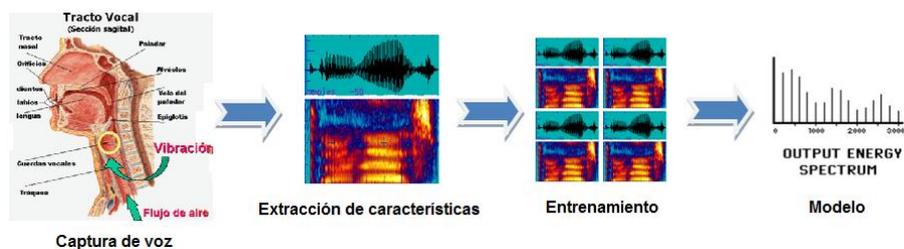


Figura 2.9: Reconocimiento de voz.

2.5.4.- Rostro

Otro sistema biométrico es el reconocimiento facial. El método más común utiliza una cámara para capturar una imagen de la cara de un sujeto, que es analizada en función de ciertos puntos clave, como la distancia entre los ojos, la anchura de la nariz, la distancia del ojo a la boca, o la longitud de la línea de la mandíbula. En función de estos parámetros se genera una plantilla única que luego es comparada con otras [18].

El reconocimiento facial puede ser menos exacto que otros sistemas como las huellas digitales, pero es menos invasivo.

Existen dos modos:

- **Autenticación de caras:** compara una imagen de la cara con otra imagen con la cara de la que se quiere saber la identidad.
- **Identificación de caras:** compara la imagen de una cara desconocida con todas las imágenes de caras conocidas que se encuentran en la base de datos.

La tecnología está mejorando más en la verificación de identidad que en la identificación de un rostro entre una multitud.

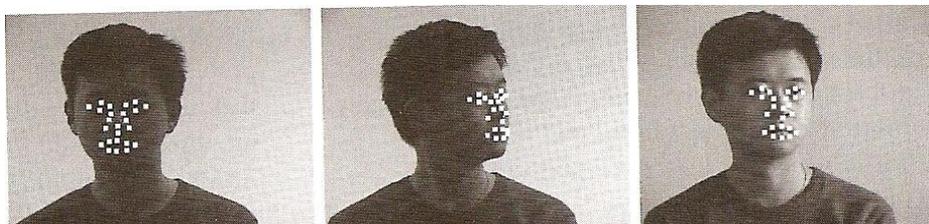


Figura 2.10: Rastreo de rostro [30].

2.5.5.- Geometría de la mano

Los sistemas basados en la geometría de la mano usan una cámara para capturar dos imágenes ortogonales bidimensionales de la palma y lados de la mano. Se realizan medidas del ancho, la altura, y longitud de la mano, etc.

Se trata de un sistema rápido, económico y simple, por ello tiene una gran aceptación y es usado en muchos ámbitos. Pero para algunos casos, como por ejemplo en lugares que se requiera alta seguridad, no es aceptada por sus falsos positivos.

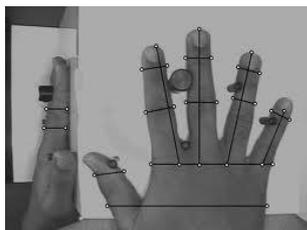


Figura 2.11: Posicionamiento geométrico de la mano [29].

2.5.6.- Andadura

Consiste en un sistema biométrico conductual que se basa en examinar la forma de andar de un sujeto. Se puede realizar reconocimiento a corta o larga distancia. También el reconocimiento puede basarse en la figura humana (de forma estática) así como en su movimiento.

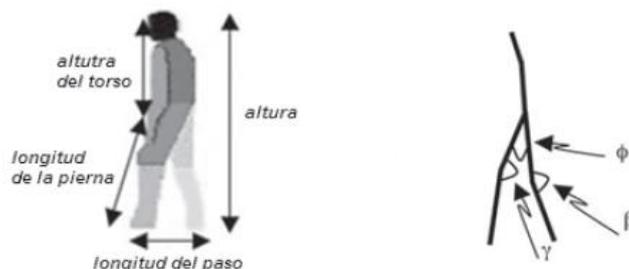


Figura 2.12: Parámetros de la forma y del movimiento [19].

2.5.7.- Dinámica de teclado

Es un sistema biométrico conductual usado para verificar la identidad de un sujeto examinando sus patrones de dinámica de teclado. Se cree que cada individuo tiene un patrón de tecleo distintivo y una cadencia de tecleo. La mayoría de los estudios usan la duración entre tecleo como una característica de verificación, aunque hay otros que utilizan el tiempo que permanece la tecla presionada. Según diversos estudios se obtienen mejores resultados usando los tiempos de presión de la tecla que los tiempos de latencia, pero lo mejor es el uso simultáneo de ambas.



Figura 2.13: Dinámica de teclado.

2.5.8.- Firma manuscrita

La firma manuscrita es una característica biométrica de comportamiento. Además, se ha usado desde la antigüedad, de ahí su gran aceptación social y que no presente rechazo por parte de los usuarios. Por ello la verificación de firma está muy extendida. En un primer momento se utilizó la verificación visual (que aún es muy utilizada) y posteriormente surgió la digital. En la verificación digital no solo se utilizan características estáticas, si no también dinámicas, como el tiempo que se tarda en firmar, la presión ejercida, la inclinación, etc.



Figura 2.14: Captura de firma.

Los sistemas biométrico basados en firma manuscrita es el objeto de estudio de este proyecto, se hará un mayor análisis de la técnica en capítulos posteriores.

2.6.- Comparativa de los principales sistemas

No existe una modalidad biométrica que sea mejor para todas las implementaciones. Hay que tener en cuenta muchos factores al implementar un dispositivo biométrico: la ubicación, los riesgos de seguridad, el tipo de tarea (identificación o verificación), cantidad de usuarios, circunstancias de utilización, base de datos, etc. Otro factor importante es que las técnicas biométricas están en distintas etapas de desarrollo.

En la Tabla 2.1, a modo de resumen, se recogen las diferentes características de los sistemas biométricos más utilizados:

	Ojo (Iris)	Ojo (Retina)	Huellas dactilares	Vascular dedo	Vascular mano	Geometría de la mano	Escritura y firma	Voz	Cara 2D	Cara 3D
Fiabilidad	Muy alta	Muy Alta	Muy Alta	Muy Alta	Muy Alta	Alta	Media	Alta	Media	Alta
Facilidad de uso	Media	Baja	Alta	Muy Alta	Muy Alta	Alta	Alta	Alta	Alta	Alta
Prevención de ataques	Muy alta	Muy Alta	Alta	Muy Alta	Muy Alta	Alta	Media	Media	Media	Alta
Aceptación	Media	Baja	Alta	Alta	Alta	Alta	Muy Alta	Alta	Muy alta	Muy alta
Estabilidad	Alta	Alta	Alta	Alta	Alta	Media	Baja	Media	Media	Alta

Tabla 2.1: Tabla comparativa sistemas biométricos [12].

En la Tabla 2.2 se muestra una comparativa con las principales ventajas e inconvenientes de los sistemas biométricos:

Tecnología	Ventajas	Inconvenientes
Huella dactilar	<ul style="list-style-type: none"> Alto grado de madurez Costes de implantación reducidos Buena aceptación 	<ul style="list-style-type: none"> Incompatibilidad con determinados trabajos manuales
Reconocimiento de voz	<ul style="list-style-type: none"> No requiere inversión en dispositivos Posibilidad de autenticación remota 	<ul style="list-style-type: none"> El ruido de fondo dificulta la captura Dificultad para reconocer ciertas formas de hablar
Reconocimiento facial	<ul style="list-style-type: none"> Reconocimiento en multitudes Identificación a media distancia Buena aceptación 	<ul style="list-style-type: none"> Escasa resistencia al fraude Unicidad limitada
Reconocimiento de iris	<ul style="list-style-type: none"> Patrones muy complejos Unicidad muy alta Alto grado de permanencia 	<ul style="list-style-type: none"> Coste de implantación alto Menor grado de aceptación
Reconocimiento de retina	<ul style="list-style-type: none"> Unicidad muy alta Alto grado de permanencia 	<ul style="list-style-type: none"> Precisa de total colaboración del usuario
Reconocimiento de la geometría de la mano	<ul style="list-style-type: none"> Alto grado de permanencia Facilidad de uso 	<ul style="list-style-type: none"> Unicidad limitada
Reconocimiento de firma	<ul style="list-style-type: none"> Buena aceptación Facilidad de uso 	<ul style="list-style-type: none"> Dificultad de captura por cambios de posición
Reconocimiento de escritura de teclado	<ul style="list-style-type: none"> No requiere inversión en dispositivos Posibilidad de realizar monitorización 	<ul style="list-style-type: none"> Tecnología emergente

Tabla 2.2: Ventajas e inconvenientes de los sistemas biométricos [20].

En la Figura 2.15 se presenta un gráfico del grado de implantación en el mercado de los principales sistemas biométricos:

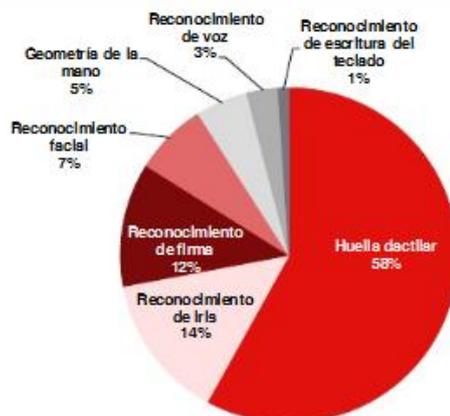


Figura 2.15: Grado de implantación sistemas biométricos (Fuente: Central Institute of Technology of Kokrajhar(2010)).

Como conclusión de la Figura 2.15 y la Tabla 2.2 se deduce que la lectura biométrica de las minucias de la huella digital es, sin duda, el sistema biométrico más avanzado y seguro del mercado.

El Instituto Nacional de Tecnologías de Comunicación (Inteco), en un estudio sobre las tecnologías biométricas aplicadas a la seguridad de 2011, destaca las ventajas de usar sistemas biométricos dados los crecientes ataques informáticos. También, destacan la escasa oferta de soluciones integradas, lo cual limita notablemente la biometría. Otra limitación de esta tecnología es la actual situación socio-económica. En términos económicos, la situación de crisis que todos los países (en mayor o menor medida) están viviendo, provoca que se paralice el desarrollo y el despliegue de estos sistemas. Por otro lado en relación a lo social, se teme el uso que pueda hacerse de estos datos.

Capítulo 3:

Introducción a la firma manuscrita

3.1.- Historia de la firma

La firma es una característica singular y dinámica propia de cada individuo. No se tiene certeza del origen exacto de ésta, pero se sabe que tiene una historia jurídica, ornamental, eclesiástica, diplomática, comercial y de identidad.

En Roma, existía una ceremonia (Manufirmatio) en la que se leía un documento por su autor (o el funcionario), se colocaba desenrollando y extendido sobre la mesa del escribano y luego se pasaba la mano abierta sobre el pergamino en actitud de jurar, pero sin hacerlo, se estampaba el nombre, signo, o una o tres cruces, por el autor o el funcionario en su nombre, después también lo hacían unos testigos.

En la Edad Media, lo que se hacía era escribir una cruz a la que se le añadían diversas letras y signos, lo cual se utilizaba como firma. Por aquel entonces la mayoría no sabía leer ni escribir, por lo que los nobles remplazaron esta práctica con el uso de sellos.

Posteriormente, la inscripción manuscrita del nombre o de los apellidos hizo que se comenzase a diferenciar entre firmas y signos. Fue entonces cuando la firma manuscrita empezó a cobrar la importancia que hoy tiene en la sociedad actual [21].

3.2.- Firma: definición, características y elementos

La firma consiste en el nombre y apellido que una persona plasma en un documento con su propio puño y letra. Cuando se refiera a firma autógrafa, debe firmar directamente la propia persona siendo inválida cualquier otra alternativa.

Una firma tiene tres características principales [21]:

- **Identificativa** → reconocer al firmante.
- **Declarativa** → indica conformidad del firmante.
- **Probatoria** → determina si el firmante es aquel que el sistema ha reconocido como autor.

Por otro lado, la firma está formada por varios elementos:

- **Elementos formales** → elementos materiales de la firma que están relacionados con los procedimientos utilizados al firmar y el grafismo.
 - Firma como signo personal.
 - Animas signandi (voluntad de asumir el contenido del documento que se firma).

- **Elementos funcionales** → indica la intención de la firma (identificar o autenticar).

3.3.- Sistema de reconocimiento de firmas

El principal objetivo de los sistemas de reconocimientos de firmas es evitar las falsificaciones que tan extendidas están en la sociedad actual y tan peligrosas son.

Como se ha visto en capítulos anteriores dentro de los sistemas de reconocimiento biométricos se distinguen dos tipos, esto mismo ocurre para el caso particular de firmas:

- **Identificación:** proceso en el cual se ve quién es el sujeto, comparando la firma con las muestras almacenadas en la base de datos de la que se dispone.
- **Verificación:** proceso que sirve para comprobar que un sujeto es realmente quien dice ser. La firma se extrae de la base de datos y se compara. (Es el objeto de estudio de este proyecto).

La verificación de firmas está formada por una serie de etapas:

- Adquisición de la firma.
- Acondicionamiento de la señal.
- Extracción de características.
- Preprocesado de datos.
- Comparación de patrones de firma.

Por otro lado, mencionar que se suelen considerar tres tipos de sujetos:

- **Sujeto verdadero** → aporta firmas genuinas.
- **Sujeto impostor casual** → intenta suplantar la identidad del verdadero, pero desconoce la firma de éste.
- **Sujeto imitador** → intenta suplantar la identidad del verdadero, conoce perfectamente su firma y consigue una reproducción muy fiel a la original.

Por último, mencionar que el reconocimiento de firmas puede llevarse a cabo a través de dos métodos:

- **Métodos estáticos (Off-line):** permiten verificar características estáticas biométricas que no varían con el tiempo. Tanto la inscripción como las pruebas se realizan con muestras previas. Se basan en el reconocimiento de patrones.

- **Métodos dinámicos (On-line):** permiten verificar características dinámicas biométricas que pueden variar con el tiempo. Tanto la inscripción como las pruebas se realizan en el momento en que se capturas las muestras. No es preciso almacenar los datos.

3.4.- Adquisición de la firma y dispositivos de captura

Como hemos visto en el apartado anterior, existen dos técnicas de adquisición de firmas: off-line y on-line [22]. A continuación, en los siguientes apartados se describen cada una de ellas.

3.4.1. - Adquisición Off-line

Se denomina captura off-line cuando en primer lugar se realiza la firma y posteriormente se captura. La captura en este caso consiste en digitalizar la firma que fue escrita en papel. Esta técnica tiene una serie de problemas, donde destaca la pérdida de información, como ocurre con la presión ejercida al firmar o la duración. Sin embargo, presenta alguna ventaja, ya que la verificación puede usarse sin la presencia del sujeto y que en general resulta más económico en cuanto al equipamiento necesario.

En este contexto, el dispositivo de captura puede ser un scanner o una cámara.



Figura 3.1: Adquisición off-line.

3.4.2.- Adquisición On-line

Se denomina captura on-line cuando la adquisición de los datos es simultánea a la ejecución de la firma. Permite contar además de con la firma original con algunos datos extras, como la presión, la velocidad, la duración, el punto de inicio y fin de los trazos, entre otros.

Al recibir la información directa e inmediatamente del lápiz, cabe la posibilidad de aprender y ajustar el sistema de reconocimiento en tiempo real (mejorando así los índices de acierto).

En este contexto, el dispositivo de captura debe ser una tableta digital. Uno de los mayores inconvenientes de estos sistemas es que requieren el uso de material especial y la posibilidad de que varíen ciertas características ya que no es lo mismo escribir con un bolígrafo sobre papel que con un lápiz de plástico sobre una pantalla táctil.



Figura 3.2: Tableta digital [9].

3.5.- Parámetros de estudio en una firma

Una firma tiene diversas características y además depende de multitud de factores como la posición, el bolígrafo, el medio donde se firma, el estado de ánimo, el espacio donde se firma, etc. Por todo esto, una persona no firma dos veces igual y ahí reside la dificultad del reconocimiento.

Las principales características de una firma, que una tableta es capaz de recoger, son las siguientes:

- ▶ Coordenada x.
- ▶ Coordenada y.
- ▶ Presión.
- ▶ Azimut.
- ▶ Elevación.
- ▶ Nº de muestras.

A partir de estas pueden obtenerse otras como las velocidades o las aceleraciones.

Gracias a estudios anteriores, como la tesis doctoral “Uso de la Firma Manuscrita Dinámica para el Reconocimiento Biométrico de Personas en Escenarios Prácticos” [23], se concluye que la coordenada Y, la presión y las velocidades de X e Y, son las 4 características más significativas y discriminantes. Por lo tanto, son las más

interesantes para desarrollar el algoritmo de reconocimiento y son las que se han usado en este proyecto.

3.6- Marco legal y tecnológico

Los rasgos biométricos son datos personales, por lo que la privacidad es un gran reto de esta tecnología. La Ley Orgánica de Protección de Datos (LOPD) contempla cómo deben ser tratados. Este tipo de información está considerada "de protección básica" según la ley. Sin embargo, según los usuarios, es percibida como información muy sensible (Inteco [12]).

Debe mencionarse la precaución que debe tenerse al almacenar datos de esta índole en cualquier sistema. La LOPD de fecha 14 de diciembre de 1999, establece que los datos de carácter físico suponen un nivel alto de protección de la información. Las leyes establecen que los datos físicos son datos sensibles, pero no todos los datos biométricos deben tratarse como datos sensibles.

Los datos personales deben estar declarados en la Agencia de Protección de Datos (APD) y deben establecerse una serie de medidas de nivel básico-medio:

- ▶ Documento de seguridad: normas a seguir por el personal para garantizar la protección de los datos.
- ▶ Registro de incidencias.
- ▶ Control de acceso.
- ▶ Gestión de soporte.

Además, se debe garantizar la confidencialidad cuando viaja la información (mediante cifrado). Por otro lado, se debe asegurar que la información contenida en el sistema biométrico solo se usará para el fin a la que ha sido destinada.

En la directiva 95/46/CE del Parlamento Europeo y del Consejo de 24 de octubre de 1995, se define como datos personales: "*toda información sobre una persona física identificada o identificable; se considerará identificable toda persona cuya identidad pueda determinarse, directa o indirectamente, en particular mediante un número de identificación o uno o varios elementos específicos, característicos de su identidad física, fisiológica, psíquica, económica, cultural o social*". Así se establece que los datos biométricos deben ser considerados y tratados como datos personales.

De esta directiva se deduce que existen restricciones a la hora de desarrollar un sistema biométrico y que requiere una regulación. Esta regulación es reciente y por ello los sistemas biométricos aún no están tan desplegados como se desearía.

En cuanto al marco tecnológico, mencionar que en rasgos generales se dispone de la tecnología necesaria y se puede observar la evolución de la sociedad hacia un mundo cada vez más digitalizado, por lo tanto se poseen todos los factores y herramientas necesarias para que en un futuro próximo el uso de los sistemas biométricos en la seguridad de la sociedad de la información esté muy presente.

Capítulo 4:

Algoritmos de reconocimiento de firmas: DTW

Los principales algoritmos, empleados en los sistemas de reconocimiento de firmas dinámicas, se basan en la clasificación de patrones. Las principales técnicas se agrupan en tres áreas [24]:

- **Métodos basados en alineamiento de características:**

Estos métodos consisten en la comparación de la muestra tomada con una muestra almacenada previamente, denominada plantilla.

El método de referencia de esta categoría es el DTW (Dynamic Time Warping) o lo que es lo mismo el Alineamiento Temporal Dinámico. Éste es el objeto de estudio de este proyecto, por ello se explicará en detalle más adelante.

- **Métodos basados en modelos estadísticos:**

Los métodos basados en modelos estadísticos se basan en el uso de patrones de referencia para elaborar modelos probabilísticos y/o estadísticos.

Los métodos más empleados de esta área son el algoritmo HMM (Hidden Markov Models), es decir, los Modelos Ocultos de Markov y el GMM o Modelos de Mezclas de Gaussianas.

- **Métodos basados en fronteras de decisión:**

En cuanto a los métodos basados en fronteras de decisión, se puede decir que consisten en crear unas fronteras entre clases en función de un determinado criterio de error, el cual se forja a partir de la relación entre los resultados obtenidos y los deseados.

En este campo destacan los árboles de decisión, las redes neuronales y las máquinas de vectores de soporte.

4.1.- Alineamiento Temporal Dinámico - DTW

En esta sección se explica la técnica de referencia en sistemas de reconocimiento de firma manuscrita: el alineamiento temporal dinámico o DTW (Dinamic Time Wrapping).

Esta técnica tiene su origen en el campo de reconocimiento de voz [25], donde tiene un gran reconocimiento. Posteriormente se descubrieron sus buenas cualidades en cuanto a reconocimiento de firma se refiere, donde destacan los estudios de Sato y Kogure [26].

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.

En primer lugar, construye una matriz de distancia entre el patrón y los puntos de muestreo y a continuación encuentra la mejor manera de alinear la muestra con el patrón, buscando el camino óptimo.

