

**UNIVERSIDAD CARLOS III DE MADRID**  
**ESCUELA POLITÉCNICA SUPERIOR**



**PROYECTO FIN DE CARRERA**  
**INGENIERÍA TÉCNICA DE TELECOMUNICACIÓN:**  
**IMAGEN Y SONIDO**

**GROUP MODEL TREE**

**Modelos de preclasificación biométrica**

**AUTOR: ALBERTO BRAVO PÉREZ**

**DIRECTOR: LUIS PUENTE RODRÍGUEZ**

**Junio de 2010**



Proyecto Fin de Carrera

Group Model Tree  
(GMT)  
Modelos de preclasificación biométrica

AUTOR:

Alberto Bravo Pérez

DIRECTOR:

Luis Puente Rodríguez

La defensa del presente Proyecto Fin de Carrera se realizó el día 28 de Junio de 2010 y fue evaluada por el siguiente tribunal:

PRESIDENTE:       María Jesús Poza Lara

SECRETARIO:       Jessica Rivero Espinosa

VOCAL:             Judith Liu Jiménez



# AGRADECIMIENTOS

Una vez que se divisa la luz al final de este túnel, esa luz que parecía no divisarse nunca, la misma luz que pretende asegurarnos que estamos listos para saltar al vacío del mundo laboral.

Ahora que puedo avistar tierra firme sólo puedo quedarme con aquellos momentos y compañeros que decidieron acompañarme en este viaje. Porque lo importante no es llegar sino disfrutar del camino, como dijo Aldoux Husley: *“Por muy lentamente que os parezca que pasan las horas, os parecerán cortas si pensáis que nunca más han de volver a pasar.”*

Me gustaría empezar agradeciendo a Luis Puente, mi director de proyecto, por enseñarme tanto durante este tiempo. Por ser, más que un director, un compañero de laboratorio y eso es algo difícil de encontrar. También agradecerle por estar siempre que se le necesitaba, por volcarse en ayudarte... en definitiva, por ser como es (aunque a veces saque de quicio).

No puedo olvidarme tampoco de la inestimable ayuda de Diego, sobre todo en el apartado de programación, que sin su ayuda aún estaría programando. Y a todos los “invejadores insucios”, en particular a Pedro, que me acompañaron en el laboratorio y los momentos del café durante todo este tiempo.

Que menos que agradecer a Sergio y a todos los compañeros de carrera que amenizaron mi paso por allí. En especial a mis compañeros de andanzas y escudería; Fran y Félix. Por todos los momentos compartidos, por nuestras luchas contra gigantes, por las batallas libradas en lejanos parajes, por la amistad recibida, por poder llamaros amigos... por estar ahí. Agradecer a Andrea por poner límite a la paciencia y apretar las tuercas cuando era necesario.

Y por último, y no menos importante, a mi familia. A mis padres, por no echarme de casa y dejarme terminar la carrera. A mi madre por leer y corregirme este memorándum escrito en arameo. A mi hermano Iván, por su constante ayuda y optimismo. A mi hermana Marién, por enfadarse y sacarme de quicio continuamente. Y a mis abuelos, por escucharme. Especial mención a mi abuelo Eugenio, cuya experiencia se necesitaba en otro lugar y no dudó en ir a echar una mano.

Es difícil hacer un resumen de tantos años y recordar a toda la gente que compartió algún momento conmigo durante este tiempo. Perdonadme aquellos que olvidé en este viaje por la memoria.

A todos, GRACIAS

*Alberto,  
Junio 2010*



*A mis Hermanos  
y Padres*



*“No es el conocimiento sino el acto de conocer, no es la posición, sino el acto de llegar allí lo que produce el mayor gozo”*

*Karl Friedrich Gauss (1777-1855)  
Matemático, astrónomo y físico alemán*

*“Por muy lentamente que os parezca que pasan las horas, os parecerán cortas si pensáis que nunca más han de volver a pasar”.*

*Aldous Huxley (1894-1963)  
Escritor inglés*

*“La teoría es asesinada tarde o temprano por la experiencia”*

*Albert Einstein (1879-1955)  
Físico alemán*



# INDICE DE CONTENIDO

<b>AGRADECIMIENTOS.....</b>	<b>5</b>
<b>PRÓLOGO.....</b>	<b>19</b>
<b>CAPÍTULO 1: INTRODUCCIÓN .....</b>	<b>21</b>
<b>CAPÍTULO 2: REVISIÓN DEL MUNDO BIOMÉTRICO .....</b>	<b>23</b>
2.1    INTRODUCCIÓN.....	23
2.2    SISTEMAS BIOMÉTRICOS.....	25
2.2.1    Tipos de operaciones .....	26
2.2.2    Procesado de datos biométricos.....	27
2.3    EVALUACIÓN DE LOS SISTEMAS BIOMÉTRICOS .....	28
2.3.1    Competiciones.....	29
2.3.2    Tipos de Evaluaciones .....	30
2.3.3    Tipos de Errores .....	32
2.4    BIOMÉTRIA UNIMODAL.....	35
2.4.1    Firma manuscrita .....	35
2.4.2    Geometría Facial .....	36
2.4.3    Huella Dactilar .....	37
2.4.4    Iris.....	38
2.4.5    DNA .....	39
2.4.6    Huella de la palma de la mano.....	39
2.4.7    Dinámica del Tecleo .....	39
2.4.8    Movimiento (caminar) .....	40
2.4.9    Voz.....	40
2.5    BIOMETRÍA MULTIMODAL .....	41
<b>CAPÍTULO 3: RECONOCIMIENTO DE LOCUTOR .....</b>	<b>45</b>
3.1    INTRODUCCIÓN.....	45
3.2    EXTRACCIÓN DE CARACTERÍSTICAS .....	47
3.2.1    Coeficientes Cepstrales .....	47
3.3    ALGORITMOS: CLASIFICACIÓN Y MODELADO DE CARACTERÍSTICAS .....	50
3.3.1    Vector Quantization (VQ).....	50
3.3.2    Hidden Markov Model (HMM).....	51
3.3.3    Gaussian Mixture Models (GMM).....	52
3.3.4    Support Vector Machines (SVM).....	54
3.3.5    Neural Networks (NN).....	55
3.4    DECISIÓN.....	56