Existen algunas restricciones, locales y globales, con las que se intenta mejorar el algoritmo, ya sea minimizando la distancia o reduciendo el número de cálculos. Las restricciones locales buscan el camino de menor distancia variando la pendiente, se muestran las distintas restricciones locales o conjuntos de producciones en la Tabla 4.1. En cuanto a los límites globales, consisten en variar la ventana, su tamaño o forma, para intentar reducir al mínimo el número de puntos a calcular. En la Figura 4.1 se presentan los dos tipos de ventana por excelencia: el paralelogramo de Itakura y la banda de Sakoe-Chiba [25].

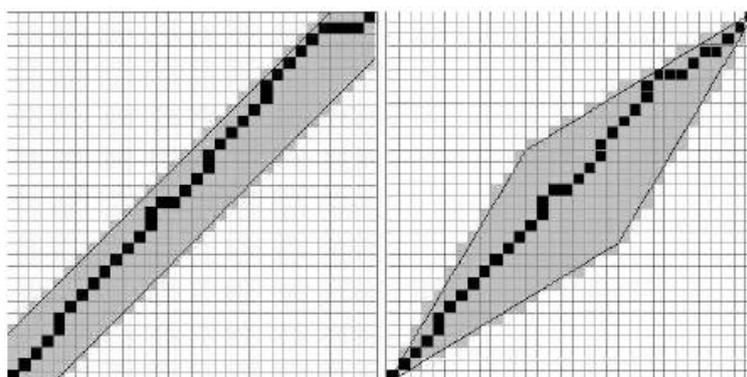


Figura 4.1: A la izquierda: Banda de Sakoe-Chiba. A la derecha: Paralelogramo de Itakura [32].

De forma genérica, se presenta la función de alineamiento a través de la Figura 4.2, en la cual se observan los límites globales, la ventana que limita por dónde se pueden realizar los cálculos y las restricciones locales, que indica la pendiente que se va eligiendo en cada instante:

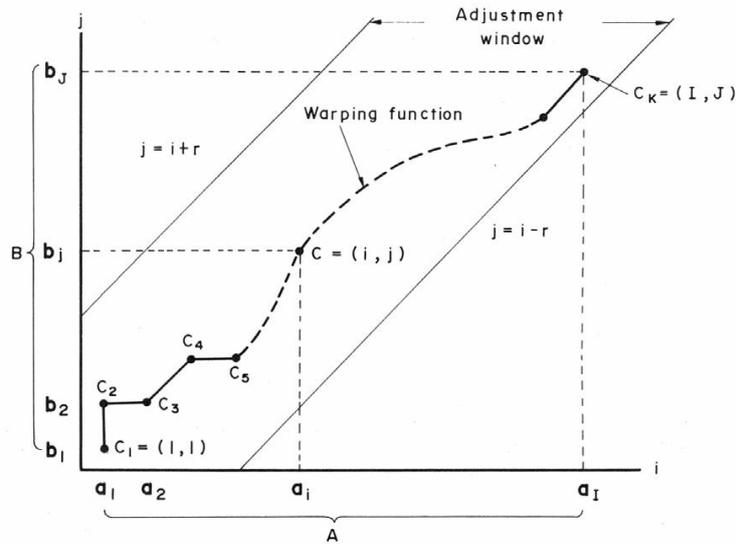


Figura 4.2: Algoritmo DTW [25].

De forma general, el algoritmo DTW consiste en una serie de pasos que se explican a continuación [1] [26] [27]:

- Se requieren dos matrices para el cálculo:
 - $D \rightarrow$ matriz de distancias.
 - $G \rightarrow$ matriz de programación dinámica, para obtener el camino mínimo entre el punto final y el inicial.

Se parte de 2 vectores:

$$A = \{a_1, a_2, a_3, \dots, a_i, \dots, a_M\}$$

$$B = \{b_1, b_2, b_3, \dots, b_j, \dots, b_N\}$$

Se denota como C a la función de alineamiento:

$$C = \{c(1), c(2), \dots, c(k), \dots, c(K)\}$$

$$c(k) = [i(k) \ j(k)] \rightarrow \text{cada } c(k) \text{ tiene una función de coste: } d\{c(k)\}.$$

Esta función muestra la diferencia entre los elementos comparados.

Como función de coste se toma la distancia euclídea, aunque podrían usarse otras distancias como la Manhattan:

$$d\{c(k)\} = (a_{i(k)} - b_{j(k)})^2$$

La función de alineamiento es aquella que minimice el coste total:

$$D(C) = \sum_{k=1}^K d\{c(k)\} \cdot w(k)$$

Donde $w(k)$ es un coeficiente de ponderación.

La función de alineamiento debe cumplir una serie de restricciones:

- 1) Monotonicidad: El camino debe tener el sentido de izquierda a derecha, y de arriba debajo de la matriz G.

$$i_k \leq i_{k+1} \quad j_k \leq j_{k+1}$$

- 2) Continuidad: El paso de un nodo al siguiente debe hacerse de modo que no haya nodos intermedios.

$$i_k - i_{k-1} \leq 1 \quad j_k - j_{k-1} \leq 1$$

- 3) Frontera: El principio y el final deben coincidir con ciertos puntos, independientemente del camino seguido.

$$i(1) = 1 \quad j(1) = 1 \quad i(K) = I \quad j(K) = J$$

- 4) Ventana: Condición de ajuste (límite global):

$$|i(k) - j(k)| \leq \Delta$$

Donde Δ es un entero positivo.

- 5) Pendiente: Se aplican restricciones (límites locales) sobre las pendientes del camino para encontrar el camino óptimo.

En la Tabla 4.1 se muestran las principales restricciones locales (conjuntos de producciones P), y dependiendo de cual se implemente se pueden obtener distintos resultados [25].

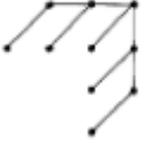
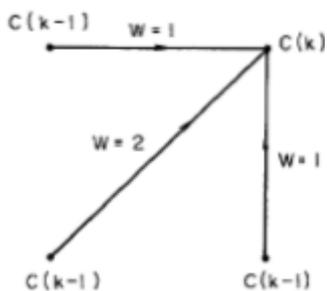
P	Esquema	Simétrico/ Asimétrico	DP – Ecuación $g(i,j)=$
0		Simétrico	$\min \begin{bmatrix} g(i, j - 1) + d(i, j) \\ g(i - 1, j - 1) + 2d(i, j) \\ g(i - 1, j) + d(i, j) \end{bmatrix}$
		Asimétrico	$\min \begin{bmatrix} g(i, j - 1) \\ g(i - 1, j - 1) + d(i, j) \\ g(i - 1, j) + d(i, j) \end{bmatrix}$
1/2		Simétrico	$\min \begin{bmatrix} g(i - 1, j - 3) + 2d(i, j - 2) + d(i, j - 1) + d(i, j) \\ g(i - 1, j - 2) + 2d(i, j - 1) + d(i, j) \\ g(i - 1, j - 1) + 2d(i, j) \\ g(i - 2, j - 1) + 2d(i - 1, j) + d(i, j) \\ g(i - 3, j - 1) + 2d(i - 2, j) + d(i - 1, j) + d(i, j) \end{bmatrix}$
		Asimétrico	$\min \begin{bmatrix} g(i - 1, j - 3) + (d(i, j - 2) + d(i, j - 1) + d(i, j))/3 \\ g(i - 1, j - 2) + (d(i, j - 1) + d(i, j))/2 \\ g(i - 1, j - 1) + d(i, j) \\ g(i - 2, j - 1) + d(i - 1, j) + d(i, j) \\ g(i - 3, j - 1) + d(i - 2, j) + d(i - 1, j) + d(i, j) \end{bmatrix}$
1		Simétrico	$\min \begin{bmatrix} g(i - 1, j - 2) + 2d(i, j - 1) + d(i, j) \\ g(i - 1, j - 1) + 2d(i, j) \\ g(i - 2, j - 1) + 2d(i - 1, j) + d(i, j) \end{bmatrix}$
		Asimétrico	$\min \begin{bmatrix} g(i - 1, j - 2) + (d(i, j - 1) + d(i, j))/2 \\ g(i - 1, j - 1) + d(i, j) \\ g(i - 2, j - 1) + d(i - 1, j) \end{bmatrix}$
2		Simétrico	$\min \begin{bmatrix} g(i - 2, j - 3) + 2d(i - 1, j - 2) + 2d(i, j - 1) + d(i, j) \\ g(i - 1, j - 1) + 2d(i, j) \\ g(i - 3, j - 2) + 2d(i - 2, j - 1) + 2d(i - 1, j) + d(i, j) \end{bmatrix}$
		Asimétrico	$\min \begin{bmatrix} g(i - 2, j - 3) + 2((i - 1, j - 2) + d(i, j - 1) + d(i, j))/3 \\ g(i - 1, j - 1) + d(i, j) \\ g(i - 3, j - 2) + d(i - 2, j - 1) + d(i - 1, j) + d(i, j) \end{bmatrix}$

Tabla 4.1: Restricciones locales (conjuntos de producciones P) [25].

En esta explicación del algoritmo se han tomado una serie de decisiones de diseño, para entenderlo en profundidad. Anteriormente, se ha mencionado que para el cálculo de la distancia se ha escogido la distancia euclídea aunque podrían usarse otras. En cuanto a la restricción local que se ha utilizado, ha sido por simplicidad la restricción P(0) Simétrica [Figura 4.3] :



Simétrico

Figura 4.3: P(0) Simétrico.

En esta restricción el valor mínimo de distancia se calcula con la siguiente expresión:

$$\min \begin{bmatrix} g(i, j - 1) + d(i, j) \\ g(i - 1, j - 1) + 2d(i, j) \\ g(i - 1, j) + d(i, j) \end{bmatrix}$$

Capítulo 5:

Diseño y Desarrollo de la aplicación

5.1.- Introducción

Como se ha mencionado en capítulos anteriores, el objetivo de este proyecto es contribuir de algún modo a la mejora de la de la verificación biométrica, que tan importante es en la sociedad de la información.

Dicha contribución consiste en evaluar el algoritmo DTW, realizando ciertas modificaciones y observando si la verificación mejora o por el contrario empeora. Se trata entonces de un proyecto de investigación con el cual se deben obtener unas conclusiones.

Para la elaboración del mismo se ha utilizado una base de datos de firmas, sin la cual habría sido imposible la realización de este proyecto. En el siguiente apartado se explica dicha base de datos.

5.2- Base de datos MCyT

La base de datos MCyT [28] es una base bimodal que contiene la huella dactilar y la firma manuscrita de 330 usuarios, sin embargo, en este proyecto se han usado exclusivamente las firmas. Se utiliza una versión libre formada por 100 usuarios.

Esta base de datos es una iniciativa de cuatro universidades españolas, cada una de ella aporta las firmas de algunos usuarios:

- ▶ Universidad Politécnica de Madrid: 145 usuarios.
- ▶ Universidad de Valladolid: 75 usuarios.
- ▶ Universidad del País Vasco: 75 usuarios.
- ▶ Escuela Universitaria Politécnica de Mataró: 35 usuarios.

Cada sujeto realiza 25 firmas originales y 25 firmas falsificadas de la firma original de otros sujetos.

La adquisición de la firma se llevó a cabo con una tableta digital *Wacom*, la cual permite firmar igual que con un bolígrafo sobre un papel. Al usuario se le hace firmar en un recuadro delimitado (área de captura: 127x97 mm).

La resolución de la tableta gráfica es de 100 líneas por milímetro y la precisión es de +/- 0.25mm. La frecuencia utilizada es 100 Hz.

La información que se obtiene de la tableta es:

- ▶ Coordenada x.
- ▶ Coordenada y.

- ▶ Presión ejercida.
- ▶ Ángulo Azimut.
- ▶ Ángulo de altitud.

En la Figura 5.2.1 se muestra un ejemplo de una firma verdadera y una falsificación de la misma:

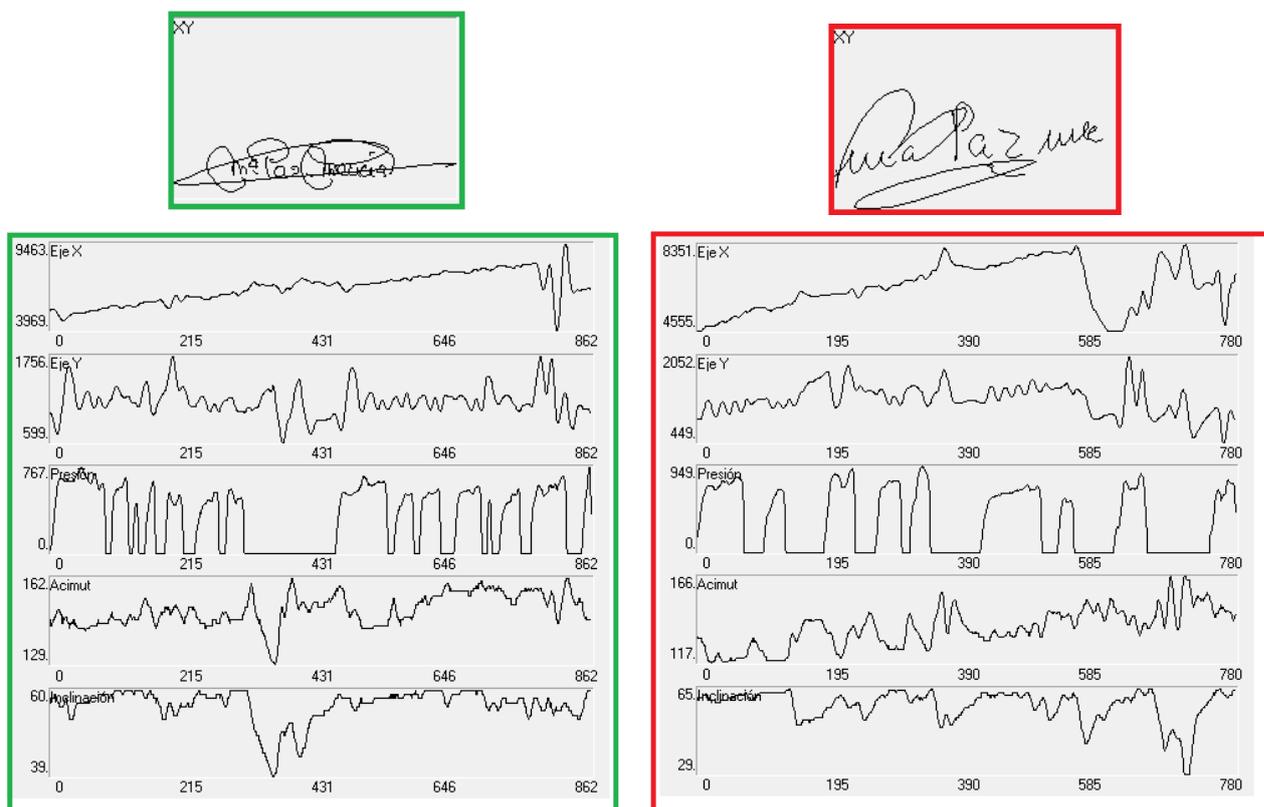


Figura 5.2.1: Izda.: Firma verdadera usuario 55 muestra 24; dcha.: Firma falsificada usuario 55 muestra 24[28].

5.3- Primera parte: Algoritmo DTW básico

El proyecto se divide en dos partes, en la primera se implementa el algoritmo DTW con unas características concretas, para entender en profundidad su funcionamiento. Este primer DTW se denotará como “básico” a lo largo del proyecto. Esta primera aproximación se ha basado en el artículo de Sakoe y Chiba [25].

En el DTW básico no se ha utilizado ningún tipo de restricción global, es decir, no se ha controlado el tamaño de la ventana. Mientras que como restricción local se ha utilizado la programación $P(0)$ Simétrica [Figura 4.3]. Por otro lado, se ha utilizado para el cálculo de distancias la distancia euclídea, por ser en general la más extendida y

utilizada. Por último, se han normalizado los datos respecto de la media y la desviación típica, para normalizarlo en función del tamaño y la posición.

Como datos de estudio para el DTW básico se ha realizado una primera aproximación solo para la coordenada x, por simplicidad. Posteriormente, se ha extendido a otras características.

El código se divide en varias funciones, que se han ido usando en los diferentes pasos. Se representa el camino crítico (mínima distancia) y una comparativa de la señal original y la alineada.

Los prototipos de las funciones que se han implementado son:

$[Xa, Ya, za, aza, ina, ppsa] = FPG_Signature_Read(firma1_path, 0, 0);$ → Función incluida en la base de datos MCyT, que extrae las características de una firma.

$[Xa, Xb] = normalizar(Xa, Xb);$ → Función que normaliza la coordenada x, respecto de la media y la desviación típica.

$[D, G, Dist, w, Dist_norm] = dtw_basico(signal1, signal2);$
→ Función que calcula la distancia mínima, el DTW propiamente dicho.

$[D, G, Dist, w, Dist_norm] = main(firma1_path, firma2_path);$
→ Función principal que va llamando a las anteriores.

A modo de ejemplo, el resultado de $[D, G, Dist, w, Dist_norm] = main('0002v00.fpg', '0002v01.fpg')$, se muestra en las gráficas Figura 5.3.1 y Figura 5.3.2.

En la Figura 5.3.1 se presenta el camino óptimo, el de mínima distancia que se ha ido generando gracias a la programación dinámica.

En la Figura 5.3.2 se representan dos gráficas, la de izquierda muestra la coordenada x de la firma 0 del usuario 2 (señal 1) y la coordenada x de la firma 1 del usuario 2 (señal 2). Al ser firmas del mismo individuo se observa que son muy parecidas. La gráfica de la derecha muestra las mismas señales pero tras ser alineada, es decir, tras aplicar el DTW básico, puede verse en este caso que ahora son aún más parecidas. Esto es así porque son de la misma persona, pero si se tratase de dos personas con firmas muy distintas la distancia entre ambas debería ser muy grande.

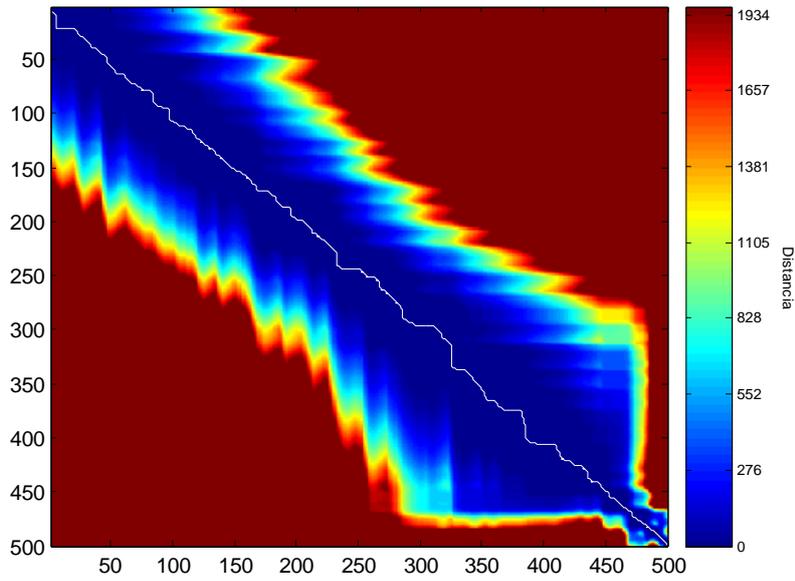


Figura 5.3.1: Distancia acumulada y camino óptimo.

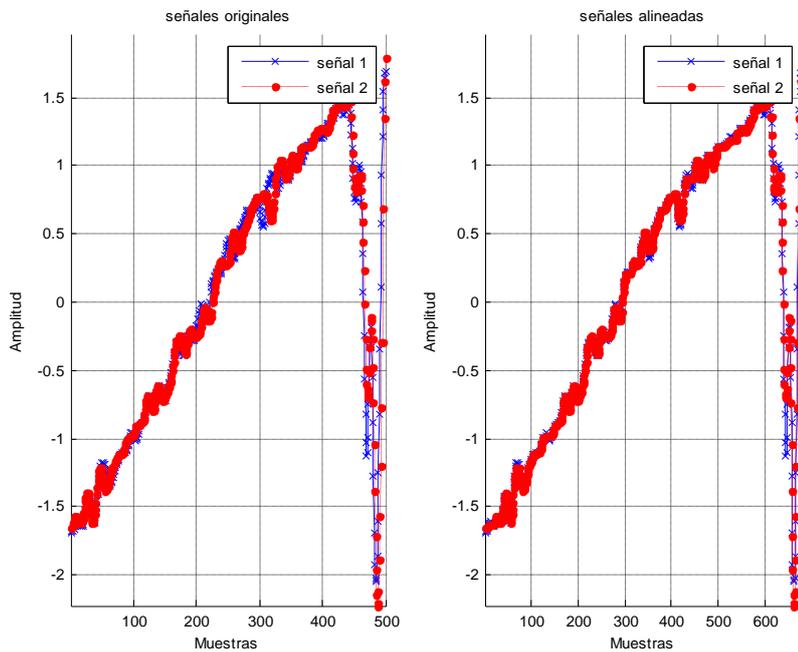


Figura 5.3.2: Señales originales y alineadas.

Posteriormente se implementó el algoritmo DTW básico de forma multidimensional, teniendo en cuenta las coordenadas x e y, y las velocidades. En la Figura 5.3.3 representamos las características utilizadas en la versión multidimensional.

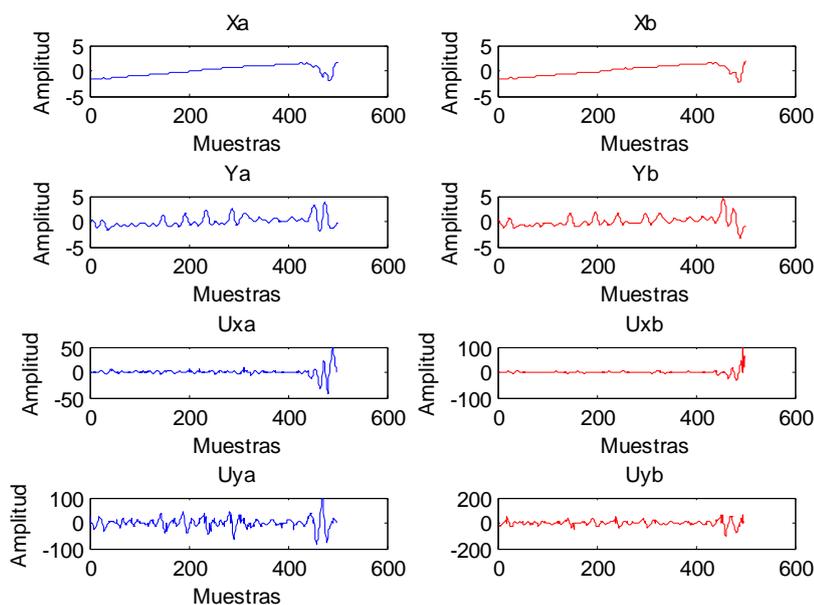


Figura 5.3.3: Variables de estudio: coordenada x e y, y velocidades (U_{xa} , U_{xb}).

Para el cálculo del algoritmo DTW multidimensional se hizo una primera versión sencilla que calculaba la distancia mínima de unas pocas firmas, usando un patrón formado por las 5 primeras firmas de un usuario. Dicho ejemplo se desarrolló solo para un usuario a modo de prueba.

Tras haber completado la primera parte del proyecto y haber estudiado en profundidad el DTW, se decidió usar para la segunda parte del proyecto un código más genérico y compacto que nos permitiese después realizar las modificaciones necesarias para realizar los experimentos objeto de este estudio de forma eficiente.

Este programa se creó teniendo en cuenta estudios previos y se tomaron unas decisiones de diseño en función de éstos, para conseguir una opción de algoritmo DTW con unas características concretas. Como decisiones más relevantes destacan:

Gracias a la tesis doctoral "Online Signature Verification Algorithms and Development of Signature International Standards" [9], la cual evalúa el funcionamiento del algoritmo variando el número de puntos, se concluye que submuestreando las firmas de manera que estén formadas sólo por 256 puntos, es suficiente para asegurar que los niveles de error obtenidos se mantengan en niveles aceptables. Por lo tanto, para los experimentos realizados en este proyecto se han empleado 256 puntos por firma.

La otra decisión de diseño, consiste en la elección de los parámetros más relevantes de una firma. Como parámetros se tienen: la coordenada x, la coordenada y, la presión ejercida al firmar, el azimut, la elevación y la frecuencia de muestreo. En este caso, la elección de los parámetros se ha hecho gracias a la tesis doctoral “Uso de la Firma Manuscrita Dinámica para el Reconocimiento Biométrico de Personas en Escenarios Prácticos” [23], en la cual se concluye que la coordenada Y, la presión y las velocidades de X e Y, son las 4 características más significativas y discriminantes.

En la Figura 5.3.4 se muestra un diagrama de flujo de las funciones, sus llamadas de forma secuencial y sus relaciones:

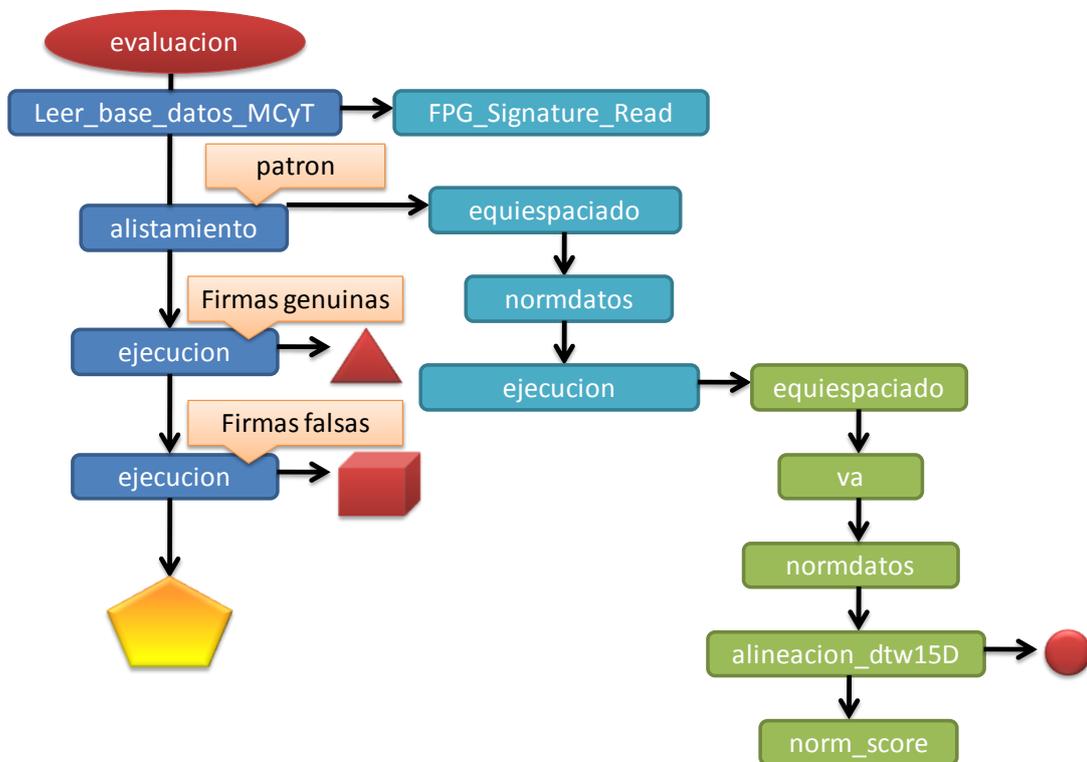


Figura 5.3.4: Diagrama de flujos funciones DTW básico (parte 1).

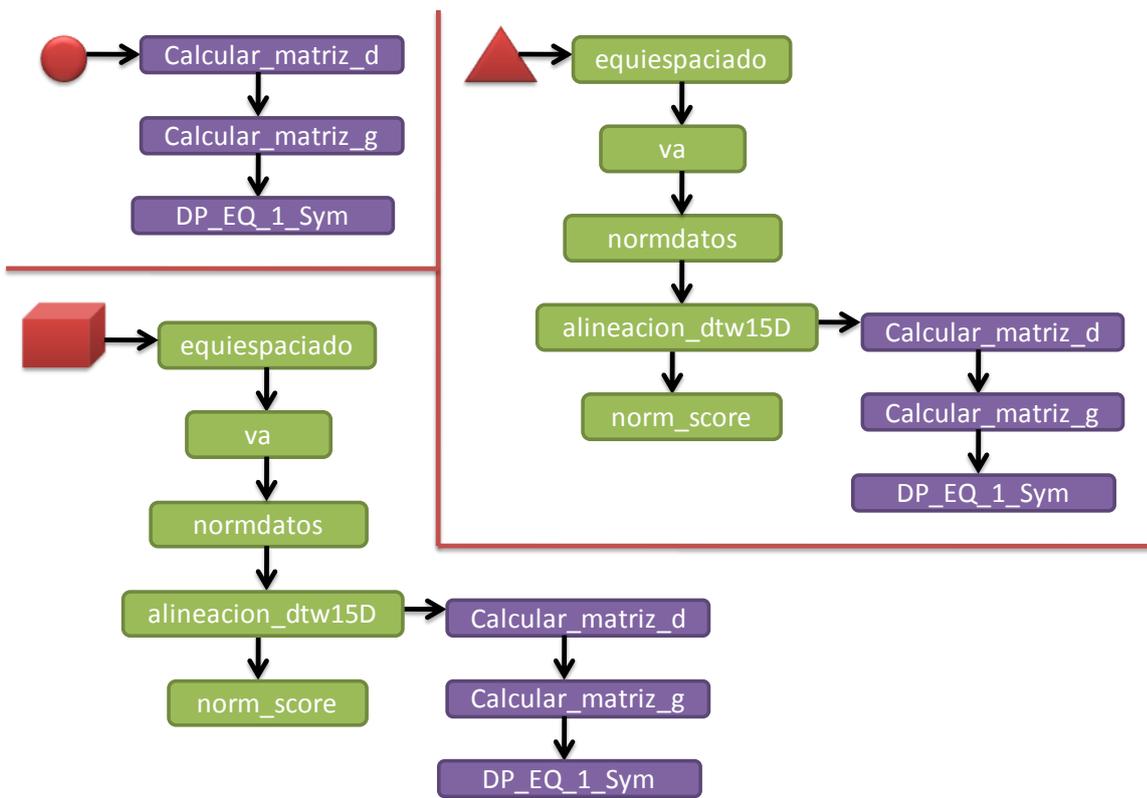


Figura 5.3.4: Diagrama de flujos funciones DTW básico (parte 2).

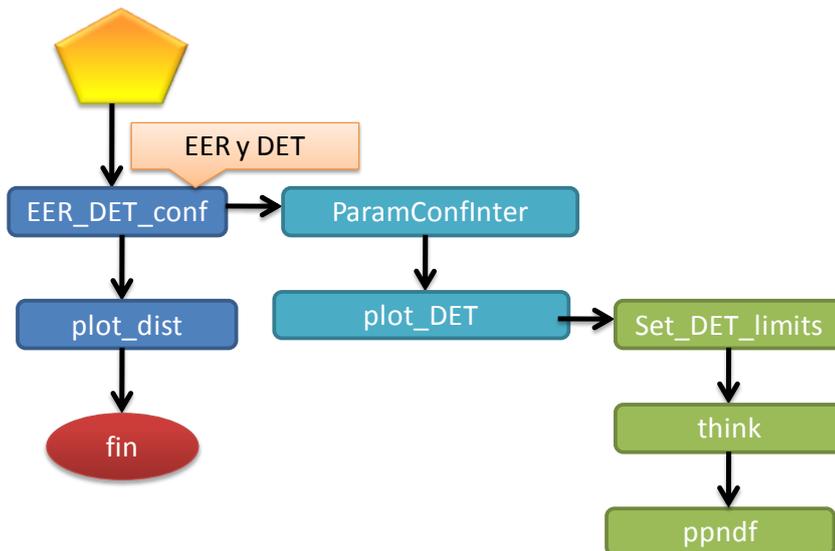


Figura 5.3.4: Diagrama de flujos funciones DTW básico (parte 3).

Los prototipos de estas funciones y sus descripciones son:

`evaluacion.m` → Script de evaluación del algoritmo DTW. Llama a las distintas funciones.

`leer_base_datos_MCyT(variables)` → Función que lee toda la base de datos de firmas MCyT.

`FPG_Signature_Read(filename, showXY, showWV)` → Función que lee una firma y extrae sus características: coordenada x, coordenada y, presión, elevación, azimut.

`alistamiento(signatures, variables)` → Función que genera el patrón, de tal forma que selecciona las 5 primeras firmas de cada usuario y va calculando el DTW quedándose con la distancia mínima.

`equiespaciado(x)` → Función que interpola para que el patrón y la muestra tengan el mismo número de puntos.

`normdatos(datos)` → Función que normaliza los datos (velocidad y aceleración) respecto del tamaño y la posición.

`ejecucion(fm, p, variables)` → Función que prepara las características de las firmas y llama a la función que calcula el DTW.

`va(datos)` → Función que calcula las derivadas (primera y segunda): velocidad y aceleración.

`alineacion_dtw15D(m, p, variables)` → Función que calcula el algoritmo DTW.

`norm_score(score, stats)` → Función que normaliza la distancia.

`calcular_matriz_d(m, p)` → Función que calcula la primera matriz de distancias.

`calcular_matriz_g(Matriz_D)` → Función que calcula la matriz de mínimas distancias.

`DP_EQ_1_Sym(i, j, Matriz_G, Matriz_D)` → Función que controla los límites de los índices para el cálculo de la mínima distancia.

`EER_DET_conf(clients, imposteurs, OPvalue, pas0, creciente), ParamConfInter(FAR, FRR, num_imposteurs, num_clients)` ,

Plot_DET (Pmiss, Pfa, plot_code, opt_thickness),
Set_DET_limits(Pmiss_min, Pmiss_max, Pfa_min, Pfa_max) ,
thick(w, lh) , ppndf (cum_prob) → Conjunto de funciones auxiliares de
representación, que interpretan los resultados.

5.4- Segunda parte: Algoritmo DTW modificado

Una vez se tiene el algoritmo DTW básico implementado se consigue el punto de partida, por lo tanto, se pueden realizar una serie de modificaciones con las que se busca mejorar de algún modo dicho algoritmo en el reconocimiento de firmas manuscritas.

El objetivo es conseguir que el DTW sea lo más óptimo posible, atendiendo al coste computacional y a la tasa de error, y para ello se ajustan algunos de sus parámetros.

Estas modificaciones son la base para conseguir unas conclusiones a cerca de la influencia de sus parámetros y de este modo optimizar el algoritmo. Dichas conclusiones suponen un aporte novedoso en el estudio del DTW.

Para la elección de las características a variar se han tomado de referencia varios estudios anteriores como el artículo de Sakoe y Chiba [25] o las tesis “Online Signature Verification Algorithms and Development of Signature International Standards” [9] y “Uso de la Firma Manuscrita Dinámica para el Reconocimiento Biométrico de Personas en Escenarios Prácticos” [23].

Se ha querido centrar este estudio en varias vías de investigación y se han elegido 3 conjuntos de experimentos:

- ◆ **1^{er} Conjunto de experimentos:** Tipos de restricciones locales.
- ◆ **2^o Conjunto de experimentos:** Tipos de distancias.
- ◆ **3^{er} Conjunto de experimentos:** Tipos de normalizaciones.

En el siguiente capítulo se analizarán en profundidad estas modificaciones, se explicarán los experimentos realizados y se expondrán las conclusiones que se han obtenido en cada uno de estos experimentos.

Capítulo 6:

Análisis del algoritmo y pruebas

6.1- Introducción

En este capítulo se explicaran los conjuntos de experimentos y las modificaciones que se han realizado, y las conclusiones que se han obtenido al variar los parámetros del algoritmo.

En esta segunda parte, se han implementado una serie de funciones nuevas que se han añadido a las ya existentes en la primera parte del proyecto. Debido al gran coste computacional que requiere la evaluación de toda la base de datos, se ha decidido implementar las funciones más “pesadas” en C++ y combinarlo con Matlab. Con esta decisión se reduce notablemente el tiempo de procesamiento, lo cual ha hecho posible realizar diversos experimentos.

Para poder usar conjuntamente C++ y Matlab se ha utilizado la herramienta MEX, la cual es capaz de compilar código en C++ en un archivo ejecutable de Matlab. Un ejemplo de compilación sería: `mex -v -c función.cpp`, tras esto ya se podría llamar a la función desde una función o script desarrollado íntegramente en Matlab.

Para poder usar dicho entorno se ha tenido que investigar sobre este tipo de programación, además de instalar y configurar una serie de herramientas.

Como se ha mencionado anteriormente, se ha decidido organizar los experimentos en 3 conjuntos:

- ◆ **1^{er} Conjunto de experimentos:** Tipos de restricciones locales. En este primer conjunto de experimentos se variará este parámetro pero se mantendrán constantes las otras dos modificaciones: la distancia, donde se ha elegido usar la euclídea, y la normalización donde se ha decidido usar la media y la desviación típica.
- ◆ **2^o Conjunto de experimentos:** Tipos de distancias. En la segunda tanda de experimentos la modificación consistirá en variar la distancia pero se mantendrán constantes las otras dos modificaciones: la restricción local, donde se escoge la óptima tras las conclusiones de los experimentos de la primera modificación. Y la normalización, donde se utiliza la media y la desviación típica.
- ◆ **3^{er} Conjunto de experimentos:** Tipos de normalizaciones. En la tercera modificación se variará la normalización pero se mantendrán constantes las otras dos modificaciones: la restricción local, donde se escoge la óptimo tras las conclusiones de los experimentos del primer conjunto de experimentos. Y el

tipo de distancia donde se utiliza la óptima tras las conclusiones de los experimentos de la segunda modificación.

6.2- Escenario de los experimentos

Como se ha explicado en el apartado 2 del capítulo 5 (5.2- *Base de datos MCyT*), para la realización de este proyecto se ha empleado la base de datos MCyT y por tanto el escenario de experimentación ha sido el mostrado en la Figura 6.2:

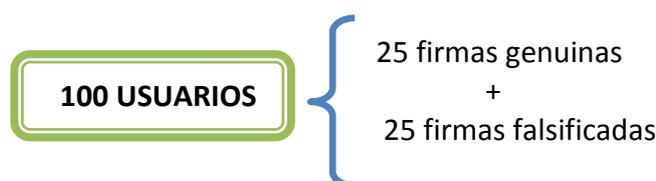


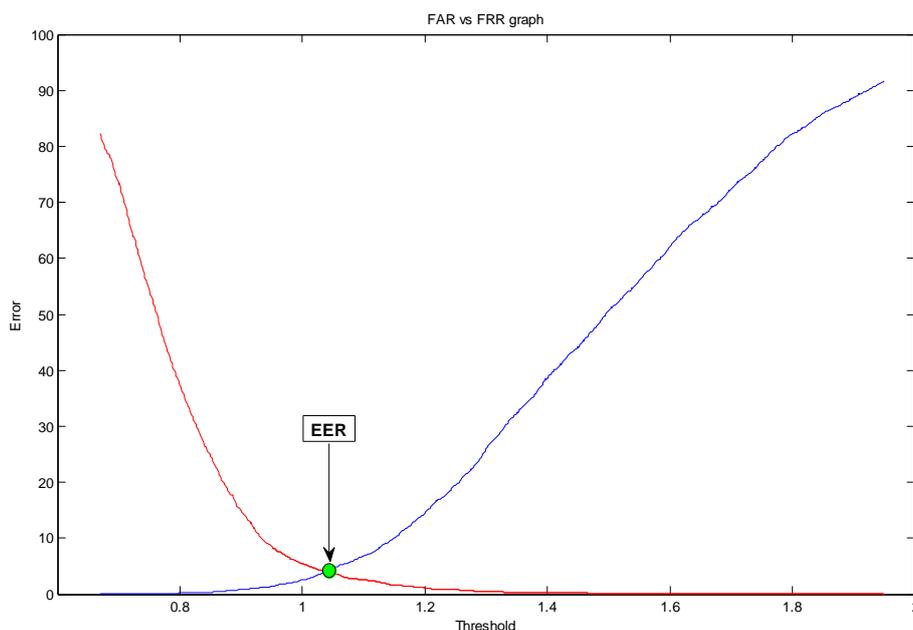
Figura 6.2: Escenario de los experimentos.

En apartados posteriores se explicarán los resultados obtenidos con dicho escenario.

6.3- Sistemas de evaluación

Para evaluar el algoritmo en cada caso se han utilizados 3 gráficas [29]: la gráfica FAR (False Acceptance Rate) vs FRR (False Reject Rate), la curva ROC (Receiver Operating Characteristic) y la curva DET (Detection Error Tradeoff).

- 1) **FAR vs FRR:** En el capítulo 2, en el apartado "2.3.2.- *Sistemas de verificación*", se ha explicado la diferencia entre la tasa de falsa aceptación (FAR) y la de falso rechazo (FRR). Representar una frente a otra aporta una información muy valiosa puesto que permite deducir cómo de seguro es el sistema. El punto en el que ambos valores se cruzan es el EER (Equal error rate) y cuanto menor sea éste mejor será el sistema de verificación, más exacto. En la Figura 6.3.1 se muestra cómo sería la representación gráfica:



- 2) **Curva ROC:** La curva ROC (Receiver Operating Characteristic) se basa en la sensibilidad y la especificidad, siendo la primera la probabilidad de clasificar correctamente a un individuo cuyo estado real sea el definido como positivo respecto a la condición que estudia la prueba. Mientras, la especificidad es la probabilidad de clasificar correctamente a un individuo cuyo estado real sea el definido como negativo. La curva representa los verdaderos positivos (VPR = Razón de Verdaderos Positivos) frente al ratio de falsos positivos. Nuestro sistema será mejor cuanto mayor sea la VPR y menor la FAR. En la Figura 6.3.2 se presenta a modo de ejemplo una gráfica que muestra una curva ROC. El sistema “ideal” representaría una curva en la que la FAR fuese igual a 0 y (1-FRR) fuese 1, por lo tanto en los casos reales el sistema es mejor cuanto más se acerque la curva a la esquina superior izquierda.