<b>CAPÍTULO 4: ESTADO DEL ARTE.....</b>	<b>57</b>
4.1 INTRODUCCIÓN.....	57
4.2 PROBLEMA DE LA IDENTIFICACIÓN.....	61
4.3 SOLUCIÓN ACTUAL DEL PROBLEMA.....	61
4.3.1 Soluciones Adoptadas en Huella Dactilar.....	61
4.3.2 Soluciones Adoptadas en Iris .....	63
 <b>CAPÍTULO 5: METODOLOGÍA DE TRABAJO .....</b>	<b>67</b>
5.1 METODOLOGÍA DE TRABAJO .....	68
 <b>CAPÍTULO 6: BASE DE DATOS .....</b>	<b>69</b>
6.1 BANCA .....	69
6.2 BIOSEC BASELINE CORPUS .....	71
6.3 MICROAES .....	72
 <b>CAPÍTULO 7: ENTORNO DE TRABAJO .....</b>	<b>75</b>
 <b>CAPÍTULO 8: LINEA BASE.....</b>	<b>77</b>
 <b>CAPÍTULO 9: PLANTEAMIENTO Y EVOLUCIÓN DEL GMT .....</b>	<b>79</b>
9.1 PLANTEAMIENTO .....	80
9.2 GMT: EVOLUCIÓN .....	82
9.2.1 B-GMT (Bottom-Up) .....	84
9.2.2 Conclusiones .....	90
9.2.3 T-GMT (Top-Down) .....	91
9.2.4 Conclusiones .....	93
9.2.5 T-NGMT (Top-Down).....	93
9.3 MEJORAS PARA EL T-NGMT .....	96
9.3.1 Distintas maneras de calcular distancias entre usuarios .....	96
9.3.2 Implementación Matemática.....	98
9.3.3 Evaluación de los Resultados .....	104
9.3.4 Conclusiones .....	106
9.3.5 Nueva forma de elección de usuarios.....	106
9.3.6 Evaluación de los Resultados .....	108
9.3.7 Conclusiones .....	110
9.4 CONCLUSIONES GENERALES .....	110
 <b>CAPÍTULO 10: CONCLUSIONES Y TRABAJOS FUTUROS.....</b>	<b>113</b>
10.1 CONCLUSIONES .....	113
10.2 TRABAJOS FUTUROS.....	115
 <b>REFERENCIAS.....</b>	<b>117</b>
 <b>ANEXO I: GESTIÓN DEL PROYECTO .....</b>	<b>131</b>
11.1 INTRODUCCIÓN.....	133

11.2	PLANIFICACIÓN DEL PROYECTO .....	134
11.3	ESTIMACIÓN DE LOS RECURSOS TEMPORALES.....	134
11.4	ESTIMACIÓN DE LOS RECURSOS ECONÓMICOS.....	135
11.4.1	Recursos Materiales .....	135
11.4.2	Recursos Humanos.....	136
11.4.3	Costes Totales .....	137
<b>ANEXO II: HERRAMIENTAS UTILIZADAS .....</b>		<b>139</b>



# INDICE DE FIGURAS

FIGURA 1: SISTEMA BIOMÉTRICO, EXTRAÍDA DE [ANIL K. JAIN ET AL. 2004A]	26
FIGURA 2: DIAGRAMA DE FLUJO DEL PROCESADO BIOMÉTRICO	28
FIGURA 3: DIAGRAMA (RECURSOS/COSTES) Vs (VERACIDAD) DE LOS TEST OFF-LINE	31
FIGURA 4: UMBRAL DE DECISIÓN	33
FIGURA 5: REPRESENTACIÓN DE FRR, FAR Y ERR	34
FIGURA 6: CURVA ROC	34
FIGURA 7: CURVA DET PARA DIFERENTES RASGOS BIOMÉTRICOS	34
FIGURA 8: ARQUITECTURA TÍPICA DE VERIFICACIÓN ON-LINE	35
FIGURA 9: EXTRACCIÓN DE CARACTERÍSTICAS DE LA BIOMETRÍA FACIAL	36
FIGURA 10: EXTRACCIÓN DE CARACTERÍSTICAS DE LA HUELLA DACTILAR	37
FIGURA 11: FASES DEL PROCESO DE RECONOCIMIENTO DE IRIS	38
FIGURA 12: EXTRACCIÓN DE CARACTERÍSTICAS DE LA PALMA DE LA MANO	39
FIGURA 13: CARACTERÍSTICAS DE LA DINÁMICA DEL TECLEO	40
FIGURA 14: EXTRACCIÓN DE CARACTERÍSTICAS DE LA SILUETA	40
FIGURA 15: EXTRACCIÓN DE LAS CARACTERÍSTICAS DE LA VOZ	41
FIGURA 16: VARIOS ESCENARIOS DE UN SISTEMA BIOMÉTRICO MULTIMODAL (EXTRAÍDO DE [ANIL K. JAIN ET AL, 2004A])	43
FIGURA 17: DIAGRAMA DEL RECONOCIMIENTO DE LOCUTOR	46
FIGURA 18: FASE ENTRENAMIENTO/IDENTIFICACIÓN	46
FIGURA 19: ESPECTRO DE LA SEÑAL VOCAL/CEPSTRUM	48
FIGURA 20: FILTROS TRIANGULARES PARA OBTENER MEL-CEPSTRUM	49
FIGURA 21: PROCESO DE EXTRACCIÓN DE LOS COEFICIENTES MEL-CEPSTRALES	49
FIGURA 22: PROCESO DE EXTRACCIÓN DE LOS COEFICIENTES CEPSTRALES MEDIANTE LPC	50
FIGURA 23: REPRESENTACIÓN ESPACIAL DEL CODEBOOK	51
FIGURA 24: TRANSICIONES DE ESTADO DEL MODELO DE MARKOV DE 5 ESTADOS	52
FIGURA 25: EJEMPLO DE MODELADO GMM	53
FIGURA 26: HIPERPLANO SEPARADOR DE CLASES	54
FIGURA 27: DIAGRAMA DE RED NEURONAL UNITARIA Y MULTICAPA	55
FIGURA 28: DIAGRAMA GENERAL DEL PROCESO DE DECISIÓN	56
FIGURA 29: DIAGRAMA DE LA BÚSQUEDA EXHAUSTIVA (ES)	59
FIGURA 30: DIAGRAMA DE LA BÚSQUEDA EXHAUSTIVA CON UMBRAL GENERAL (ES-GT)	59
FIGURA 31: DIAGRAMA DE LA BÚSQUEDA EXHAUSTIVA CON UMBRAL POR USUARIO (ES-TPU)	60
FIGURA 32: TIPOS DE HUELLA (DACTILOSCOPIA DE HENRY)	62
FIGURA 33: DIAGRAMA DE PROCESOS DE LA METODOLOGÍA DE TRABAJO	67
FIGURA 34: METODOLOGÍA DE AGRUPACIÓN PROPUESTA	80
FIGURA 35: METODOLOGÍA DE AGRUPACIÓN CON CRITERIO DE PARADA	81
FIGURA 36: MÓDULO PARA LA CREACIÓN DEL GMT	81
FIGURA 37: ÁRBOL GMT	82
FIGURA 38: ESQUEMA DE UN GMT CON SIETE USUARIOS	83
FIGURA 39: DIAGRAMA DE PROCESOS B-GMT	84
FIGURA 40: ESQUEMA DE CREACIÓN DEL B-GMT	85
FIGURA 41: DISTANCIA PUNTO MEDIO A GAUSSIANAS	87
FIGURA 42: FORMA IDEADA PARA CÁLCULO DE DISTANCIAS (CASO PARTICULAR DE 3 GAUSSIANAS/USUARIO)	88
FIGURA 43: DISTANCIA EUCLÍDEA Vs MANHATTAN	89
FIGURA 44: CÁLCULO DE DISTANCIA CITY-BLOCK (BOTTOM-UP)	90
FIGURA 45: ESQUEMA DE CREACIÓN DEL T-GMT	91
FIGURA 46: DIAGRAMA DE PROCESOS T-GMT	92

FIGURA 47: ESQUEMA DE CREACIÓN DEL T-NGMT.....	94
FIGURA 48: DIAGRAMA DE PROCESOS T-NGMT .....	95
FIGURA 49: REPRESENTACIÓN GRÁFICA DE LA DISTANCE BETWEEN MIDDLES 1.....	96
FIGURA 50: REPRESENTACIÓN GRÁFICA DE LA DISTANCE BETWEEN GAUSSIANS 1.....	97
FIGURA 51: REPRESENTACIÓN GRÁFICA DE DISTANCE MIDDLE TO GAUSSIANS 1 .....	98
FIGURA 52: FRONTERA DE DECISIÓN VS OUTLIERS.....	107
FIGURA 53: PROCESO DE DESARROLLO SOFTWARE .....	133
FIGURA 54: DIAGRAMA GANTT.....	138

# INDICE DE TABLAS

TABLA 1: COMPARACIÓN DE VARIAS CARACTERÍSTICAS BIOMÉTRICAS (A=ALTO, M=MEDIO, B=BAJO) .....	25
TABLA 2: EDADES DE LOS DONANTES.....	73
TABLA 3: COMPARATIVA DE LAS DIFERENTES VARIANTES DE IDENTIFICACIÓN EXHAUSTIVA.....	78
TABLA 4: FORMA DE LA LISTA QUE CONSTITUYE EL GMT .....	83
TABLA 5: COMPARATIVA DE FNMR DEL PRIMER EXPERIMENTO .....	104
TABLA 6: COMPARATIVA DE FMR DEL PRIMER EXPERIMENTO.....	105
TABLA 7: COMPARATIVA DE HTER DEL PRIMER EXPERIMENTO .....	105
TABLA 8: COMPARATIVA TIEMPO TOTAL DEL PRIMER EXPERIMENTO .....	105
TABLA 9: COMPARATIVA DE TIEMPO PROMEDIO POR IDENTIFICACIÓN EN EL PRIMER EXPERIMENTO .....	106
TABLA 10: COMPARATIVA DE FNMR DEL SEGUNDO EXPERIMENTO .....	108
TABLA 11: COMPARATIVA DE FMR DEL SEGUNDO EXPERIMENTO.....	108
TABLA 12: COMPARATIVA DE HTER DEL SEGUNDO EXPERIMENTO .....	109
TABLA 13: COMPARATIVA TIEMPO TOTAL DEL SEGUNDO EXPERIMENTO .....	109
TABLA 14: COMPARATIVA DE TIEMPO PROMEDIO POR IDENTIFICACIÓN EN EL SEGUNDO EXPERIMENTO .....	109
TABLA 15: RECURSOS TEMPORALES POR FASES DEL PROYECTO .....	135
TABLA 16: RECURSOS MATERIALES PARA EL DESARROLLO DEL PROYECTO .....	136
TABLA 17: RECURSOS HUMANOS .....	136
TABLA 18: COSTES TOTALES DEL PROYECTO .....	137

# INDICE DE FÓRMULAS

(FÓRMULA 1) .....	88
(FÓRMULA 2) .....	89
(FÓRMULA 3) .....	89
(FÓRMULA 4) .....	90
(FÓRMULA 5) .....	99
(FÓRMULA 6) .....	99
(FÓRMULA 7) .....	99
(FÓRMULA 8) .....	99
(FÓRMULA 9) .....	100
(FÓRMULA 10) .....	100
(FÓRMULA 11) .....	100
(FÓRMULA 12) .....	100
(FÓRMULA 13) .....	101
(FÓRMULA 14) .....	101
(FÓRMULA 15) .....	101
(FÓRMULA 16) .....	102
(FÓRMULA 17) .....	102
(FÓRMULA 18) .....	102
(FÓRMULA 19) .....	103
(FÓRMULA 20) .....	103
(FÓRMULA 21) .....	103
(FÓRMULA 22) .....	104
(FÓRMULA 23) .....	104
(FÓRMULA 24) .....	104
(ECUACIÓN 1) .....	86
(ECUACIÓN 2) .....	86
(ECUACIÓN 3) .....	87
(ECUACIÓN 4) .....	87
(ECUACIÓN 5) .....	87
(ECUACIÓN 6) .....	88
(ECUACIÓN 7) .....	89

# PRÓLOGO

El reconocimiento biométrico ha sido un parte importante en la creación de la sociedad. Desde el comienzo de la civilización se ha utilizado el reconocimiento biométrico para establecer clases y grupos sociales, estatus personales y de seguridad... en definitiva, para la estructuración de la sociedad.

Como consecuencia de todo lo acontecido durante estos últimos años, se ha producido un incremento en la demanda de sistemas de identificación humana fiables. Todo esto repercutió en un crecimiento de la industria biométrica, según el informe publicado por Frost & Sullivan<sup>1</sup> se prevé como la industria que más crecerá durante los próximos años.

Este crecimiento puede atribuirse a una serie de situaciones civiles y militares, tanto dentro como fuera de la industria biométrica. Entre las mismas están los continuos perfeccionamientos de la tecnología ofertada, lo que incluye mejores índices de precisión y niveles de rendimiento. Otras surgen de las necesidades de seguridad requeridas por la población, importando más los mayores logros en la protección contra el delito. También se puede atribuir a un incremento en la demanda y más altas tasas de aceptación de las ofertas a nivel de consumidor. Como bien refleja el informe anteriormente mencionado.

Todo este avance sufrido en la industria biométrica debe venir acompañado por un incremento en la eficiencia de los métodos de reconocimiento biométrico. A lo largo de este documento se presentará un enfoque a la identificación biométrica, llamada GMT (Group Model Tree). Este planteamiento pretende aligerar los tiempos de cálculo ofreciendo así una mayor seguridad y reducción de los recursos computacionales.

---

<sup>1</sup> <http://www.nfcnews.com/2010/06/03/informe-pronostica-crecimiento-del-mercado-biometrico-en-norteamerica>



---

# CAPÍTULO 1

---

## INTRODUCCIÓN

A día de hoy la seguridad es una parte importante en la vida de las personas. Los mecanismos de seguridad están tan asumidos por la sociedad que se utilizan continuamente sin que ello nos resulte extraño, desde encender un ordenador hasta conducir un coche. Se utilizan prácticamente en cualquier lugar, coches, casas, oficinas de trabajo, edificios gubernamentales, museos, aeropuertos... Pero no siempre fue así, hubo un punto de inflexión que ayudó a intensificar los mecanismos de seguridad. Este consiguió reflotar y ampliar el ámbito de uso de una industria en concreto: “La industria biométrica”.

Los atentados de las Torres Gemelas ayudaron a la implantación definitiva de la industria biométrica en el mercado actual. A partir de ese momento las exigencias de seguridad se fueron intensificando gradualmente. Las tarjetas de identidad y los pasaportes ya no eran suficientes ya que pueden falsificarse. Se necesitaba identificar individuos con participación tanto activa como pasiva de los mismos. Por lo tanto, se necesitaban procedimientos nuevos e infalsificables. Convirtiendo a la biometría en una creciente línea de investigación tecnológica y convirtiéndose en una industria con mucho futuro.

A raíz de aquellos acontecimientos se generó una sensación de inseguridad pública en aeropuertos y estaciones de tren. Probablemente la identificación de individuos es el cometido más importante para aumentar la seguridad en las zonas de acceso público o de tránsito masivo de personas. Detectar sujetos potencialmente peligrosos antes de su

actuación es un “plus” añadido a la seguridad personal. La introducción de la biometría, y más concretamente la identificación biométrica, consiguió disminuir esa sensación de inseguridad generada.

La biometría ofrece múltiples formas de reconocimiento de personas, y mediante la tecnología ofrecida por la industria biométrica se pretende evitar nuevos desastres. Este tipo de reconocimiento basado en rasgos intrínsecos a la persona, lleva el hándicap añadido de no unicidad. Bajo esa característica no existen dos personas con todos los rasgos biométricos iguales. El reconocimiento biométrico de personas ofrece dos alternativas: la verificación y la identificación. Con la verificación se pretende corroborar la autenticidad de la persona, mientras que con la identificación se pretende reconocer a una persona.