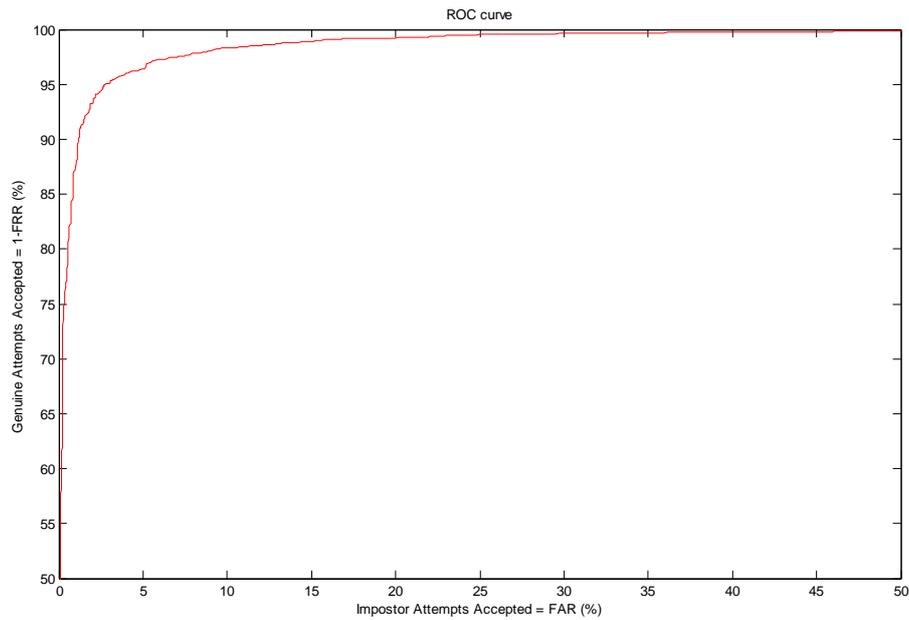


Figura 6.3.2: Curva ROC.

- 3) **Curva DET:** La curva DET (Detection Error Tradeoff) es una variación de la ROC pero los ejes son en escala logarítmica. Permite una comparación visual entre sistemas, más clara y fácil de realizar. Cuanto más cerca de cero se encuentre la curva mejor será el sistema.

En la Figura 6.3.3 se presenta un ejemplo de cómo sería una curva DET.

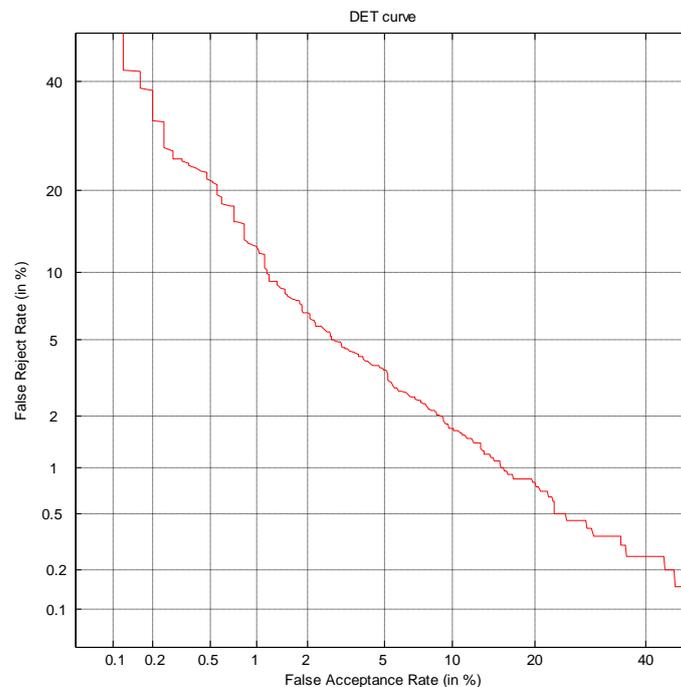


Figura 6.3.3: Curva DET.

En estas gráficas se ha utilizado un umbral de decisión, que define el grado de seguridad deseado. Se trata de un número que indica el mínimo grado de similitud permitido para aceptar a un usuario como verdadero. En este proyecto se ha calculado un umbral distinto en cada experimento puesto que se calcula como la distancia máxima obtenida para las firmas genuinas, menos la distancia mínima obtenida para las firmas falsas (tanto las entrenadas como las aleatorias). Además, para normalizar el umbral se divide entre un factor 1000. Con este umbral, se calcula la FAR y la FRR, con lo se obtiene la tasa de equierror.

6.4- Experimentos y pruebas

Como se ha mencionado anteriormente, para el desarrollo de los experimentos se han establecido dos escenarios de estudio: escenario casual y escenario seguro.

- ▶ **Escenario casual:** es el escenario en el que se utilizan firmas falsas aleatorias de la base de datos, es decir, propias de otros usuarios. En este caso puede darse más prioridad al coste computacional, puesto que requiere menos seguridad, se trata de un entorno menos abierto a intrusismo.
- ▶ **Escenario seguro:** en este escenario se utilizan firmas falsas entrenadas, es decir, otros usuarios entrenan la falsificación de firmas. Este escenario requiere máxima seguridad, la prioridad es no permitir que falsos usuarios sean reconocidos como verdaderos. Por lo tanto, se requiere una tasa de equierror lo más baja posible.

Además, para la evaluación de las prestaciones, los experimentos se centran en obtener la tasa de equierror y el coste computacional. Es importante resaltar que en el coste computacional tiene un papel muy importante las características del ordenador utilizado en los experimentos, puesto que éstas influyen en el tiempo de ejecución y procesamiento del programa. En este caso, se ha utilizado un ordenador de sobremesa con procesador AMD Athlon 64 X2 Dual Core 2,60Hz, con 4GB de memoria RAM y sistema operativo Windows Vista Home Premium de 32 bits.

En los siguientes apartados se expondrán los conjuntos de experimentos que se han realizado y las conclusiones de cada una de las modificaciones realizadas.

6.4.1- 1^{er} Conjunto de experimentos: Tipos de restricciones locales

En el capítulo 4, donde se ha descrito el algoritmo DTW, se han explicado los límites locales (restricciones) que se usan para obtener el camino de mínima distancia. En dicho capítulo también puede verse una tabla (Tabla 4.1: Restricciones locales (conjuntos de producciones P) [25]) que muestra los 8 caminos posibles.

Como primera modificación se ha decidido implementar el algoritmo con las 8 restricciones locales posibles y evaluar los resultados con cada una de ellas. Para ello se ha implementado la función `alineacion_dtw15D_Modificacion1_Caminos`, en la cual se ha ido modificando la subfunción que controla los límites. Dicha subfunción, se ha implementado en C++, puesto que es la que más coste computacional conlleva.

A continuación, se analiza cada tipo de restricción local y se presentan los resultados que se han obtenido al utilizar cada una de ellas:

◆ Restricción Local 0 Simétrica:  $\min \begin{bmatrix} g(i, j - 1) + d(i, j) \\ g(i - 1, j - 1) + 2d(i, j) \\ g(i - 1, j) + d(i, j) \end{bmatrix}$

En este experimento se ha utilizado la función `calcular_g_Sym_0`.

Restricción local	P(0) Simétrico 
Norma	L ₂
Normalización	Media y desv. típica
EER_FA (%)	0,35
EER_FE (%)	3,63
Tiempo (s)	4404,85

Tabla 6.4.1.1: Resultados experimento P(0) Simétrico.

En la Tabla 6.4.1.1 se observa que el EER de las firmas falsas aleatorias (FA) es 0,35%, mientras que el de las firmas falsas entrenadas (FE) es 3,63%. El equierror de las aleatorias no llega a ser el 10% del valor de las entrenadas, lo cual es coherente, como se ha explicado anteriormente al entrenar la falsificación detectarla es mucho más difícil.

En la Figura 6.4.1.1 se representan las curvas DET y ROC que se han obtenido en este experimento. Las líneas azules representan el resultado cuando se han utilizado las falsas entrenadas (FE) y las rojas con las falsas aleatorias (FA). Se concluye que ambas gráficas son lógicas, puesto que en el caso de la DET el error es menor cuanto más se acerque al punto (0,0) y se puede apreciar que la curva de las falsas aleatorias está más cerca de dicha posición. Esto mismo ocurre con la ROC, la de las falsas aleatorias está muy próxima a 1.

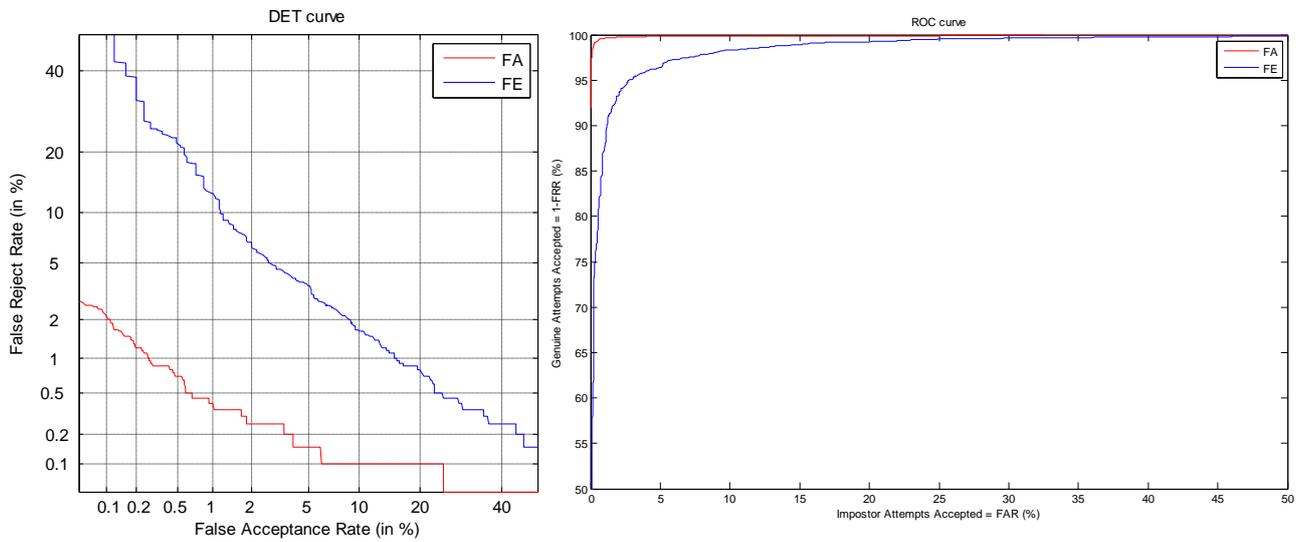


Figura 6.4.1.1: Curva DET y ROC para P(0) Simétrico.

En cuanto al tiempo de procesamiento, se ha obtenido un tiempo total de 4004,85 segundos, que si se divide entre 100, que es el número total de usuarios y ese valor se divide a su vez entre 50, que es el número total de firmas utilizadas, se obtiene un tiempo total en promedio por firma de 0,8 segundos.

◆ Restricción Local 0 Asimétrica: $\nearrow \min \begin{bmatrix} g(i, j - 1) \\ g(i - 1, j - 1) + d(i, j) \\ g(i - 1, j) + d(i, j) \end{bmatrix}$

En este experimento se ha utilizado la función *calcular_g_Asym_0*.

Restricción local	P(0) Asimétrico \nearrow
Norma	L_2
Normalización	Media y desv. típica
EER_FA (%)	0,59
EER_FE (%)	3,98
Tiempo (s)	4329,78

Tabla 6.4.1.2: Resultados experimento P(0) Asimétrico.

En la Tabla 6.4.1.2 re resumen los resultados obtenidos para este experimento. El EER para las firmas falsas entrenadas es 3,98% y para las falsas aleatorias 0,59%. En este caso, la tasa de equierror de las aleatorias es casi el 15% de la tasa de las entrenadas. Si se compararan estos valores con los del caso anterior, el experimento del P(0) Simétrico, observamos un incremento de casi el 40% en las falsas aleatorias y en torno

al 9% en las entrenadas. La diferencia entre ambos caminos es mucho más significativa para el caso de firmas aleatorias.

En la Figura 6.4.1.2 se muestran las curvas DET y ROC para este experimento. En ellas se observa la importante diferencia que existe entre las falsas entrenadas.

En cuanto al coste computacional, para este caso se obtiene un tiempo total de 4329,78 segundos, por lo que cada firma se tarda en evaluar 0,87 segundos. Comparándolo con el caso simétrico, se observa que es solo 7 centésimas de segundo mayor, lo cual no es relevante.

Por todo ello, se puede concluir que el caso simétrico es más aconsejable que el asimétrico, puesto que, en general, la tasa de equierror es significativamente menor.

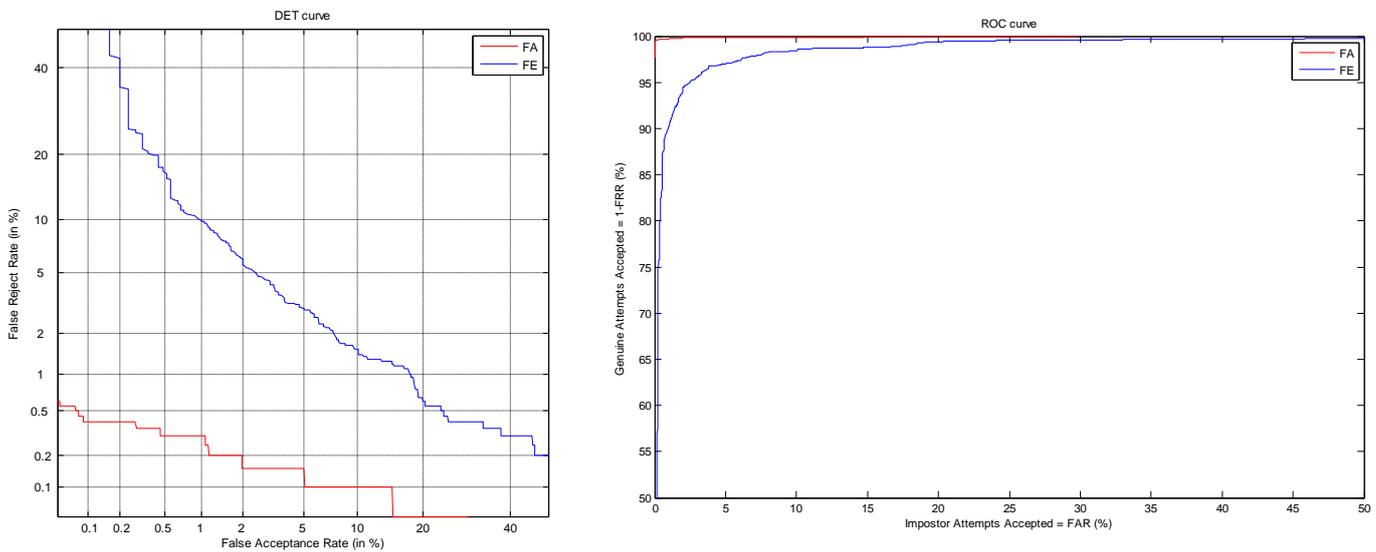
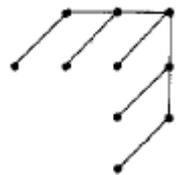


Figura 6.4.1.2: Curva DET y ROC para P(0) Asimétrico.

◆ Restricción Local 1/2 Simétrica:



$$\min \begin{bmatrix} g(i-1, j-3) + 2d(i, j-2) + d(i, j-1) + d(i, j) \\ g(i-1, j-2) + 2d(i, j-1) + d(i, j) \\ g(i-1, j-1) + 2d(i, j) \\ g(i-2, j-1) + 2d(i-1, j) + d(i, j) \\ g(i-3, j-1) + 2d(i-2, j) + d(i-1, j) + d(i, j) \end{bmatrix}$$

En este experimento se ha utilizado la función `calcular_g_Sym_05`.

Restricción local	P(1/2) Simétrico
Norma	L_2
Normalización	Media y desv. típica
EER_FA (%)	0,4
EER_FE (%)	4,07
Tiempo (s)	5045,82

Tabla 6.4.1.3: Resultados experimento P(1/2) Simétrico.

En la Tabla 6.4.1.3 se recogen los resultados obtenidos para el experimento P(1/2) Simétrico. El EER para las firmas falsas aleatorias es aproximadamente 0,4% y para las entrenadas 4,07%. En esta situación el valor de las aleatorias no llega a ser el 10% del de las entrenadas. Por otro lado, el tiempo de ejecución total es 5045,82 segundos y para una sola firma es de 1,01 segundos.

En la Figura 6.4.1.3 se presenta la representación de las curvas DET y ROC. Una vez más resalta la evidente diferencia entre el caso de las entrenadas y las aleatorias.

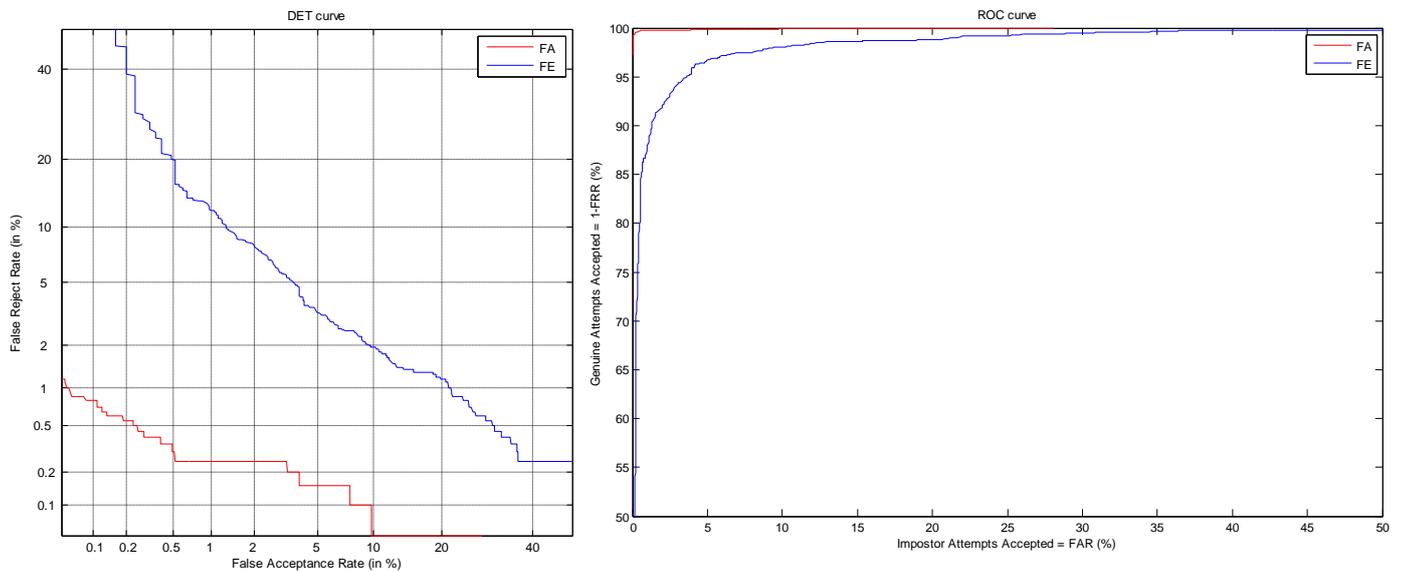
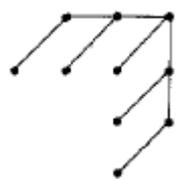


Figura 6.4.1.3: Curva DET y ROC para P(1/2) Simétrico.

◆ Restricción Local 1/2 Asimétrica:



$$\min \begin{bmatrix} g(i-1, j-3) + (d(i, j-2) + d(i, j-1) + d(i, j))/3 \\ g(i-1, j-2) + (d(i, j-1) + d(i, j))/2 \\ g(i-1, j-1) + d(i, j) \\ g(i-2, j-1) + d(i-1, j) + d(i, j) \\ g(i-3, j-1) + d(i-2, j) + d(i-1, j) + d(i, j) \end{bmatrix}$$

En este experimento se ha utilizado la función *calcular_g_Asym_05*.

Restricción local	P(1/2) Asimétrico	
Norma	L ₂	
Normalización	Media y desv. típica	
EER_FA (%)	0,31	
EER_FE (%)	3,99	
Tiempo (s)	5080,3	

Tabla 6.4.1.4: Resultados experimento P(1/2) Asimétrico.

En la Tabla 6.4.1.4 se presentan los resultados que se han obtenido para el experimento P(1/2) Asimétrico. El EER para las firmas falsas aleatorias está en torno al 0,31% mientras que el de las entrenadas es cercano al 4%. En este caso, la diferencia es aún mayor puesto que el equierror de las aleatorias es aproximadamente el 7,7% del equierror de las entrenadas. Comparando estos valores con el caso simétrico, se concluye que con el caso asimétrico se obtiene una tasa de equierror un 2% menor para las aleatorias y un 22% menor para las entrenadas. Además, la diferencia entre las entrenadas y las aleatorias se reduce.