Para realizar una identificación biométrica se debe comparar la muestra biométrica con las existentes en la base de datos. Por lo tanto, se necesita un elevado número de comparaciones hasta conseguir una identidad, en el caso de que estuviera registrada. Lo cual genera un tiempo elevado de cómputo variable con el tamaño de los usuarios registrados en el sistema. Todo ello repercute también en un elevado coste computacional requerido por el sistema en cada identificación.

El objetivo del proyecto es estudiar alternativas para disminuir los “elevados tiempos de cálculo” necesarios en cada identificación. Este proyecto se encargará de establecer una metodología de preclasificación de los modelos existentes. Esta metodología pretende ser de ámbito general, aplicable para cualquier tipo de tecnología biométrica sin atender a la semántica de las características extraídas de la muestra biométrica.

Este documento se estructura de la siguiente manera: En primera lugar se describen las motivaciones y objetivos del proyecto. Se continúa con un acercamiento al mundo biométrico y al reconocimiento de voz. Se prosigue con una revisión sobre el estado del arte del problema “elevados tiempos de cómputo”. En el siguiente capítulo se expone la metodología de trabajo desarrollada para conseguir los objetivos. En los tres capítulos siguientes se describirán la elección de la base de datos, la línea base y se concluirá con un análisis, evaluación y evolución de la solución adoptada. Se concluye la memoria con un apartado de conclusiones y futuras vías de investigación. Las últimas secciones del documento incluirán los diferentes anexos elaborados durante la creación del proyecto.

---

# CAPÍTULO 2

---

## REVISIÓN DEL MUNDO BIOMÉTRICO

En este capítulo se ofrece una visión resumida de la biometría. En primer lugar, se revisa el concepto de biometría para continuar con la descripción de un sistema biométrico y sus operaciones. Se prosigue analizando los sistemas de experimentación y las medidas de evaluación de los mismos. Y se termina por la enumeración de las tecnologías biométricas más utilizadas y los sistemas multimodales.

### 2.1 INTRODUCCIÓN

Desde el comienzo de la civilización la identificación de humanos ha sido crucial para la construcción de una sociedad. Desde hace cientos de años se fueron utilizando las características de nuestro propio cuerpo para reconocernos entre nuestros semejantes [Anil K. Jain et al. 2004a]. Entre ellas la cara, la voz, el modo de caminar...

Biometría se refiere al concepto de reconocimiento de individuos mediante algunos rasgos o características personales (biológicas, físicas, psicológicas o de comportamiento) y unívocas [Anil K. Jain et al. 2004a]. Según [Dessimoz et al, 2005] estas características biológicas deben de ser mensurables o estadísticamente medibles.

El santo y seña de los siglos pasados será en el siglo XXI el propio cuerpo y sus características físicas, únicas y distintas a las de cualquier otro ser humano. La industria biométrica abre un amplio abanico de posibilidades tanto en accesibilidad como en

seguridad, está en continuo desarrollo y en los últimos años se han producido importantes avances tecnológicos. El reconocimiento biométrico es un concepto antiguo, véase [Raphael et al, 1974], pero se podría decir que los sistemas biométricos y el cuerpo humano serán la llave maestra que nos abrirá muchas puertas en el futuro próximo.

Para que un rasgo humano sea útil desde el punto de vista biométrico debe de ser parametrizable o mensurable. Además de ello, para que ese mismo rasgo pueda ser utilizado como característica biométrica debe cumplir una serie de requisitos que aparecen en [Anil K. Jain et al. 2004a]:

- ✓ *Universalidad*: El rasgo biométrico debería existir para todas las personas de una población.
- ✓ *Distintivos o inconfundibles*: Dos personas deben ser suficientemente diferentes en términos de rasgos.
- ✓ *Permanentes*: Las características deben de ser suficientemente invariantes con el tiempo dentro de un sujeto.
- ✓ *Accesibles*: Las características deben ser cuantificadas mediante una operación moderadamente sencilla.

A menudo, a la hora de elegir que rasgos biométricos son los más adecuados para una aplicación concreta existen otros tipos de requisitos que deberían ser tomados en cuenta:

- ✓ *Rendimiento*: Hace referencia a la precisión y velocidad del proceso de reconocimiento como a los recursos necesarios para conseguir los niveles de rendimiento deseados.
- ✓ *Aceptabilidad*: Representa el grado de aceptación que la sociedad presenta ante que un determinado rasgo sea capturado.
- ✓ *Invulnerabilidad*: Refiere a la robustez que ofrece el rasgo al sistema frente a métodos falsificables.

Una única característica biométrica no es la herramienta de reconocimiento final y la elección de la misma dependerá de la aplicación. En la (Tabla 1: Comparación de varias técnicas biométricas) se muestra las características biométricas más comunes en los sistemas biométricos actuales, así como el grado de cumplimiento de los requisitos anteriores.

Aunque en el ámbito forense (judicial, policial y pericial) el análisis científico de evidencias biométricas se ha venido utilizando desde hace más de un siglo, el reconocimiento biométrico automático como medio de autenticación personal en

aplicaciones comerciales o civiles, es un área de investigación y desarrollo bastante más reciente y en pleno crecimiento.

TIPO	Rasgos Biométricos	Universalidad	Distintivos o Inconfundibles	Permanentes	Accesibles	Rendimiento	Aceptabilidad	Invulnerabilidad
Fisiológicas	Voz <sup>2</sup>	M	B	B	M	B	A	A
	Huella dactilar	M	A	A	M	A	M	M
	Iris	A	A	A	M	A	B	B
	Retina	A	A	M	B	A	B	B
	Geometría facial	A	B	M	A	B	A	A
	Venas del dorso de la mano	M	M	M	M	M	M	B
	Geometría de la palma de la	M	M	M	A	M	M	M
	Huella de la palma de la mano	M	A	A	M	A	M	M
	Geometría de la oreja	M	M	A	M	M	A	M
	Olor corporal	A	A	A	B	B	M	B
Comportamiento	Voz <sup>1</sup>							
	Firma	B	B	B	A	B	A	A
	Dinámica de tecleo	B	B	B	M	B	M	M
	Movimiento y/o cadencia del	M	B	B	A	B	A	M
	DNA	A	A	A	B	A	B	B

Tabla 1: Comparación de varias características biométricas (A=Alto, M=Medio, B=Bajo)

## 2.2 SISTEMAS BIOMÉTRICOS

En esencia el núcleo del sistema biométrico es un sistema de reconocimiento de patrones [Anil K. Jain et al, 2000]. De forma vaga, [Watanabe, 1985] un patrón biométrico es “una entidad a la que se le podría adjudicar un nombre”. Por tanto, un sistema biométrico es un sistema que reconoce individuos mediante un vector de características derivado de algún rasgo físico o de comportamiento que el individuo posee [Anil K. Jain et al, 2004a].

Los sistemas biométricos se caracterizan por realizar su trabajo en tres fases. [Anil K. Jain et al, 2000] [Blackburn, 2004] [Kresimir Delac and Mislav Grgic, 2004] En una primera fase, se obtienen muestras biométricas de un individuo mediante algún sensor o

<sup>2</sup>La voz encierra tanto componentes físicas como componentes psicológicas y de comportamiento

sensores específicos. Seguidamente se extrae el conjunto de valores correspondientes a las características relevantes de la muestra (vector de características) adquirida que se pueda manipular y medir mediante algún algoritmo o fórmula matemática. Por último, se compara los datos originados anteriormente con los almacenados y se decide la aceptación o rechazo del individuo (Figura 1).

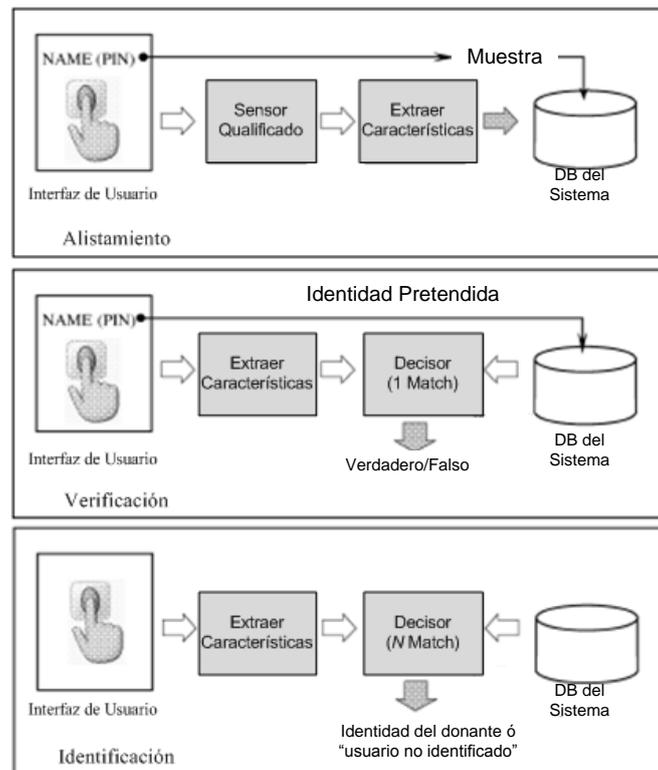


Figura 1: Sistema Biométrico, extraída de [Anil K. Jain et al. 2004a]

### 2.2.1 Tipos de operaciones

Los pasos comunes explicados anteriormente son utilizados en una serie de operaciones biométricas de alto nivel [D. Dessimoz et al, 2005] [Anil K.Jain et al, 2004a]: Alistamiento, Verificación e Identificación.

- *Alistamiento:* Con esta operación se añade un nuevo usuario al sistema. Mediante un cierto número de representaciones biométricas se entrena un modelo para el usuario y se adapta el modelo de los usuarios de fondo en caso de ser necesario (Figura 1). También se pueden dar de baja a usuarios que ya no pertenecen al sistema biométrico [Arslan Brömme, 2005].
- *Verificación:* La verificación consiste en comparar la representación biométrica de un usuario donante frente al modelo de un usuario pretendido. Una vez contrastada la información con el modelo se obtiene una puntuación que se

coteja con un umbral para aceptar o denegar la identidad (Figura 1) [J. L. Wayman, 2001].

- *Identificación:* Mediante esta operación se busca la representación biométrica más parecida en el sistema a la muestra de entrada (Figura 1). El sistema reconoce a un individuo mediante la comparación de sus características biométricas con los modelos de todos los usuarios almacenados en el sistema. A partir de las puntuaciones obtenidas para cada modelo se establece la identidad [Karthik Nandakumar et al. 2009] [Felipe Ochoa et al. 2008]. La identificación es un proceso relativamente lento debido al elevado número de comparaciones que se deben realizar para dar con una identidad.

### 2.2.2 Procesado de datos biométricos

El procesado de datos biométricos es una sucesión lineal de operaciones. Para cualquier modalidad biométrica, los datos serán transformados siguiendo una serie de pasos que se pueden observar de manera esquematizada en la Figura 2 tomada de [D. Dessimoz et al. 2006]:

- ✓ *Captura:* La información es digitalizada por un sensor específico y almacenada.
- ✓ *Pre-Procesado:* Se realiza una normalización de la información en el dominio de la señal con el fin de eliminar información corrupta. Quedando la información preparada para la posterior extracción de características.
- ✓ *Extracción de características:* De la señal digitalizada y pre-procesada se extrae la información más relevante consiguiendo una nueva representación de la señal en un nuevo dominio. De esta forma, se reduce la dimensionalidad de la señal de entrada y se crea una representación a nivel de características del patrón de entrada. Esta nueva señal se emplea en el posterior reconocimiento de patrones y se la conocerá en el documento como vector de características.
- ✓ *Post-Procesado:* Las características obtenidas en el paso anterior se vuelven a normalizar para eliminar el posible sesgo, o bien para conseguir una mejor adaptación de las mismas al clasificador.
- ✓ *Creación del modelo:* Una vez se extrajo la información relevante de la señal y se obtuvo la representación en el nuevo dominio, se procede a la generación del modelo para dicho usuario. La generación del modelo no es más que una representación matemática y genérica de un usuario a partir de su vector de características. Este modelo será usado para futuras comparaciones.
- ✓ *Creación del modelo de fondo:* También llamado modelo del mundo. Algunos sistemas necesitan de la creación de un modelo para normalizar las

puntuaciones. Este modelo representa una “media” de los usuarios con respecto a la población del sistema.

- ✓ *Almacenamiento de modelos:* Una vez estimados los datos que definen la identidad de un usuario los modelos son almacenados para su futuro uso en operaciones biométricas.
- ✓ *Comparación de modelos:* Se compara un conjunto de vectores de características con el modelo biométrico almacenado de un usuario. Esta comparación da como resultado una puntuación que indica cuanto se asemejan los datos biométricos de entrada con las identidades de los modelos biométricos que han sido cotejadas. En ocasiones, usar el modelo del mundo en la comparación mejora los resultados.
- ✓ *Establecimiento de umbrales:* A partir de los datos de usuarios reales e impostores se calcula un umbral de decisión que marca la aceptación de la identidad en el sistema.

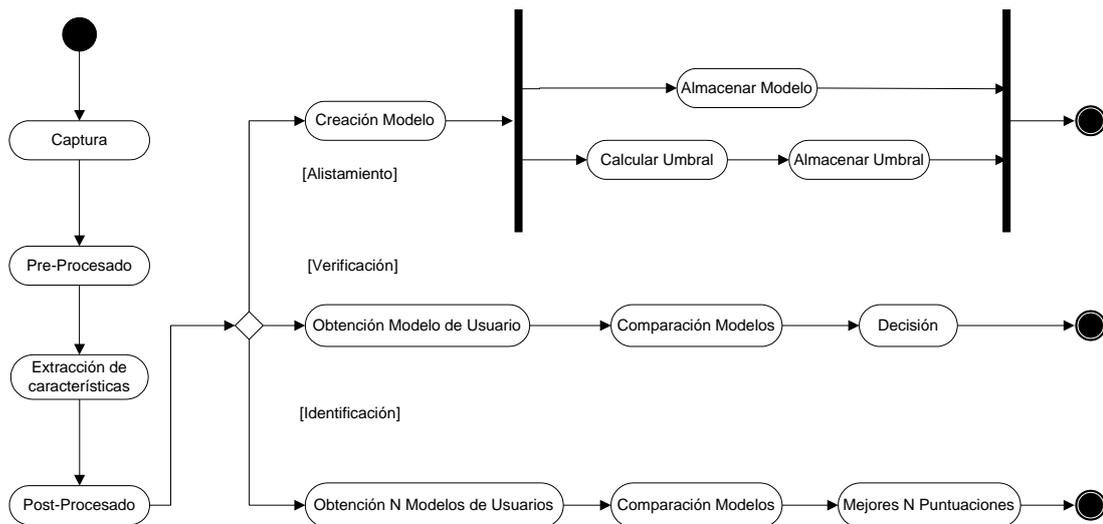


Figura 2: Diagrama de flujo del Procesado Biométrico

## 2.3 EVALUACIÓN DE LOS SISTEMAS BIOMÉTRICOS

Durante años se han ido introduciendo multitud de avances en los sistemas biométricos, por parte de la comunidad científica y organizaciones comerciales. Como bien describe [R. Capelli et al, 2006], las prestaciones estaban siendo medidas mediante protocolos de evaluación no estandarizados y bases de datos propias. Obteniendo así un conjunto de datos no susceptibles de ser comparados y de entredicha importancia.