En cuanto al tiempo de procesamiento, se obtienen 5080,3 segundos, lo cual se transforma en 1,02 segundos aproximadamente, de media para cada usuario y firma. Anteriormente, en el caso simétrico se obtenía un tiempo medio de 1,01 segundos por usuario y firma, la diferencia entre ambos valores es muy pequeña, por lo que compensa ese incremento de tiempo frente a los beneficios que nos aporta reducir el error.

En la Figura 6.4.1.4 se muestran las curvas DET y ROC para esta evaluación.

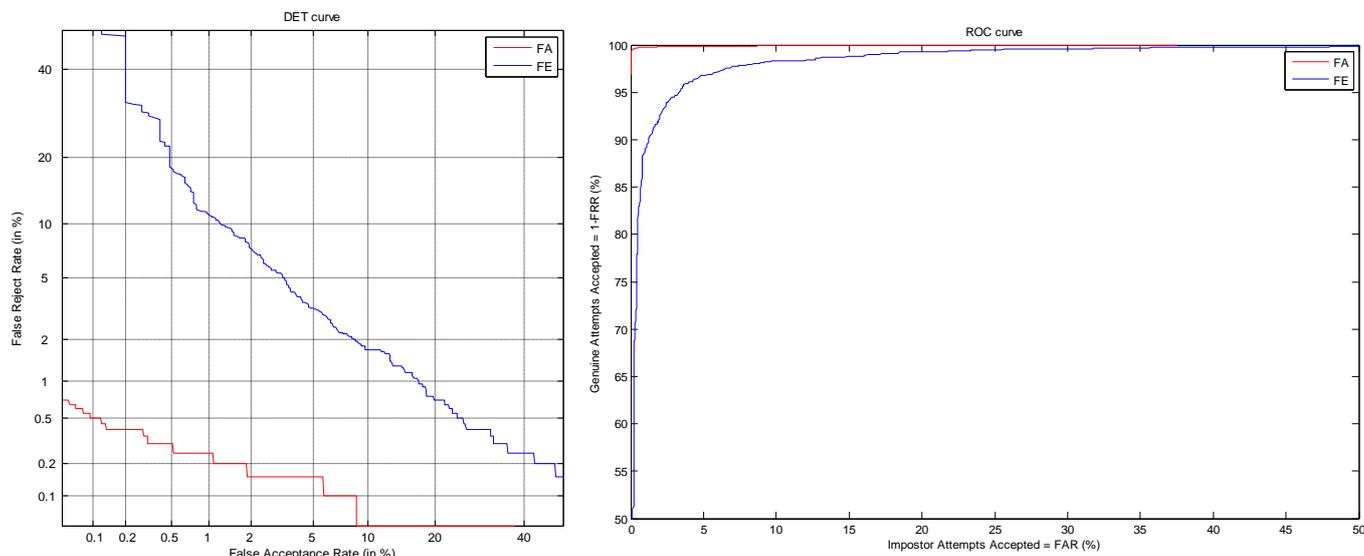


Figura 6.4.1.4: Curva DET y ROC para P(1/2) Asimétrico.

◆ Restricción Local 1 Simétrica:
$$\min \begin{bmatrix} g(i-1, j-2) + 2d(i, j-1) + d(i, j) \\ g(i-1, j-1) + 2d(i, j) \\ g(i-2, j-1) + 2d(i-1, j) + d(i, j) \end{bmatrix}$$

En este experimento se ha utilizado la función *calcular_g_Sym_1*.

Restricción local	P(1) Simétrico
Norma	L ₂
Normalización	Media y desv. típica
EER_FA (%)	0,46
EER_FE (%)	4,76
Tiempo (s)	4618,93

Tabla 6.4.1.5: Resultados experimento P(1) Simétrico.

En la Tabla 6.4.1.5 se agrupan los resultados obtenidos con el experimento P(1) Simétrico, donde el equierror para las firmas falsas aleatorias es 0,46% y para las entrenadas 4,76%. La diferencia entre ambas es muy elevada, la tasa de equierror de las entrenadas es más de 10 veces el valor de las aleatorias.

En lo referente al tiempo de ejecución, se tiene un tiempo total de 4618,93 segundos que en promedio para una firma de un usuario es menor a un segundo (0,92 segundos).

En la Figura 6.4.1.5 se muestran las curvas DET y ROC para este experimento. Se aprecia un comportamiento lógico y acorde a los resultados esperados.

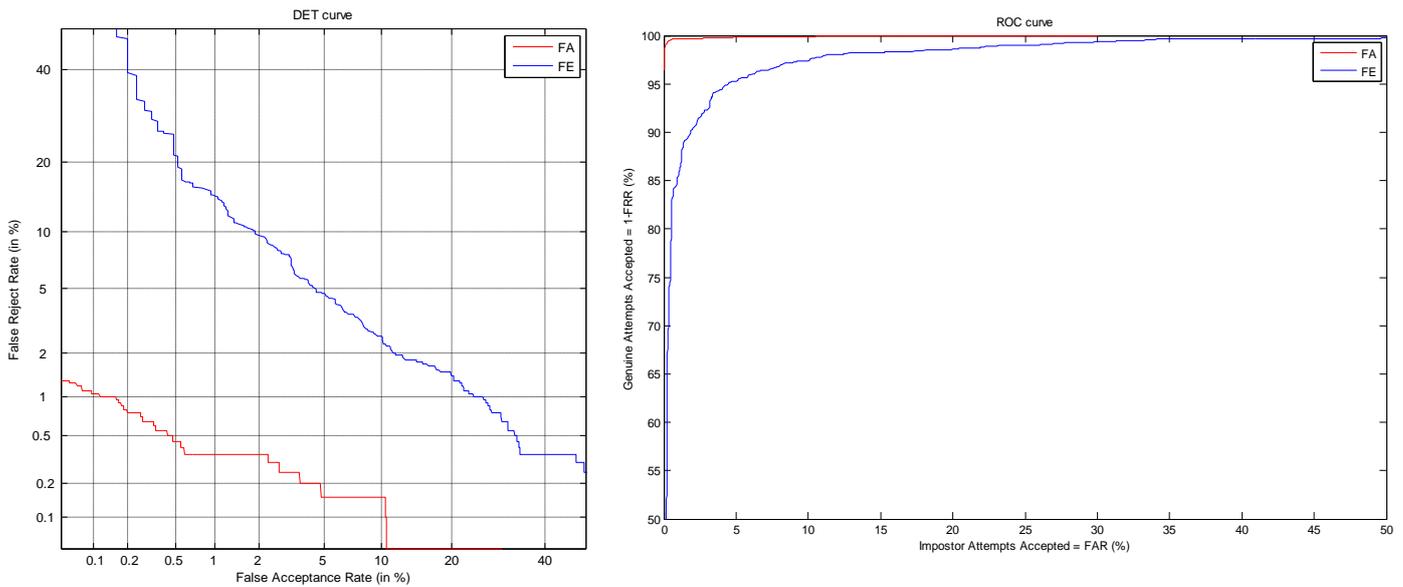
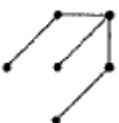


Figura 6.4.1.5: Curva DET y ROC para P(1) Simétrico.

◆ Restricción Local 1 Asimétrica: 
$$\min \begin{bmatrix} g(i-1, j-2) + (d(i, j-1) + d(i, j))/2 \\ g(i-1, j-1) + d(i, j) \\ g(i-2, j-1) + d(i-1, j) \end{bmatrix}$$

En este experimento se ha utilizado la función *calcular_g_Asym_1*.

Restricción local	P(1) Asimétrico 
Norma	L ₂
Normalización	Media y desv. típica
EER_FA (%)	0,35
EER_FE (%)	4,65
Tiempo (s)	4498,69

Tabla 6.4.1.6: Resultados experimento P(1) Asimétrico.

En la Tabla 6.4.1.6 se muestran los resultados para el caso Asimétrico del P(1). Se ha obtenido un valor de EER de 0,35% para las firmas falsas aleatorias y un valor de 4,65%

para las entrenadas. El equierror de las aleatorias es el 7,5% del de las entrenadas. Si se compara con el caso simétrico, se aprecia un incremento del error en las entrenadas (aproximadamente un 2% más) y un decremento en las aleatorias (en torno al 24% menos).

Si se analiza el tiempo medio de procesamiento, en este caso, se obtiene que una firma de un usuario consume un tiempo de 0,9 segundos (4498,69 segundos para toda la base de datos), un valor muy parecido al del caso simétrico. Debido a la escasa diferencia de tiempo, a que el aumento del equierror en las entrenadas es bajo y a que la reducción del equierror en las aleatorias es considerable, cuando se pueda elegir entre el caso simétrico y asimétrico, se debería optar por el asimétrico, puesto que en general proporciona mejores resultados.

En la Figura 6.4.1.6 se representan las gráficas para este caso y se puede concluir que son muy similares a las del caso simétrico.

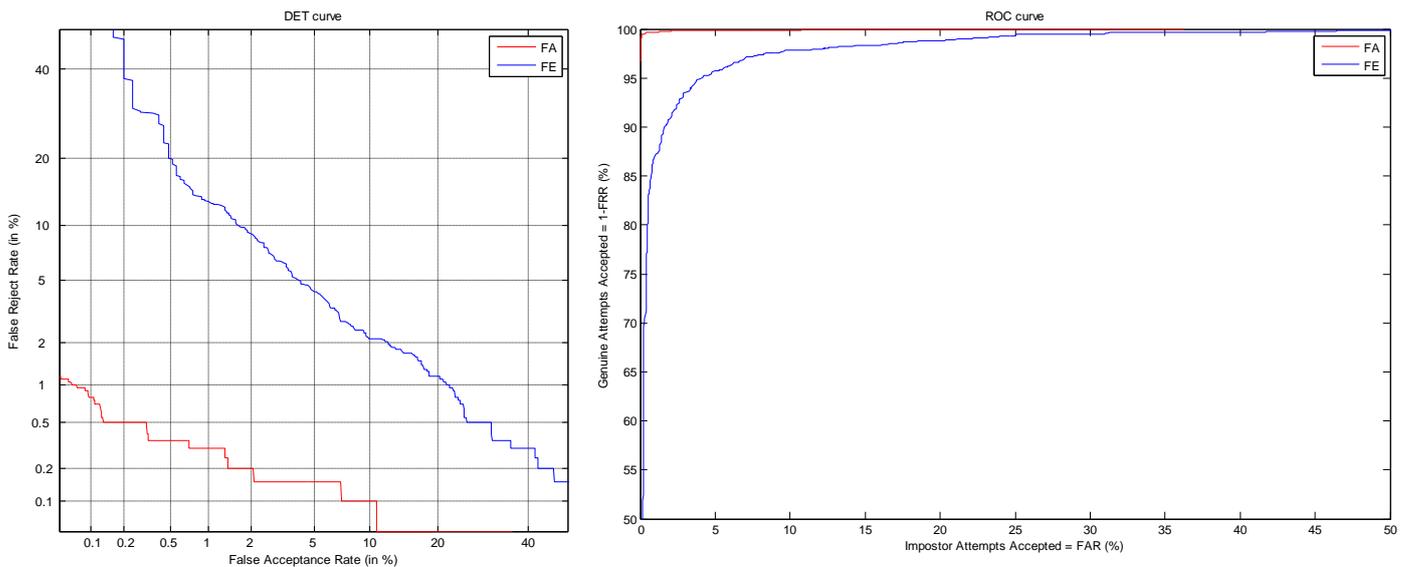
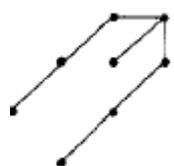


Figura 6.4.1.6: Curva DET y ROC para P(1) Asimétrico.

◆ Restricción Local 2 Simétrica:



$$\min \begin{bmatrix} g(i-2, j-3) + 2d(i-1, j-2) + 2d(i, j-1) + d(i, j) \\ g(i-1, j-1) + 2d(i, j) \\ g(i-3, j-2) + 2d(i-2, j-1) + 2d(i-1, j) + d(i, j) \end{bmatrix}$$

En este experimento se ha utilizado la función *calcular_g_Sym_2*.

Restricción local	P(2) Simétrico
Norma	L_2
Normalización	Media y desv. típica
EER_FA (%)	0,82
EER_FE (%)	5,6
Tiempo (s)	4847,84

Tabla 6.4.1.7: Resultados experimento P(2) Simétrico.

En la Tabla 6.4.1.7 se resumen los resultados para el caso P(2) Simétrico. Para el caso de las firmas falsas aleatorias se ha obtenido un equierror de 0,82% y para las entrenadas de 5,6%. La tasa de equierror de las entrenadas es casi 7 veces la de las aleatorias, es decir, el EER de las aleatorias es menos del 7% del equierror de las entrenadas.

El coste computacional total es 4847,84, lo cual se traduce en 0,97 segundos para una sola firma.

En la Figura 6.4.1.7 se presentan las curvas DET y ROC para este caso particular. Como se ha visto los valores de los EER son ligeramente superiores a casos anteriores, lo cual también se traduce en las gráficas. La curva DET se aleja un poco del origen de coordenadas y la ROC del 100. En este caso se obtienen más falsas aceptaciones y falsos rechazos.

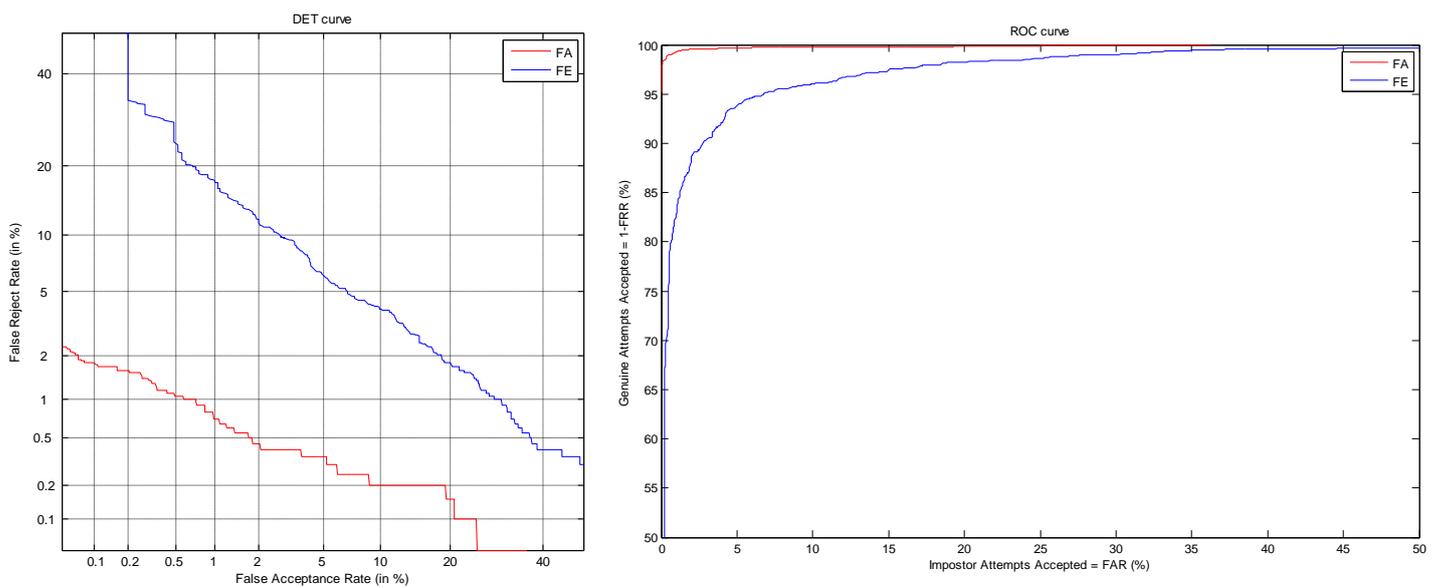
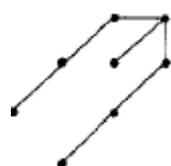


Figura 6.4.1.7: Curva DET y ROC para P(2) Simétrico.

◆ Restricción Local 2 Asimétrica:



$$\min \begin{bmatrix} g(i-2, j-3) + 2((i-1, j-2) + d(i, j-1) + d(i, j))/3 \\ g(i-1, j-1) + d(i, j) \\ g(i-3, j-2) + d(i-2, j-1) + d(i-1, j) + d(i, j) \end{bmatrix}$$

En este experimento se ha utilizado la función *calcular_g_Asym_2*.

Restricción local	P(2) Asimétrico 
Norma	L ₂
Normalización	Media y desv. típica
EER_FA (%)	0,65
EER_FE (%)	5,4
Tiempo (s)	4917,96

Tabla 6.4.1.8: Resultados experimento P(2) Asimétrico.

En la Tabla 6.4.1.8 se agrupan los resultados para el último experimento del primer conjunto, el caso P(2) Asimétrico. El EER de las firmas falsas entrenadas es 5,4% y el de las aleatorias 0,65%, es decir, el 12% del equierror de las entrenadas. Comparándolo con el caso simétrico se observa una reducción de casi el 21% en las aleatorias y del 3,6% en las entrenadas, aquí en ambos caso es mejor el asimétrico puesto que se obtienen tasas de equierror más bajas.

El tiempo medio de procesamiento de una firma es 0,98 segundos (4917,96 segundos en total para la base de datos completa) frente a los 0,97 segundos del caso simétrico, esta leve variación hace que compense el uso del caso asimétrico frente al simétrico para el camino P(2).

En la Figura 6.4.1.8 se representan las curvas DET y ROC para este caso, resultan similares a las del experimento simétrico.

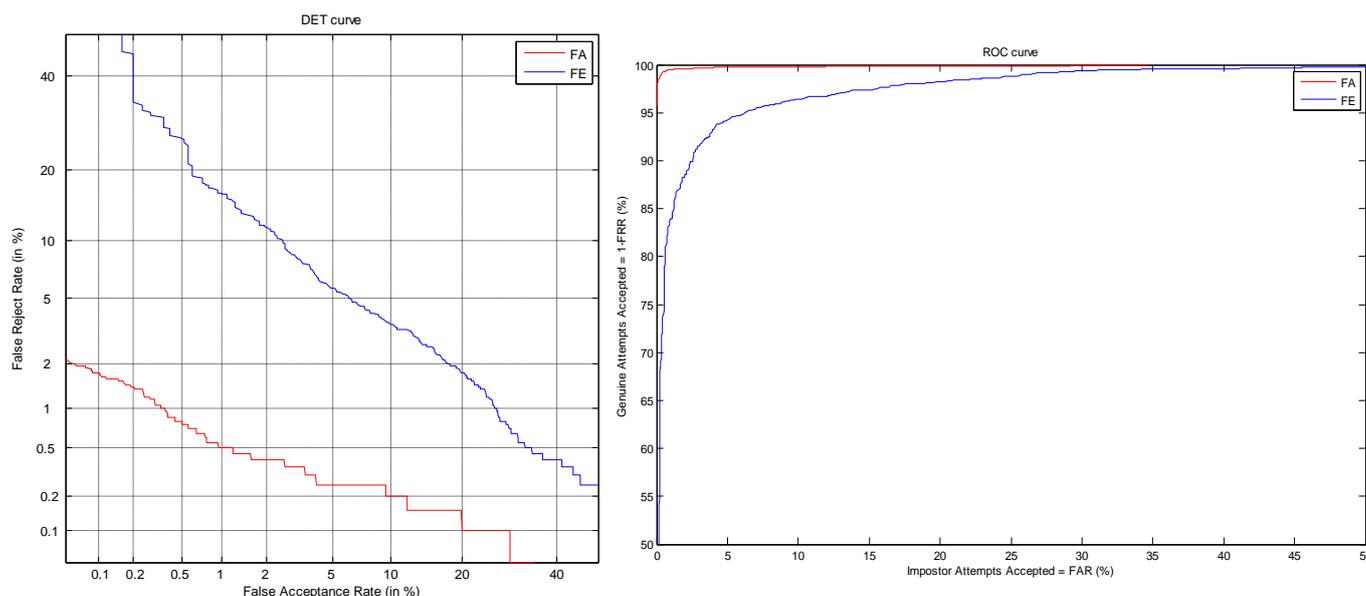


Figura 6.4.1.8: Curva DET y ROC para P(2) Asimétrico.

Como conclusión general del análisis de los experimentos de las restricciones locales, se deduce que salvo para el camino P(0), en el cual el caso simétrico presenta mejores resultados que el asimétrico, para el resto de restricciones el caso asimétrico es más óptimo en rasgos generales, puesto que la tasa de equierror de las falsas aleatorias o entrenadas es menor y la diferencia de tiempos no es significativa puesto que suele variar en torno a alguna centésima de segundo.

6.4.2- 2º Conjunto de experimentos: Tipos de distancias

Una vez concluidos los experimentos cuyo objetivo era analizar los distintos tipos de restricciones locales, se ha decidido reducir el área de actuación para el segundo conjunto de experimentos. De este modo, se han elegido dos restricciones locales para la realización de las próximas pruebas, por un lado se utilizará la restricción P(0) Simétrica, ya que con ella se ha obtenido el menor EER (3,63%) para las firmas falsas entrenadas. La otra restricción local que se ha elegido es la P(1/2) Asimétrica, puesto que tiene una tasa de equierror de 0,31%, la cual es la menor que se ha obtenido para las falsas aleatorias en la totalidad de los experimentos realizados hasta el momento.