Un protocolo de evaluación definido por [P.J. Phillips et al, 2000] determina tanto la forma en la que se testea un sistema y se miden sus prestaciones como la información utilizada para evaluarlo. Para que un sistema de evaluación sea aceptado debe de cumplir una serie de requisitos y someterse a una metodología de evaluación. Todo este trabajo ha sido realizado por diferentes organismos nacionales y finalmente se aunaron en la Common Criteria (CC) <sup>3</sup> para posteriormente ser estandarizada por la ISO 15408 [R. S. Reñillo et al, 2007]. Para que un protocolo de evaluación sea aceptado por la comunidad científica, los detalles de sus procedimientos han de estar publicados junto con las medidas de prestaciones propuestas y algunos ejemplos representativos del conjunto de datos.

### 2.3.1 Competiciones

Durante el siglo XXI se han desarrollado multitud de iniciativas con el claro fin de poder establecer una medida de evaluación única y fiable. Las iniciativas más importantes son:

- ✓ La Universidad de Bolonia desarrolla desde el año 2000 una competición bianual, la Fingerprint Verification Competition<sup>4</sup>, para evaluar el Estado del Arte de los sistemas de verificación basados en huella dactilar.
- ✓ El Instituto Nacional de Estándares y Tecnología (NIST)<sup>5</sup> desarrolla desde el año 2005 una competición para evaluar el estado del arte de los sistemas de reconocimiento de iris.
- ✓ El Instituto de Ciencias de la Academia China de las Ciencias realiza desde el año 2004 competiciones para evaluar las prestaciones de los sistemas de verificación de identidad basados en iris. Así mismo, desarrolla la Signature Verification Competition para valorar los sistemas de verificación de identidad basados en firma on-line.
- ✓ El Face Recognition Vendor Test<sup>6</sup> desarrolla competiciones para la evaluación de sistemas comerciales basados en geometría facial desde el año 2000. En sus inicios llamado FERET [P.J. Phillips et al, 2000b].
- ✓ El Instituto Nacional de Estándares y Tecnología (NIST)<sup>7</sup> desarrolla desde el año 2005 una competición para evaluar el estado del arte de los sistemas de reconocimiento de la cara.

---

<sup>3</sup> <http://www.commoncriteriaportal.org/>

<sup>4</sup> <http://bias.csr.unibo.it/fvc2006/>

<sup>5</sup> <http://iris.nist.gov/ice/>

<sup>6</sup> <http://www.frvt.org/>

<sup>7</sup> <http://face.nist.gov/frgc/>

### 2.3.2 Tipos de Evaluaciones

La evaluación de los sistemas biométricos es una parte muy importante del reconocimiento biométrico. Una evaluación es correcta cuando las puntuaciones obtenidas permiten distinguir entre los distintos enfoques y tecnologías existentes. Las fortalezas y debilidades encontradas en el proceso de evaluación indican las aplicaciones para las que las tecnologías son más adecuadas. Se describen a continuación los distintos tipos de evaluaciones existentes según [P. J. Phillips et al, 2000] [A. J. Mansfield et al, 2002]:

- *Evaluación de Tecnología:* Es la tecnología de evaluación más generalizada. Se comparan los diferentes algoritmos existentes y prototipos de laboratorio para una misma tecnología. Todos los algoritmos son testados con una base de datos estandarizada y recogida mediante un sensor universal. Se realiza para determinar el estado del arte de una tecnología e identificar los progresos tecnológicos.
- *Evaluación de Escenario:* Valora las prestaciones globales de los sistemas para un prototipo de escenario o aplicación simulada. El propósito principal de este tipo de evaluación es determinar si una tecnología biométrica es lo suficientemente madura para alcanzar los requisitos de prestaciones en un tipo de aplicaciones. Se suele evaluar el comportamiento de los diferentes algoritmos y combinación de sensores bajo condiciones reales.
- *Evaluación Operacional:* Aunque es similar a la evaluación de escenario su diferencia radica en que valora las prestaciones de un algoritmo específico para un tipo de aplicación específica. El propósito principal de la evaluación operacional es determinar si un sistema biométrico alcanza los requisitos para una aplicación específica.

La evaluación de un sistema biométrico requiere una captura de información [R. Capelli et al, 2006]. Esta información se utiliza tanto para el entrenamiento del sistema biométrico, como para la obtención de las puntuaciones en la etapa de reconocimiento [Anil K. Jain et al, 2004a]. Esa información puede ser utilizada en tiempo real o por el contrario se puede almacenar para un uso posterior. Dependiendo de la manera de utilizar la información según [A. J. Mansfield et al, 2002] [Capelli et al, 2006] se pueden clasificar los test de evaluación en dos grupos:

- *Off-line:* Este tipo de evaluación se realiza con la presencia del usuario. El alistamiento y el cálculo de las puntuaciones se realizan inmediatamente después de obtener la información biométrica. Posee la ventaja de que sólo se almacena la información estrictamente necesaria debido a que la muestra se elimina tras su utilización. Desafortunadamente, una misma operación no puede ser repetida por falta de la muestra descartada.

- *On-line*: Este tipo de evaluación se basa en una colección de datos capturados y almacenados previamente necesitando así un mayor almacenamiento de información. Este tipo de evaluación ofrece un mayor control sobre qué información será utilizada en cada operación. Permitiendo reproducir un mismo test bajo las mismas condiciones.

A su vez, los test de evaluación off-line se subdividen en diferentes tipos como bien recoge [D. Carrero et al., 2009]. No todos ellos consumen los mismos recursos computacionales (Figura 3) [Capelli et al, 2006]:

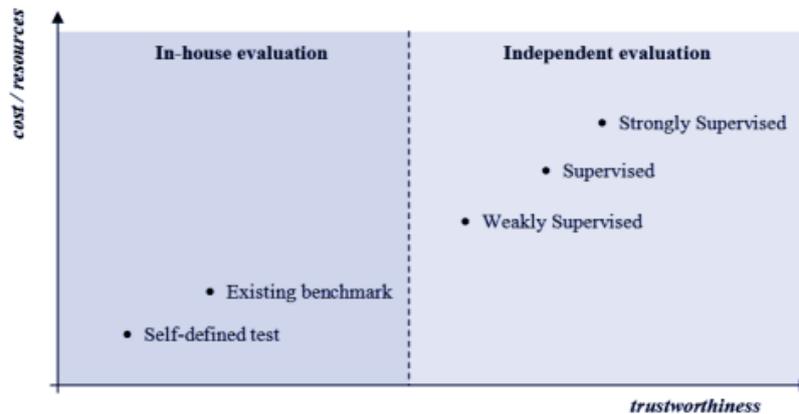


Figura 3: Diagrama (recursos/costes) Vs (veracidad) de los test off-line

- ✓ *In-house – self defined test*: La recolección de la información biométrica y la definición del protocolo de prueba es definido por el investigador. Como consecuencia, no permite ningún tipo de comparación debido a que los datos de test no pueden ser reproducidos por terceras personas.
- ✓ *In-house – existing Benchmark*: El test se realiza sobre una base de datos de carácter público y de acuerdo con los protocolos existentes. En este caso, los resultados obtenidos son comparables con otros obtenidos mediante el mismo protocolo y la misma base de datos. El principal inconveniente reside en la imposibilidad de predecir las verdaderas prestaciones del sistema en aplicaciones reales debido al posible sobreajuste realizado sobre los datos utilizados.
- ✓ *Independent – weakly supervised*: La base de datos no se hace pública hasta el comienzo de la evaluación. Por un lado, el test se realiza en el entorno del investigador mediante datos sin etiquetar y con limitaciones temporales. Por otro lado, las prestaciones del sistema se determinan por evaluadores independientes a partir de las puntuaciones obtenidas por el investigador durante la prueba del sistema.
- ✓ *Independent – supervised*: Este enfoque es muy similar al anterior con la salvedad de que el test se realiza en la localización del evaluador sobre el

hardware del investigador. En este test el evaluador puede mejorar el control de la evaluación pero no hay forma de comparar la eficiencia computacional, algunas estadísticas de interés no se obtienen y no se puede impedir la normalización de las puntuaciones.

- ✓ *Independent – strongly supervised*: La base de datos es aislada y no se comparte hasta la conclusión del test. Los componentes software se amoldan para dar unos protocolos de entrada/salida que se testean en la zona el evaluador sobre su hardware. Los algoritmos son testeados sobre un entorno completamente controlado donde las operaciones de entrada/salida son estrictamente monitorizadas.

### 2.3.3 Tipos de Errores

No existen sistemas biométricos que realicen perfectas comparaciones [P. Rosenzweyz et al, 2004]. El rendimiento de los sistemas biométricos también es medido, cuantificado y expresado mediante una serie de indicadores [Dessimoz et al, 2005]. Muchos de estos indicadores son utilizados para medir el rendimiento de las operaciones de los sistemas biométricos (véase 2.2.1).

Durante el proceso de alistamiento pueden aparecer una serie de dificultades a la hora de capturar los datos biométricos. Dificultando la creación de modelos de usuarios en el sistema biométrico [R. L. van Renesse, 2006]. Las medidas de rendimiento que reflejan estos problemas son [Anil k. Jain et al, 2004a]:

- ✓ *FTC (Failure to Capture) o FTA (Failure to Aquisition)*: Es el porcentaje de veces que el sistema no puede obtener una muestra biométrica de suficiente calidad.
- ✓ *FTE (Failure to Enroll) o FER (Failure to Enroll Rate)*: Se define como el porcentaje de veces que el sistema no ha podido crear un modelo de usuario de suficiente calidad.
- ✓ *TTE (Time to Enroll)*: Es definido como el tiempo transcurrido desde que se capturan la información biométrica hasta que se crea el modelo.

La captura de los datos se ve influenciada por la calidad del sensor, cambios en las características psíquicas y de comportamiento del individuo, las condiciones ambientales en el momento de la captura. Todos estos factores pueden llevar a errores en la identificación o reconocimiento de un individuo. Debido a esta peculiaridad, se establece un umbral que cuantifica la similitud entre los datos de entrada y el modelo de usuario almacenado (Figura 4). Como consecuencia, se producen los siguientes tipos de errores de decisión:

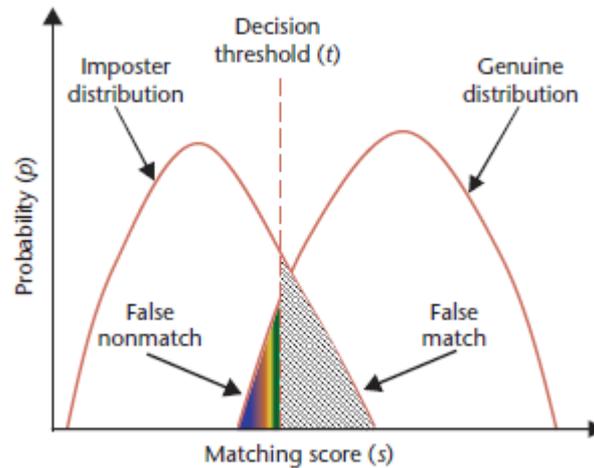


Figura 4: Umbral de decisión

- ✓ *FRR (False Rejection Rate) o Tasa de Error de Tipo I*: Se define como la probabilidad de que se rechace a un usuario legítimo por ser el sistema incapaz de identificarle.
- ✓ *FAR (False Acceptance Rate) o Tasa de Error de Tipo II*: Es la probabilidad de que se autentique a un usuario ilegítimo.

Estos dos errores dan lugar a la aparición de estos otros errores de comparación [S. Prabhakar et al, 2003]:

- ✓ *FMR (False Match Rate) o Falso Positivo*: Se define como la frecuencia con la que se produce FAR.
- ✓ *FNMR (False Non-Match Rate) o Falso Negativo*: Es la frecuencia con la que se produce FRR.

Existen otras medidas que pueden ser usadas como una medida general de las prestaciones de un sistema biométrico:

- ✓ *EER (Equal Error Rate)*: Hace referencia al valor del umbral de decisión para el que los valores de FAR y FRR son iguales (Figura 5).
- ✓ *HTER (Half Total Error Rate)*: Es la semisuma de los valores de FAR y FRR.
- ✓ *Coste*: Es la suma ponderada de los valores de FAR y FRR. En este cálculo se asignan costes a los diferentes tipos de errores del sistema.

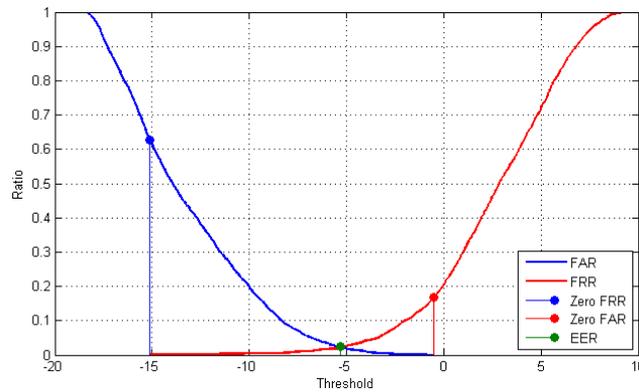


Figura 5: Representación de FRR, FAR y ERR

Una forma alternativa de evaluar las prestaciones de un sistema biométrico es mediante la curva ROC (Receiver Operating Curve) [Fogarty et al, 2005] o mediante la curva DET (Detection Error Trade-Off). La curva ROC (Figura 6) consiste en la representación gráfica del ratio de falsa aceptación (FAR) frente al ratio de correcta aceptación (1-FRR) en función del umbral de decisión. Cuanto mayor sea el área por debajo de la curva ROC mayor confianza posee el sistema biométrico, lo ideal es que el área sea igual a la unidad. Por otro lado, la curva DET (Figura 7) consiste en la representación grafica de FNMR frente a FMR. Una de las principales ventajas de estas curvas es la posibilidad de contrastar las prestaciones de diferentes sistemas en un solo gráfico.