Ante esta perspectiva, con la segunda modificación se busca evaluar si al cambiar el tipo de distancia empleada en el cálculo de la matriz de distancias [ver 4.1.- *Alineamiento Temporal Dinámico - DTW*], se puede reducir la tasa de equierror y cómo afecta al coste computacional.

Las distintas distancias que se han utilizado en este proyecto se explican a continuación [31]:

◆ Norma L_1 : Conocida también como distancia Taxicab o Manhattan.

La distancia Taxicab o Manhattan entre dos vectores, en un espacio vectorial real n-dimensional con un sistema de coordenadas cartesianas fijo, es la suma de las longitudes de las proyecciones del segmento de línea entre los puntos sobre el sistema de ejes coordenados.

Su ecuación matemática se muestra en la Ecuación [1]:

$$d_1(\mathbf{p}, \mathbf{q}) = \|\mathbf{p} - \mathbf{q}\|_1 = \sum_{i=1}^n |p_i - q_i|, \quad \text{Ec. [1]}$$

◆ Norma L_2 : Conocida también como distancia euclídea.

Es la distancia entre dos puntos de un espacio euclídeo, la cual se deduce a partir del teorema de Pitágoras.

Su ecuación matemática se muestra en la Ecuación [2]:

$$d_E(P, Q) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \cdots + (p_n - q_n)^2} = \sqrt{\sum_{i=1}^n (p_i - q_i)^2}. \quad \text{Ec. [2]}$$

◆ Norma $L_{1/2}$:

Norma fractal en la que el exponente es $\frac{1}{2}$.

Su ecuación matemática se muestra en la Ecuación [3]:

$$\|\mathbf{x}\|_p := \left(\sum_{i=1}^n |x_i|^p \right)^{1/p}. \quad \text{Ec. [3]}$$

En función de la distancia se han realizado 4 experimentos, 2 con la distancia L_1 y otros 2 con la $L_{1/2}$, puesto que la L_2 es la que se ha utilizado en todos los experimentos del primer conjunto de experimentos.

◆ Restricción Local 0 Simétrica y distancia L_1 :

Restricción local	P(0) Simétrico ↗
Norma	L_1
Normalización	Media y desv. típica
EER_FA (%)	0,3
EER_FE (%)	3,71
Tiempo (s)	3650,38

Tabla 6.4.2.1: Resultados experimento P(0) Simétrico y L_1 .

En la Tabla 6.4.2.1 se resumen los resultados obtenidos para este caso, si se comparan con los obtenidos en el experimento en el que se utilizó esta misma restricción local pero la norma L_2 , se observa que el EER para las firmas falsas aleatorias se ha reducido de un 0,35% a 0,3%, es decir, un 14% aproximadamente. Para las falsas entrenadas se ha pasado de tener un error de 3,63% a 3,71%, se ha visto incrementado en un 2%. El tiempo de procesamiento promedio para una firma ha pasado a ser 0,73 segundos, frente a 0,8 segundos del primer experimento.

Como conclusión se extrae que si se tiene que elegir entre usar el P(0) Simétrico con la norma L_2 o la misma restricción con la norma L_1 , en términos generales, compensa el uso de la norma L_1 , puesto que se reduce en un 14% los falsos aleatorios, lo que proporciona mayor seguridad al sistema. Además, también existe una diferencia de tiempo, que aunque sea pequeña en el análisis de una firma en particular, en el análisis de la base de datos completa supone más de 12 minutos. Este resultado es lógico, puesto que el tiempo de procesamiento es mayor cuanto más compleja es una operación.

En la Figura 6.4.2.1 se representan las curvas DET y ROC para este experimento, se observa que resultan similares a las del experimento con L_2 .

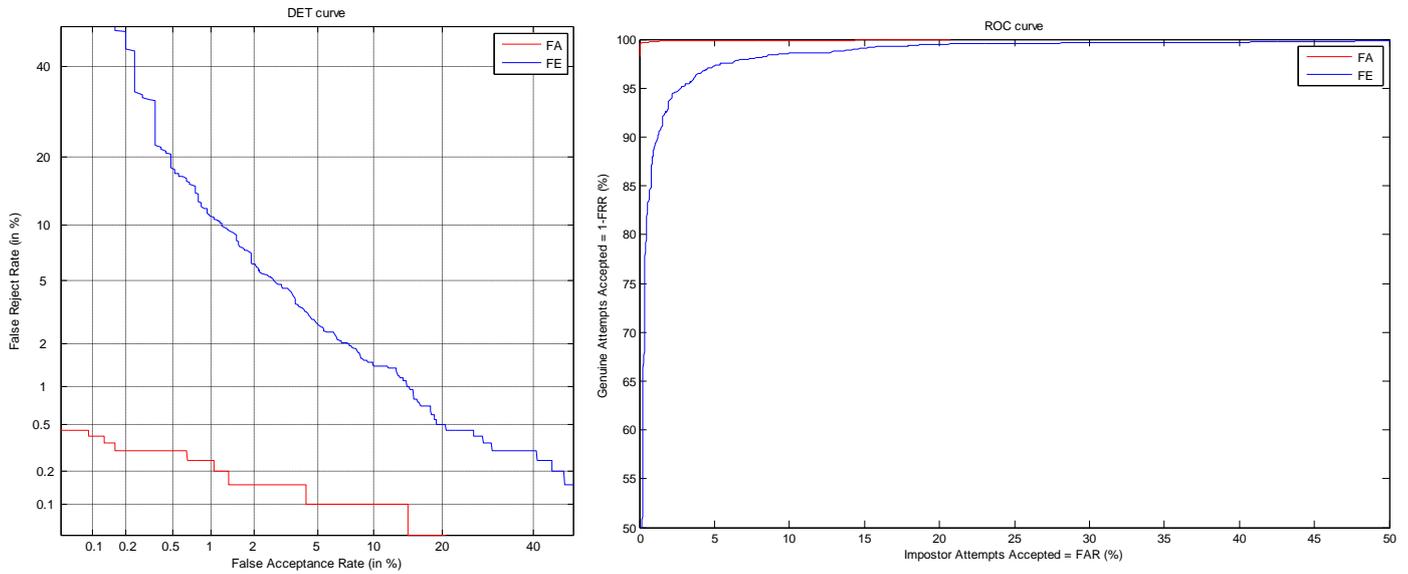


Figura 6.4.2.1: Curva DET y ROC para P(0) Simétrico y L_1 .

◆ Restricción Local 1/2 Asimétrica y distancia L_1 :

Restricción local	P(1/2) Asimétrico
Norma	L_1
Normalización	Media y desv. típica
EER_FA (%)	0,3
EER_FE (%)	4
Tiempo (s)	4469,33

Tabla 6.4.2.2: Resultados experimento P(1/2) Asimétrico y L_1 .

En la Tabla 6.4.2.2 se recogen los resultados de este experimento. Si se comparan con los obtenidos en el experimento de la restricción P(1/2) Asimétrica utilizando la distancia L_2 , se observa que el equierror de las firmas falsas aleatorias se ha reducido de 0,31% a 0,3% y el de las entrenadas ha aumentado de 3,99% a 4%. Son valores muy similares, cuyas pequeñas variaciones son casi inapreciables. El tiempo de procesamiento medio disminuye ligeramente, pasando de 1,02 segundos por firma a 0,9 segundos. También es una variación muy pequeña.

Se pueden concluir que, en general, para este conjunto de experimentos se aprecian ciertos beneficios al aplicar la distancia L_1 . La tasa de equierror para las falsas aleatorias se reduce y el coste computacional también, por ello el utilizar esta norma

para un escenario casual es interesante, puesto que se puede priorizar el coste computacional. Por el contrario, si se desea un escenario seguro, se podría utilizar la norma L_2 , puesto que la tasa de equierror para las falsas entrenadas es ligeramente menor.

En la Figura 6.4.2.2 se muestran las curvas DET y ROC para este experimento, son similares a las del experimento con L_2 .

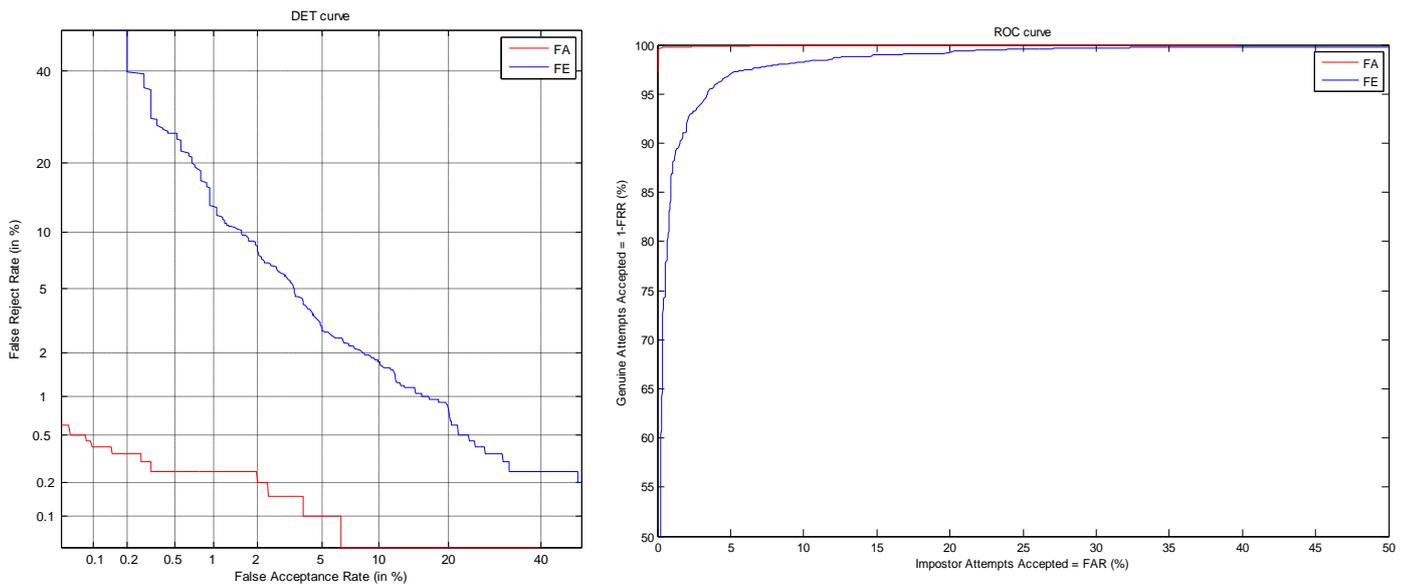


Figura 6.4.2.2: Curva DET y ROC para P(1/2) Asimétrico y L_1 .

◆ Restricción Local 0 Simétrica y distancia $L_{1/2}$:

Restricción local	P(0) Simétrico ↗
Norma	$L_{1/2}$
Normalización	Media y desv. típica
EER_FA (%)	0,3
EER_FE (%)	4,11
Tiempo (s)	14476,73

Tabla 6.4.2.3: Resultados experimento P(0) Simétrico y $L_{1/2}$.

En la Tabla 6.4.2.3 se resumen los resultados que se han obtenido con la norma $L_{1/2}$. Al compararlo con los obtenidos en el experimento en el que se utilizó esta misma restricción local pero la norma L_2 , se observa que el EER para las firmas falsas aleatorias se ha reducido de 0,35% a 0,3% y si se compara con el obtenido

con la norma L_1 vemos que es el mismo valor. Por otro lado, para las falsas entrenadas se ha pasado de tener un equierror de 3,63% a 4,11%, se ha visto incrementado en un 12% aproximadamente. Si se compara también con el EER de la norma L_1 , se observa que también se incrementa, esta vez 9,7%. El coste computacional promedio para una firma pasa de ser 0,8 segundos para la L_2 , a casi 3 segundos con $L_{1/2}$. Si se compara con el tiempo medio obtenido con L_1 , al ser aún menor que el obtenido con L_2 , existe aún más diferencia con el de $L_{1/2}$. Por lo tanto, el coste computacional se ha incrementado aproximadamente un 75%.

Puesto que no supone una mejora en la tasa de equierror y se produce un altísimo incremento del coste computacional se puede concluir que es más conveniente, en general, el uso de la norma L_1 .

En la Figura 6.4.2.3 se muestra gráficamente las curvas DET y ROC para el experimento P(0) Simétrico con la norma $L_{1/2}$.

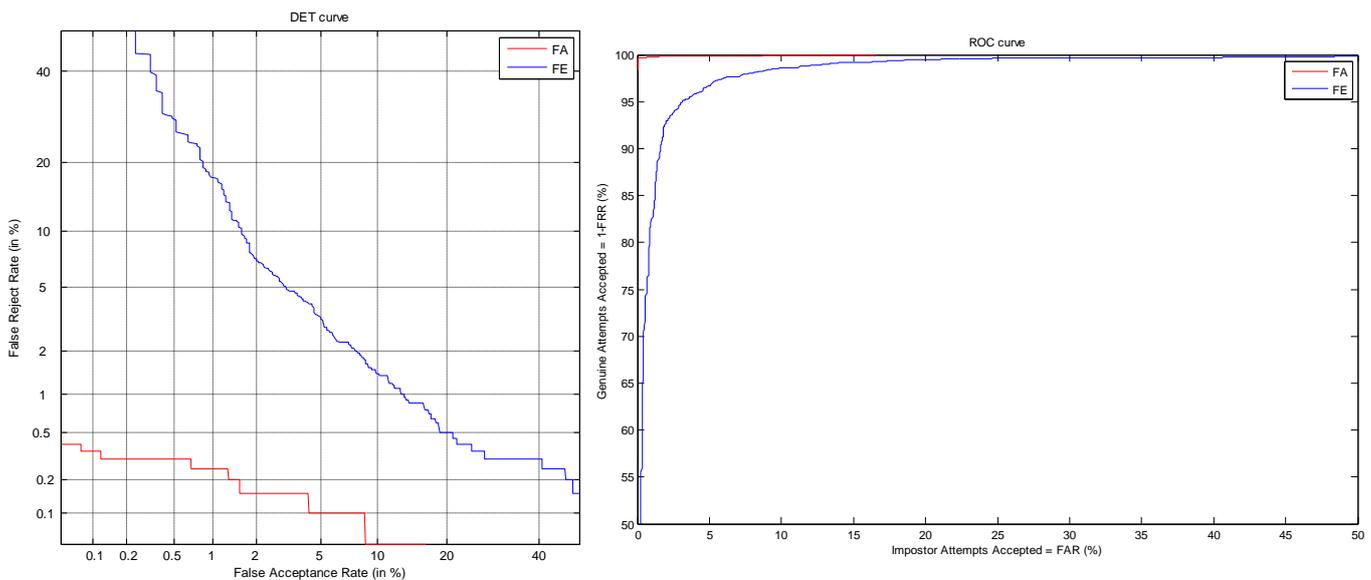


Figura 6.4.2.3: Curva DET y ROC para P(0) simétrico y $L_{1/2}$.

◆ Restricción Local 1/2 Asimétrica y distancia $L_{1/2}$:

Restricción local	P(1/2) Asimétrico
Norma	$L_{1/2}$
Normalización	Media y desv. típica
EER_FA (%)	0,29
EER_FE (%)	4,38
Tiempo (s)	15476,5

Tabla 6.4.2.4: Resultados experimento P(1/2) Asimétrico y $L_{1/2}$.

En la Tabla 6.4.2.4 se recogen los resultados que se han obtenido con la norma $L_{1/2}$ para este experimento. Comparando con la obtenida en el experimento en el que se utilizó esta misma restricción local pero la norma L_2 , se observa que el equierror para las falsas aleatorias se ha reducido de 0,31% a 0,29% y si se compara con el obtenido con la norma L_1 se observa que se ha reducido en 0,01%, por lo tanto entre L_2 , L_1 y $L_{1/2}$ permanece prácticamente constante puesto que varía muy poco. Por otro lado, se ha visto que para las falsas entrenadas se ha pasado de tener un error de 3,99% a 4,38%, lo que supone un incrementado en un 9%. Si se compara con la norma L_1 , se observa que se obtiene un valor prácticamente igual. Además, el coste computacional se ha incrementado considerablemente, ya que se ha pasado de tener 1,02 segundos por firma, con la norma L_2 , a 3,1 segundos con la $L_{1/2}$. Si se compara con la L_1 el incremento es ligeramente superior.

Tras analizar estos datos se puede concluir que el uso de la norma $L_{1/2}$ no mejora las prestaciones, pues no supone una reducción del error significativa y el coste computacional aumenta demasiado.

En la Figura 6.4.2.4 se muestra gráficamente las curvas DET y ROC para el experimento P(1/2) Asimétrico y la norma $L_{1/2}$.

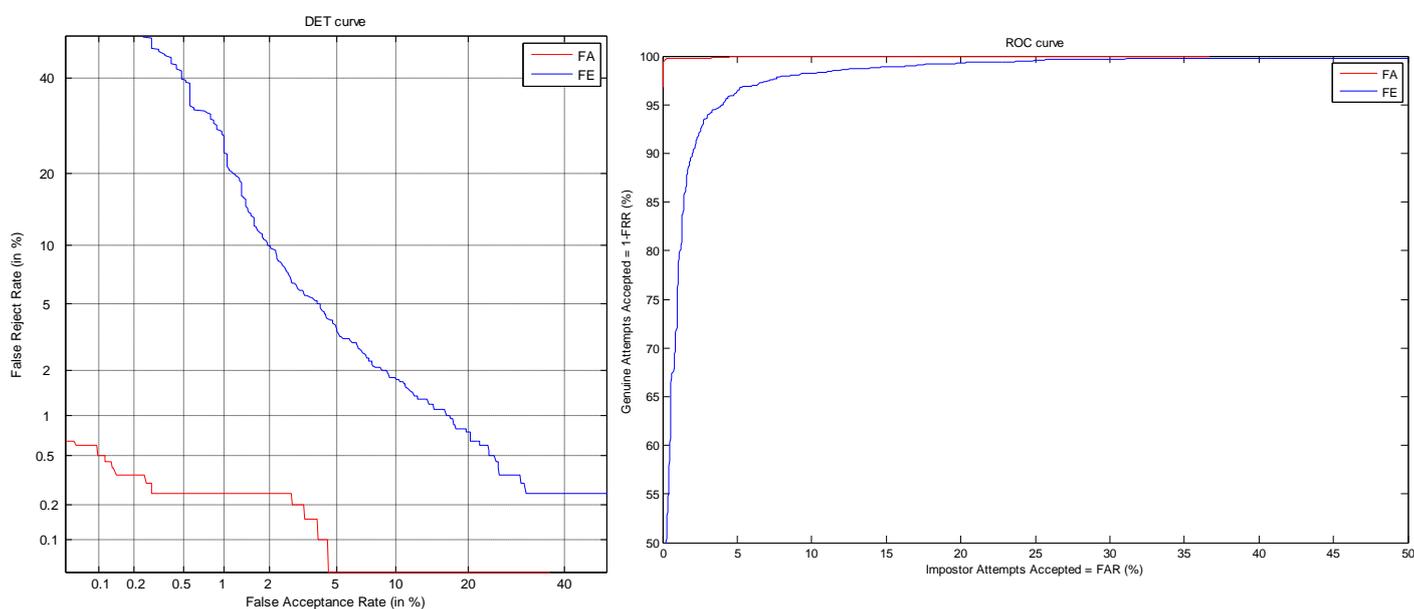


Figura 6.4.2.4: Curva DET y ROC para P(1/2) Asimétrico y $L_{1/2}$.

Una vez concluidos los conjuntos de experimentos 1 y 2, se pueden obtener ciertas conclusiones como que, en general, la distancia L_1 es la más recomendable puesto que presenta ciertas mejoras frente a las otras distancias (menor tasa de equierror en general para las falsas aleatorias (EER de 0,3%) y menor tiempo de procesamiento

medio (0,73 segundos por firma)). A pesar de ello, la menor tasa de equierror se ha obtenido hasta ahora, para las falsas entrenadas, con la restricción local $P(0)$ Simétrica y la norma L_2 (EER de 3,63%). Mientras, para las falsas aleatorias se ha obtenido con la restricción $P(1/2)$ Asimétrica y la norma $L_{1/2}$ (EER de 0,29%). No obstante, como se ha mencionado antes si se analizan en profundidad los resultados obtenidos y se mantiene un compromiso entre las tasas de equierror y el coste computacional, hasta el momento la mejor opción sería la norma L_1 .

6.4.3- 3^{er} Conjunto de experimentos: Tipos de normalizaciones

Por otro lado se quiso observar cómo afectaba variar el tipo de normalización que se aplicaba a los datos, para ello se decidió usar distintas normalizaciones y volver a probar el algoritmo.

Como se mencionó anteriormente en los experimentos 1 y 2, se utilizaron para normalizar la media y la desviación típica. Esto se ha hecho así para controlar las muestras y poder evaluarlas todas, independientemente de su posición y su tamaño.

Para este tercer conjunto de experimentos, se han utilizado las conclusiones de las dos primeras modificaciones, se ha tomado la restricción local óptima y la distancia óptima, considerando como óptimo el caso con menor tasa de equierror y teniendo en cuenta el coste computacional. Por tanto, siendo rigurosos y aunque algunas diferencias al usar unos parámetros u otros no hayan sido muy significativas, el escenario de ensayo de menor tasa de equierror ha sido, por un lado, la restricción $P(1/2)$ Asimétrica con la norma $L_{1/2}$, y por otro lado se ha probado la restricción $P(0)$ Simétrica con la norma L_1 . A pesar de obtenerse una tasa de equierror menor con la norma L_2 , se ha decidido usar la norma L_1 en los siguientes experimentos, por la reducción significativa del coste computacional.

Ante esta perspectiva, se ha probado a cambiar la normalización usando en este caso el máximo y el mínimo de los datos procesados. Los datos han sido normalizados restando el mínimo y dividiendo entre la diferencia del máximo menos el mínimo.

◆ Restricción Local 0 Simétrica y distancia L_1 :

Restricción local	P(0) Simétrico ↗
Norma	L_1
Normalización	Máximo y mínimo
EER_FA (%)	3
EER_FE (%)	9,74
Tiempo (s)	3724,3

Tabla 6.4.3.1: Resultados experimento P(0) Simétrico, L_1 y normalización de máximo y mínimo.

En la Tabla 6.4.3.1 se observan los resultados obtenidos al utilizar la restricción local P(0) Simétrica, la norma L_1 y la normalización de la diferencia del máximo y el mínimo. Comparando estos datos con los obtenidos en pruebas anteriores se aprecia un claro aumento de las tasas de equierror tanto para las firmas falsas entrenadas como para las aleatorias. En concreto si se compara con los valores obtenidos con el mismo camino y norma pero con la normalización de la media y la desviación típica, se observa que se ha producido un incremento muy significativo, se ha pasado de tener un equierror de 0,3% a 3% (diez veces más), para las aleatorias. Mientras que para las entrenadas de 3,71% a un EER cercano a 10%. En cuanto al tiempo computacional, solo mencionar que es ligeramente superior. Por lo tanto, se puede decir que al menos para este caso es mejor usar la normalización de la media y la desviación típica, puesto que se obtienen menores tasas de equierror.

En la Figura 6.4.3.1 se presentan las curvas DET y ROC obtenidas con el camino P(0) Simétrico, la norma L_1 y la normalización del máximo y el mínimo. Se aprecia que las curvas son coherentes con los resultados obtenidos, puesto que la DET se aleja del punto (0,0) y la ROC del (0,100), es decir, las curvas se alejan de sus puntos óptimos, por lo tanto el sistema es peor que en los casos anteriores.

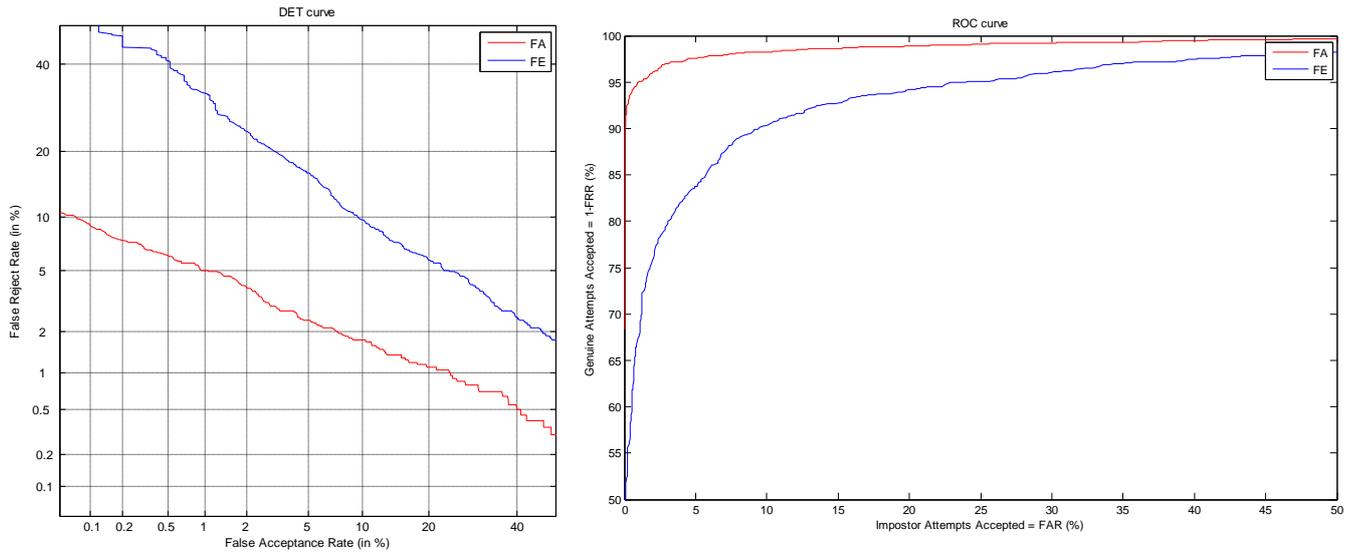


Figura 6.4.3.1: Curva DET y ROC para P(0) simétrico, L_1 y normalización de máximo y mínimo.

◆ Restricción Local 1/2 Asimétrica y distancia $L_{1/2}$:

	P(1/2)	
Restricción local	Asimétrico	
Norma	$L_{1/2}$	
Normalización	Máximo y mínimo	
EER_FA (%)	2,15	
EER_FE (%)	8,38	
Tiempo (s)	15500,7	

Tabla 6.4.3.2: Resultados experimento P(1/2) Asimétrico, $L_{1/2}$ y normalización de máximo y mínimo.

En la Tabla 6.4.3.2 se resumen las tasas de equierror y el tiempo computacional para el caso en el que se usa la restricción local P(1/2) Asimétrica, la norma $L_{1/2}$ y la normalización de máximo y mínimo. Si se compara con el caso en el que se usa el mismo camino y distancia pero la normalización de la media y la desviación típica, se observa que se ha incrementado notablemente la tasa de equierror y en menor medida el coste computacional. El EER para las falsas aleatorias ha aumentado más del 80% en comparación con el otro caso, y para las falsas entrenadas se obtiene un valor de más del doble. Ante estos malos resultados se puede concluir que es mejor utilizar la normalización de la media y la desviación típica.

En la Figura 6.4.3.2 mostramos las curvas DET y ROC obtenidas con el camino $P(1/2)$ Asimétrico, $L_{1/2}$ y normalización de máximo y mínimo.

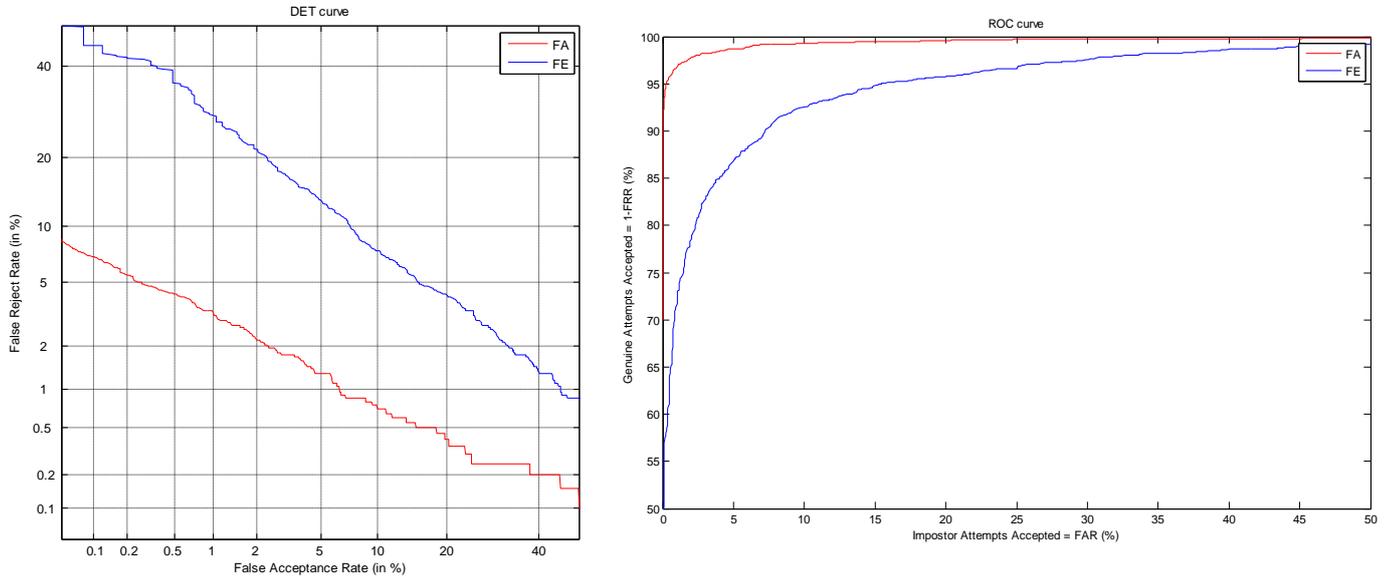


Figura 6.4.3.2: Curva DET y ROC para $P(1/2)$ Asimétrico, $L_{1/2}$ y normalización de máximo y mínimo.

Una vez finalizado el análisis de estos experimentos en base a la normalización, para ambos casos se ha concluido que se obtienen mejores resultados con la normalización de la media y la desviación típica.

Capítulo 7:

Conclusiones y trabajos futuros

7.1- Introducción

En este capítulo se exponen las conclusiones a las que se ha llegado tras realizar este proyecto, además las posibles mejoras o trabajos futuros que podrían realizarse a partir de este proyecto.

7.2- Conclusiones

A lo largo de este documento se ha ido explicando el proceso que se ha llevado a cabo para la realización de este proyecto, cómo se ha estudiado en profundidad el algoritmo DTW, cómo se le han aplicado las modificaciones necesarias para realizar los experimentos, así como el desarrollo de los experimentos que se han hecho y las conclusiones a las que se ha llegado con éstos.

Con el trabajo realizado en este proyecto se ha comprobado que la firma manuscrita dinámica es una modalidad biométrica que puede ser usada para reconocer personas de forma eficiente. A partir de los experimentos realizados en este proyecto, los cuales han sido presentados en el capítulo 6, se han obtenido unas tasas de equierror que se encuentran en general en torno al 3-5% para firmas falsas entrenadas y valores entre 0-1% con firmas falsas aleatorias. Los resultados obtenidos son coherentes ya que se han obtenido tasas de equierror mayores al analizar las firmas falsas entrenadas que en las firmas falsas aleatorias. Estos resultados tienen su lógica ya que las firmas falsas entrenadas han tenido un ensayo previo, es decir, el usuario ha practicado la falsificación de la firma, por lo que la tasa de equierror debe ser mayor que si se compara con la tasa de equierror obtenida con las firmas falsas aleatorias, las cuales son firmas de otros usuarios.

A partir de las conclusiones de los experimentos expuestos en el capítulo 6, se concluye, en general, que el DTW óptimo tendría la combinación de características con menor tasa de equierror y menor coste computacional, pero manteniendo un compromiso. Dicha combinación es la siguiente:

- ▶ **1^{er} Conjunto de experimentos:** Restricción local $P(0)$ Simétrica.
- ▶ **2^o Conjunto de experimentos:** Norma L_1 .
- ▶ **3^{er} Conjunto de experimentos:** Normalización usando la media y la desviación típica.

Con esta configuración se ha obtenido una tasa de equierror para firmas falsas entrenadas de 3,71%, para las falsas aleatorias 0,3% y un tiempo de procesamiento de 0,73 segundos por firma.

Se ha elegido esta configuración del algoritmo porque en general presenta los resultados más beneficiosos frente al resto de configuraciones. A pesar de ello, la menor tasa de equierror se ha obtenido para las falsas entrenadas con la restricción local $P(0)$ Simétrica y la norma L_2 , mientras que para las falsas aleatorias se ha obtenido con la restricción $P(1/2)$ Asimétrica y la norma $L_{1/2}$.

Si el objetivo es el escenario más seguro posible, se debe elegir la configuración que presenta la menor tasa de equierror para las firmas falsas entrenadas, en este caso se tendría que utilizar la restricción local $P(0)$ Simétrica y la norma L_2 , sin embargo, analizando el resto de resultados, es decir, la tasa de equierror de las falsas aleatorias y el coste computacional, podría ser una buena alternativa el uso de la restricción local $P(0)$ Simétrica y la norma L_1 , puesto que solo supone un aumento del 2% de la tasa de equierror de las falsas entrenadas y presenta una reducción significativa del coste computacional.

Por el contrario, si el objetivo es un escenario casual, lo prioritario es el coste computacional, por lo que la mejor opción sería la restricción local $P(0)$ Simétrica y la norma L_1 , puesto que con ésta se obtiene el menor tiempo de procesamiento.

Por último, si se compara el experimento con el que, en general, se han obtenido los mejores resultados y el experimento con los peores resultados, se tiene:

- ▶ Experimento con los mejores resultados en cuanto a tasa de equierror y coste computacional:
 - Restricción local $P(0)$ Simétrica.
 - Norma L_1 .
 - Normalización usando la media y la desviación típica.

EER (FE) = 3,71%**EER (FA) = 0,3%****TIEMPO = 3650,38 s**

Tabla 7.2.1: Resumen del experimento con los mejores resultados en cuanto a tasa de equierror y coste computacional.

- ▶ Experimento con los peores resultados en cuanto a tasa de equierror:
 - Restricción local $P(0)$ Simétrica.
 - Norma L_1 .
 - Normalización usando el máximo y el mínimo.

EER (FE) = 9,74%

EER (FA) = 3%

TIEMPO = 3724,3 s

Tabla 7.2.2: Resumen del experimento con los peores resultados en cuanto a tasa de equierror.

- ▶ Experimento con los peores resultados en cuanto a coste computacional:
 - Restricción local $P(1/2)$ Asimétrica.
 - Norma $L_{1/2}$.
 - Normalización usando el máximo y el mínimo.

EER (FE) = 8,38%

EER (FA) = 2,15%

TIEMPO = 15500,7s

Tabla 7.2.3: Resumen del experimento con los peores resultados en cuanto a coste computacional.

Por lo tanto, se ha conseguido una mejora significativa, en general, se ha reducido aproximadamente el 90% la tasa de equierror de las falsas aleatorias, el 62% para las falsas entrenadas y en cuanto al coste computacional, se ha reducido un 76%, es decir se ha reducido unos 11850 segundos (para la base de datos completa).

Finalmente cabe mencionar que se han cumplido los objetivos del proyecto:

- ◆ Subobjetivo I: Estudio e implementación del algoritmo DTW. ✓
- ◆ Subobjetivo II: Análisis de ciertas características para optimizar el reconocimiento de firmas manuscritas on-line, en función del coste computacional y la tasa de equierror. A partir de este análisis se han evaluado los resultados y se han obtenido las conclusiones pertinentes. ✓

7.3- Trabajos futuros

Existen numerosas líneas de investigación para trabajos futuros en el mundo del reconocimiento biométrico mediante firma manuscrita, puesto que es un campo muy amplio y como se ha mencionado varias veces a lo largo del documento, está en auge en la sociedad de la información.

Una línea interesante podría ser realizar un experimento donde se aplicase este algoritmo DTW modificado en un entorno directo a modo de prueba, por ejemplo simulando un pago con tarjeta en una tienda.

Otro trabajo futuro podría ser utilizar distintas bases de datos, como puede ser las SVG2004 [23] o MylDea [23]. Estas tienen características distintas por lo que el estudio podría extenderse y conseguir así un algoritmo más genérico. En esta línea de investigación, podrían también utilizarse firmas de otras nacionalidades o alfabetos, como el chino o el árabe. Para ello, podría utilizarse, por ejemplo, la base de datos SVG2004 que contiene firmas en chino. De este modo, se procesarían y se extraerían las características de las firmas de forma que pudiese adaptarse al algoritmo, y por lo tanto al código, de este proyecto. Se conseguiría así un algoritmo más genérico y por lo tanto más eficaz.

Por otro lado, a este proyecto se le podrían aplicar ciertas mejoras, algunas de ellas son:

- ◆ Realizar cada experimento varias veces, por ejemplo 10, luego se tomaría la media y así se conseguirían resultados más fiables y genéricos.
- ◆ Analizar más modificaciones del algoritmo DTW, como podría ser variar el tipo de ventana, reduciendo así el área de cálculo, con el objetivo de conseguir posibles mejoras en coste computacional.
- ◆ Implementar todo el código en un lenguaje más genérico y eficiente, como podría ser en el lenguaje C.
- ◆ Se podría crear una interfaz de usuario que permitiese realizar los experimentos a cualquier persona que desconozca el algoritmo y además se podría desarrollar una pestaña en la que se pudiesen elegir las modificaciones de una lista y variar los valores de los parámetros de forma visual, sencilla y práctica.
- ◆ Por otro lado, podría ampliarse este estudio a otras modalidades biométricas y ver si se obtienen buenos resultados. En general, podría aplicarse a cualquier aplicación en la que se procesen series temporales, donde el algoritmo DTW sea de utilidad para medir distancias entre dichas series temporales.

ANEXOS

Anexo 1: Planificación

Anexo 1.1 Diagrama EDT

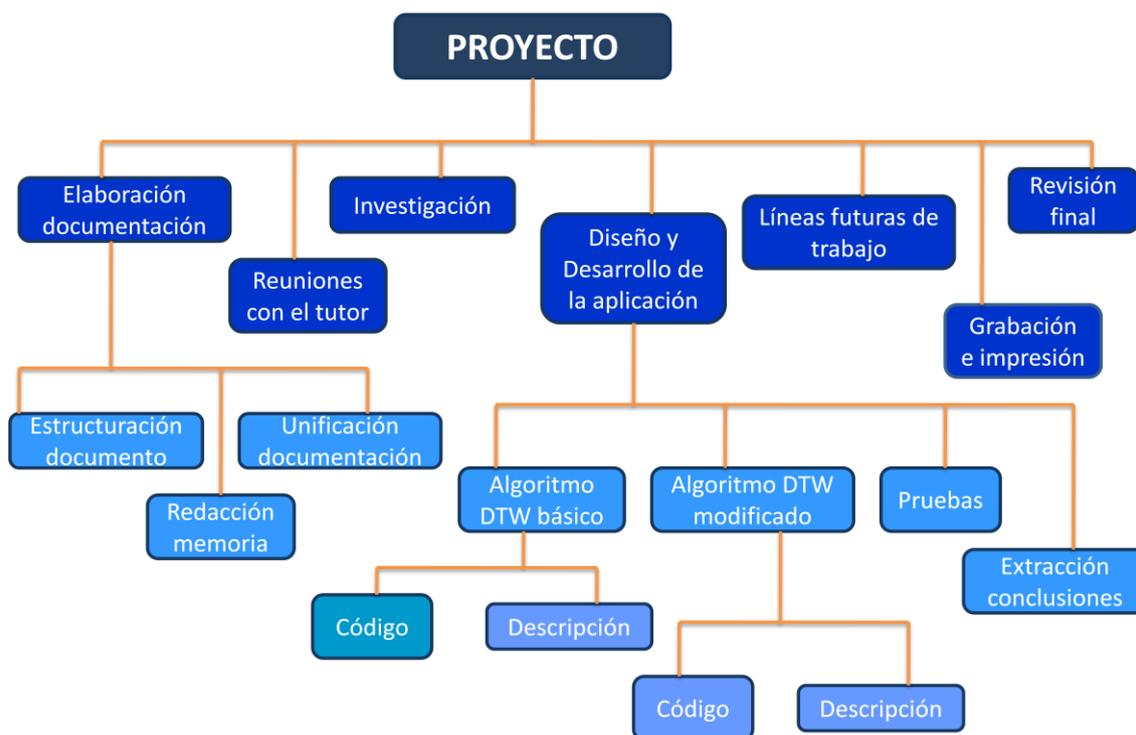


Figura A.1.1.1: Diagrama EDT.

Anexo 1.2 Diagrama PERT

Identificador de tarea	Tarea	Predecesora	Duración (días)	ES	EF	LS	LF	HT	HL
1	Estructuración documento	--	1	0	1	94	95	94	94
2	Unificación documentación	10	1	95	96	108	109	13	0
3	Reuniones con el tutor	--,1,2,5,9,7,8	1	96	97	109	110	13	0
4	Investigación	--	45	0	45	0	45	0	0
5	Algoritmo DTW básico código	4	15	45	60	45	60	0	0
6	Algoritmo DTW documentación	5,4	5	60	65	60	65	0	0
7	Algoritmo DTW modificado código	6,4	15	65	70	65	70	0	0
8	Algoritmo DTW modificado documentación	7,4	5	70	75	70	75	0	0
9	Extracción conclusiones	11	7	87	94	87	94	0	0
10	Líneas futuras	9	1	94	95	94	95	0	0
11	Pruebas	8,4	12	75	87	75	87	0	0
12	Grabación e impresión	13	3	114	117	114	117	0	0
13	Revisión final	10,3,14	4	110	114	110	114	0	0
14	Redacción memoria	-- ,1,4,6,8,9,10,11	15	95	110	95	110	0	0

Tabla A.1.2.1: Datos diagrama PERT.

ES _i	EF _i
I	D
LS _i	LF _i

$$ES_1=0 \quad EF_i = ES_i + d_i \quad ES_j = \max (EF_i + n)$$

$$LF_n = EF_n \quad LS_j = LF_j - d_j \quad LF_i = \min (LS_j - n)$$

$$HT = LS_i - ES_i \quad HL = \min (ES_j) - EF_i$$

Camino crítico → α-4-5-6-7-8-11-9-10-14-13-12

→ Camino crítico = duración mínima proyecto = 127 días

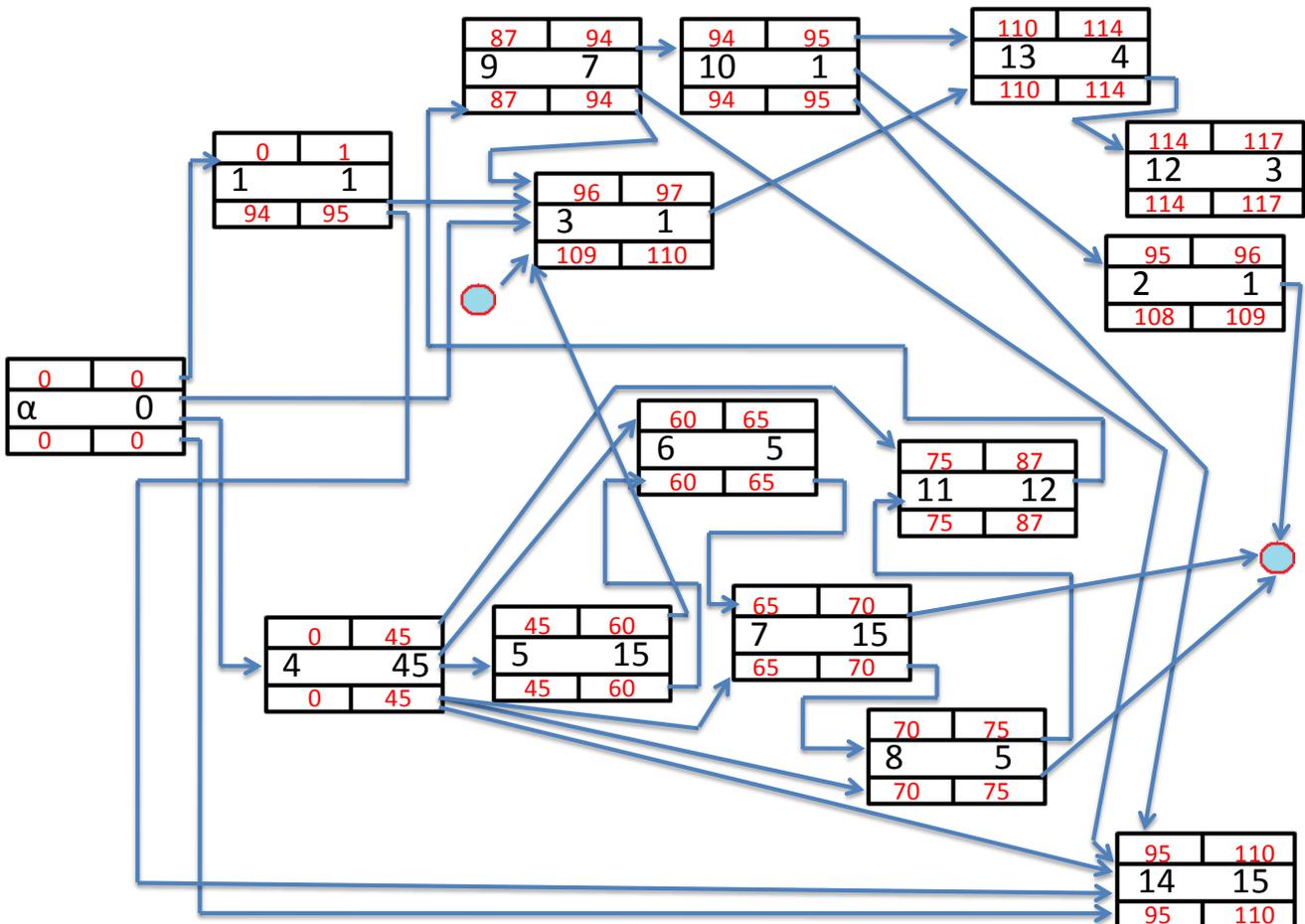


Figura A.1.2.1: Diagrama PERT.

Anexo 1.3 Planificación de tareas y diagrama GANTT

Para la planificación del proyecto se ha utilizado un software específico, el Microsoft Project 2010.

► Planificación de tareas:

Id	Nombre de tarea	Duración	Comienzo	Fin
1	Inicio de proyecto	0 días	lun 06/08/12	lun 06/08/12
2	Elaboración documentación	121 días	mié 15/02/12	lun 06/08/12
3	Estructuración del documento	1 día	mié 15/02/12	mié 15/02/12
4	Unificación del documento	1 día	lun 16/07/12	lun 16/07/12
5	Redacción memoria	15 días	mar 17/07/12	lun 06/08/12
6	Reuniones tutor	1 día	mar 07/08/12	mar 07/08/12
7	Investigación	45 días	mié 15/02/12	jue 19/04/12
8	Diseño y desarrollo aplicación	59 días	vie 20/04/12	jue 12/07/12
9	Algoritmo DTW básico	20 días	vie 20/04/12	vie 18/05/12
10	Programar algoritmo DTW básico	15 días	vie 20/04/12	vie 11/05/12
11	Algoritmo DTW básico documentación	5 días	lun 14/05/12	vie 18/05/12
12	Algoritmo DTW modificado	20 días	lun 21/05/12	vie 15/06/12
13	Programar algoritmo DTW modificado	15 días	lun 21/05/12	vie 08/06/12
14	Algoritmo DTW modificado documentación	5 días	lun 11/06/12	vie 15/06/12
15	Pruebas	12 días	lun 18/06/12	mar 03/07/12
16	Extracción de conclusiones	7 días	mié 04/07/12	jue 12/07/12
17	Líneas futuras	1 día	vie 13/07/12	vie 13/07/12
18	Impresión y grabación	3 días	mar 14/08/12	jue 16/08/12
19	Revisión final	4 días	mié 08/08/12	lun 13/08/12
20	Fin de proyecto	0 días	jue 16/08/12	jue 16/08/12

Tabla A.1.3.1: Planificación tareas.

► Diagrama GANTT:

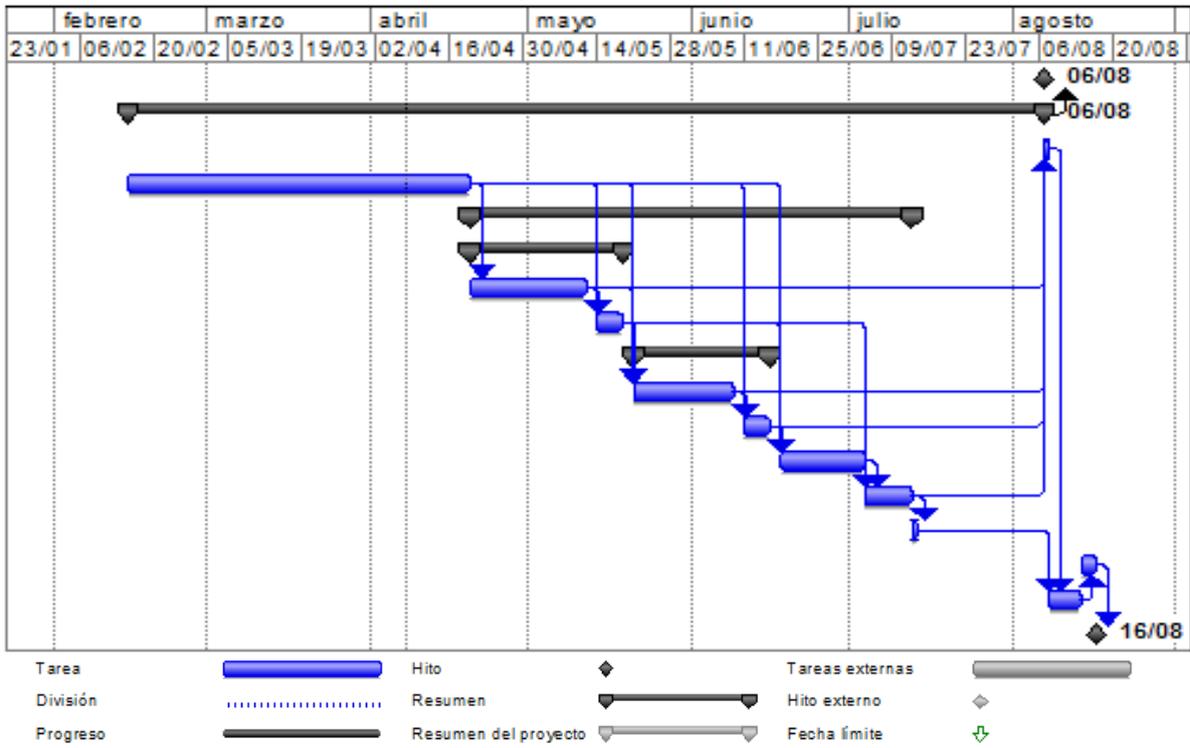


Figura A.1.3.1: Diagrama GANTT.

Anexo 2: Presupuesto

En este anexo se detalla el presupuesto del proyecto, el cual se divide en una serie de apartados en función de su naturaleza.

Anexo 2.1 Personal

En este apartado se incluyen los gastos relativos al personal que ha participado en el proyecto. En este caso, se ha contado con un Ingeniero de Sistemas Audiovisuales, trabajando a jornada completa (8h/días-40h/semanales). Además se cuenta con un director de proyecto, el cual actúa como consultor.

El coste total por personal asciende a: 11.000,00€.

Anexo 2.2 Equipos

En este apartado se consideran los costes relativos tanto al hardware como al software. Se ha tenido en cuenta la amortización.

El coste total por equipos asciende a: 461,78€.

Anexo 2.3 Subcontratación de tareas

Como tarea subcontratada se ha designado la grabación, impresión y encuadernación del proyecto.

El coste total por subcontratación de tareas asciende a: 120,00€.

Anexo 2.4 Otros costes directos del proyecto

En este apartado se ha contabilizado los costes por material de oficina, por electricidad, Internet y desplazamientos.

El coste total por otros costes directos asciende a: 450,00€.

Anexo 2.5 Costes indirectos

Se ha aplicado una tasa del 20% para el cálculo de los costes indirectos, que se calcula sobre el total de los costes directos.

El coste total por costes indirectos asciende a: 2.406,00€.

Anexo 2.6 Presupuesto desglosado del proyecto

El presupuesto estimado total es: 14.438,00€.

PRESUPUESTO DEL PROYECTO

1.- Autor: Patricia de los Ángeles Alonso Fuertes

2.- Departamento: Tecnología Electrónica

3.- Descripción del Proyecto:

Título: Análisis del algoritmo DTW para reconocimiento biométrico de personas mediante firma manuscrita on-line

Duración (meses): 6

Tasa de costes Indirectos: 20%

4.- Presupuesto total del Proyecto (€):

14.438 €

5.- Desglose presupuestario (costes directos):

PERSONAL

Apellidos y nombre	N.I.F. (no rellenar - solo a título informativo)	Categoría	Dedicación (hombres mes) ^{a)}	Coste hombre mes	Coste (Euros/total meses)	Firma de conformidad
Alonso Fuertes, Patricia de los Ángeles		Ingeniero	1	2.000,00	12.000,00	
Mengíbar Pozo, Luís		Director del proyecto	1		2.000,00	
Hombre mes 1					11.000,00	

a) 1 Hombre mes = 131,25 horas.

EQUIPOS

Descripción	Coste (Euro)	% Uso dedicado proyecto	Dedicación (meses)	Periodo de depreciación	Coste imputable ^{d)}
Ordenador sobre mesa HP Compaq Pro 3400	529,82	100	6	60	52,98
Monitor LCD 22" HP Compaq LA 2205wg	169,00	100	6	60	16,90
Microsoft Office 2010 (1 licencia)	219,00	100	6	60	21,90
Microsoft Project Estándar 2010 (1 licencia)	700,00	100	6	60	70,00
Matlab 2011 (1 licencia)	3.000,00	100	6	60	300,00
Base de datos firmas MCyT (versión libre)	0,00	100	6	60	0,00
Total					461,78

^{d)} Fórmula de cálculo de la Amortización:

$$\frac{A}{B} \times C \times D$$

A = nº de meses desde la fecha de facturación en que el equipo es utilizado

B = periodo de depreciación (60 meses)

C = coste del equipo (sin IVA)

D = % del uso que se dedica al proyecto (habitualmente 100%)

SUBCONTRATACIÓN DE TAREAS

Descripción	Empresa	Coste imputable
Impresión y grabación		120,00
Total		120,00

OTROS COSTES DIRECTOS DEL PROYECTO

Descripción	Empresa	Costes imputable
Material de oficina		50,00
Electricidad, Internet, transporte.		400,00
Total		450,00

6.- Resumen de costes:

Presupuesto Costes Totales	Presupuesto Costes Totales
Personal	11.000
Amortización	462
Subcontratación de tareas	120
Costes de funcionamiento	450
Costes Indirectos	2.406
Total	14.438

Tabla A.2.6.1: Presupuesto.

Bibliografía

[1] MÜLLER, Meinard. Information Retrieval for Music and Motion. Springer. 2007, XVI, 318 p. 136 illus., 39 in color. ISBN 978-3-540-74047-6.

[2] ORTEGA, J. DTW_AlinTempDinam_ASAL [en línea]. Disponible en Web: <http://arantxa.ii.uam.es/~jortega/DTW_AlinTempDinam_ASAL.pdf>.

[3] SOTUYO ESPINOSA, Alejandra; VARGAS CONTRERAS, Ma. Guadalupe. "Reconocimiento de dígitos aislados en español". Asesoras: Alma E. Martínez; Fabiola Martínez. Universidad autónoma metropolitana, ciencias básicas e ingeniería. Noviembre 2003. Disponible en Web:< <http://148.206.53.231/UAMI10801.PDF>>.

[4] WIKIPEDIA. Biometría [en línea]. Disponible Web: <<http://es.wikipedia.org/wiki/Biometr%C3%ADa>>.

[5] ÁVILA JIMÉNEZ, María Antonia. "PFC: Evaluación Biométrica de Algoritmos de Autenticación de Firma Manuscrita". Director: Óscar Miguel Hurtado. Universidad Carlos III de Madrid, 2010.

[6] FERNÁNDEZ GARCÍA-OBLEDO, Juan. "PFC: Diseño de una interfaz de algoritmos de firma manuscrita". Director: Óscar Miguel Hurtado. Universidad Carlos III de Madrid, 2010.

[7] MARTÍNEZ SACRISTÁN, Jesús. "PFC: Implementación del procesado de datos mediante VHDL y FPGAs para reconocimiento de firma manuscrita". Director: Luís Mengíbar Pozo. Universidad Carlos III de Madrid.

[8] PERAITA GÓMEZ DE BALUGERA, Juan. "PFC: Implementación del algoritmo DTW mediante VHDL en FPGA". Director: Luís Mengíbar Pozo. Universidad Carlos III de Madrid.

[9] MIGUEL HURTADO, Óscar. "Online Signature Verification Algorithms and Development of Signature International Standards". Director/es: Raúl Sánchez Reíllo; Luis Mengibar Pozo. Universidad Carlos III de Madrid, Departamento de Tecnología Electrónica, 2011.

[10] L. Hong and A. Jain, "Integrating Faces and Fingerprints for Personal Identification", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 20, no. 12, 1998.

[11] MORALES, Domingo; RUÍZ-DEL-SOLAR, Javier. Sistemas Biométricos: Matching de huellas dactilares mediante transformada de Hough generalizada [en línea]. Disponible en Web: <http://www2.ing.puc.cl/~iing/ed429/sistemas_biometricos.html>.

[12] Instituto Nacional de Tecnologías de la Comunicación (Inteco) [en línea]. Disponible Web: <http://www.inteco.es/Seguridad/Observatorio/Estudios/estudio_biometria>.

[13] GALVIS TRASLAVIÑA, Carlos Mauricio. Introducción a la biometría [en línea]. Disponible en Web: <<http://www.monografias.com/trabajos43/biometria/biometria.shtml>>.

[14] INBIOSYS E.U. Disponible en Web: <<http://inbiosys.wordpress.com/2009/09/16/historia-de-la-biometria/>>.

[15] WIKIPEDIA. Huella dactilar [en línea]. Disponible en Web: <http://es.wikipedia.org/wiki/Huella_dactilar>.

[16] UNAM-Facultad de Ingeniería. Capítulo 3: Clasificación de los sistemas biométricos [en línea]. Disponible Web: <<http://redyseguridad.fip.unam.mx/proyectos/biometria/clasificacionsistemas/clasificaciontipo.html>>.

[17] REAL ACADEMIA ESPAÑOLA. Definición Voz [en línea]. Disponible en Web: <<http://buscon.rae.es/draeI/SrvltObtenerHtml?LEMA=voz&SUPIND=0&CAREXT=10000&NEDIC> No>.

[18] BIOMETRICS-ON. Información de biometría y de seguridad personal [en línea]. Disponible en Web: <<http://biometrics-on.com/es/reconocimiento-biometria-facial-escaner-de-rostro.asp>>.

[19] LAMAR LEÓN, Javier; GARCÍA REYES, Edel. Reconocimiento de personas por la forma de caminar “Gaint”: Estado del arte. Reconocimiento de patrones [en línea]. CENATAV (Cuba), 2009. Disponible en Web: <http://www.cenatav.co.cu/doc/RTecnicos/RT%20SerieAzul_024web.pdf>. ISSN 2072-6287.

[20] INTECO, Observatorio de la Seguridad de la Información. Estudio sobre las tecnologías biométricas aplicadas a la seguridad [en línea]. España, 2011. Disponible en Web: <<http://es.scribd.com/doc/75655062/Estudio-sobre-las-tecnologias-biometricas-aplicadas-a-la-seguridad>>.

[21] ALAMBRE, Weblog de Tecnología y Sociedad. Orígenes de la firma autógrafa. Disponible en Web:< <http://www.alambre.info/2003/12/08/origenes-de-la-firma-autografa/>>.

[22] PÉREZ HERNÁNDEZ, Alejandro. Reconocimiento y verificación de firmas manuscritas off-line basados en el seguimiento de sus trazos componentes [en línea]. España, 2003. Disponible en Web:< <http://www.esolva.com/firmas/index.html>>.

[23] PASCUAL GASPAS, Juan. "Uso de la Firma Manuscrita Dinámica para el Reconocimiento Biométrico de Personas en Escenarios Prácticos". Director/es: Valentín Cardeñoso Payo; Marcos Faúdez Zanuy. Universidad de Valladolid, Departamento de Informática. Disponible en Web:<<http://uvadoc.uva.es/bitstream/10324/130/1/TESIS52-100223.pdf>>.

[24] ESTEBAN AHMEDT, David; DELGADO TREJOS, Edilson; VARGAS, Jesús Francisco. Generación de espacios de representación de firmas dinámicas: una revisión enfocada al análisis de complejidad [en línea]. Disponible en Web:< http://edilsondt.com/wp-content/filebox/D01-Articulo_Scientia_et_Technica-Final_version.pdf>.

[25] SAKOE; CHIBA. Dynamic Programming Algorithm Optimization for spoken word recognition, 1978.

[26] SATO, Yukio; KOGURE, Kiyoshi. Online Signature Verification Base on Shape, Motion and Writing Pressure. Tokio (Japón), 1982.

[27] Diapositivas: "DTW: Alineamiento Temporal Dinámico". Universidad Autónoma de Madrid [En línea].

[28] Ortega-Garcia; Fierrez-Aguilar; Simon; Gonzalez; Faundez-Zanuy; Espinosa; Satue; Hernaez; Igarza; Vivaracho; Escudero; Moro. MCYT baseline corpus: a bimodal biometric database. IEE Proc.-Vis. Image Signal Process., Vol. 150, No. 6. Diciembre 2003.

[29] Tapiador Mateos, Marino; Sigüenza Pizarro, Juan A. Tecnologías biométricas aplicadas a la seguridad. RA-MA 2005. ISBN 84-7897-636-1.

[30] Jain, Anil K; Li; Stan Z. Encyclopedia of Biometrics Vol1, Vol2. Springer 2009. ISBN 978-0-387-73002-8.

[31] WIKIPEDIA. [en línea]. Disponible Web:<http://en.wikipedia.org/wiki/Norm_%28mathematics%29>.

[32] ROCHE, Ariel. “Árboles de decisión y Series de tiempo”. Director/es: Dr. Badih Ghattas. Facultad de Ingeniería, UDELAR, 2009.

[33] Biografía: Francis Galton. [en línea]. Disponible Web: <<http://galton.org/>>
<http://es.wikipedia.org/wiki/Francis_Galton>.