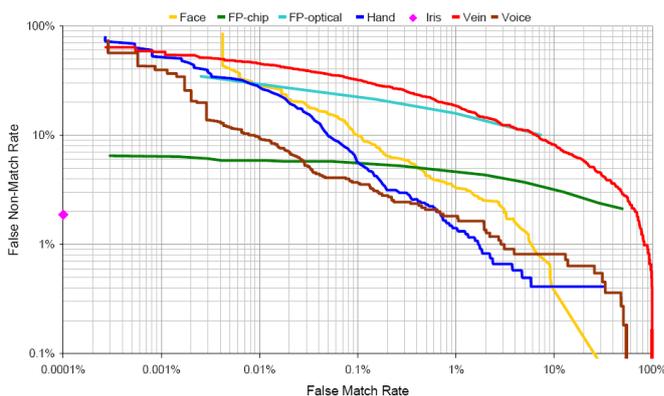


Figura 7: Curva DET para diferentes rasgos biométricos

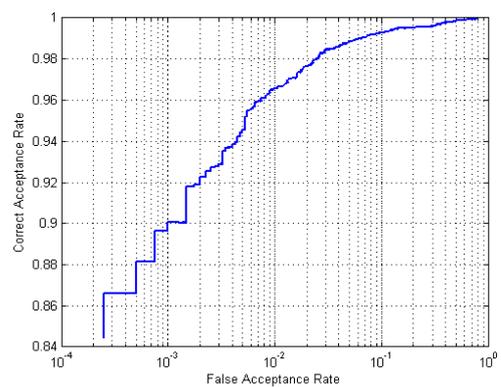


Figura 6: Curva ROC

## 2.4 BIOMETRÍA UNIMODAL

El mundo de la biometría abarca un amplio abanico de tecnologías diferentes. En este apartado se intentara estudiar algunas de las tecnologías biométricas unimodales más conocidas.

### 2.4.1 Firma manuscrita

La firma manuscrita es un rasgo que plasma el comportamiento de una persona. Está caracterizado por movimientos rápidos de la mano debidos a la actividad del sistema neuromuscular. Se puede trabajar con una imagen previamente guardada, llamada firma estática u off-line. O por el contrario, con la información instantánea (trayectoria, presión, velocidad, fuerza) recogida mediante algún sensor (tableta de firmas), llamada firma dinámica u on-line, ambas recogidas en [V. S. Nalwa, 1997]. La Figura 8 recoge el proceso de verificación on-line.

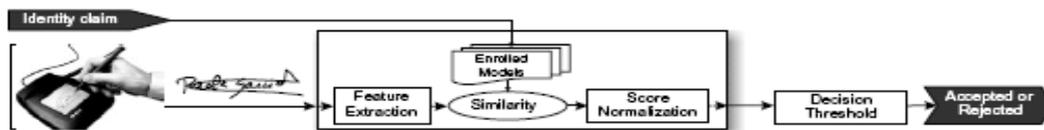


Figura 8: Arquitectura típica de verificación on-line

Los principales métodos para realizar la verificación de forma on-line se basan en los algoritmos Dynamic Time Wrapping (DTW) que se pueden encontrar en [Y. Sato et al, 1982] [Hao Feng et al, 2003], Hidden Markov Model (HMM) [J. Fierrez et al, 2007] con distribuciones de probabilidad gaussiana (GMM) [J. Richiardi et al, 2003] y redes neuronales (NN) descritas en [H.-D. Chang et al, 1993]. También se han utilizado, aunque con peores resultados que los anteriores, la distancia Euclídea de la manera descrita en [R. S. Kashi et al, 1996] [T. Rhee et al, 2001] y la correlación global y local [V. S. Nalwa, 1997]. Los sistemas que trabajan con firma offline pueden extraer las características estáticas mediante la geometría de la imagen y verificar usando NN [Huang et al, 1997]. También pueden utilizar HMM y la Discrete Radon Transform (DRT) tal como describe [J. Coetzer et al, 2004]. El inconveniente de este tipo de biometría es la variabilidad temporal entre firmas y se pueden producir falsificaciones entrenadas.

## 2.4.2 Geometría Facial

Probablemente, el reconocimiento facial sea el método natural más utilizado con el que identificamos a las personas. Se pueden obtener características biométricas sin que el individuo se percate de ello (método no intrusivo). En la actualidad, la mayoría de los sistemas actuales de reconocimiento facial determinan la apariencia de un sujeto a través de la obtención de puntos especialmente significativos de la cara (llamados puntos nodales) mediante información 2D o 3D [C. Conde et al, 2008]. A partir de estos puntos, como describe [W. Zhao et al, 2003] se determina 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. La Figura 9 recoge el proceso de extracción de características de la biometría facial.

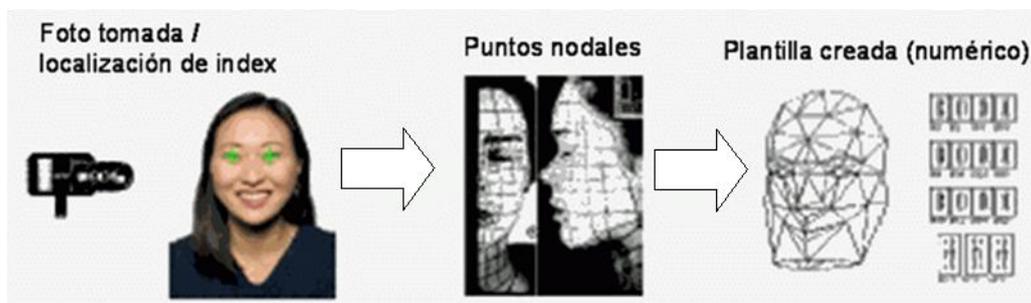


Figura 9: Extracción de Características de la Biometría Facial

Existen múltiples formas de realizar reconocimiento facial pero las más implantadas se basan en la obtención de características geométricas de la cara. Uno de los métodos más populares (Land Marks) se basa en la localización y forma de distintas características biométricas de la cara como la distancia entre los ojos o la anchura de la nariz. Los métodos basados en la correlación espacial establecen un indicador de parecido mediante la comparación de la correlación entre dos imágenes. Otros métodos, los descritos en [M. Padilla et al, 2003], se basan solo en el uso de características estáticas de la imagen (tamaño, color, forma...). Finalmente, la técnica de Eigen Faces codifica una imagen facial de tal forma que puede ser representada como la combinación lineal de otras imágenes [L. Sirovich et al, 1987]. Esta técnica puede ir aplicada con el algoritmo Hill-Climbing (descrito en [J. Galbally et al, 2009]) mejorando el resultado.

Mientras que las técnicas de correlación imponen un elevado coste computacional, Eigen Faces permite optimizar el sistema y aumenta la velocidad de procesado. Como desventaja resulta complicado reconocer una cara en una imagen tomada en diferentes puntos de vista [K. Niinuma et al, 2009], también es importante la iluminación con que es tomada la imagen y el fondo [Anil K. Jain et al, 2004a].

### 2.4.3 Huella Dactilar

El reconocimiento biométrico por huella dactilar es la tecnología más implantada y evolucionada. Aunque tienen un amplio abanico de uso, tradicionalmente se utiliza en la ciencia forense. Este tipo de tecnología posee una mayor aceptabilidad entre la población. Está basada en los estudios realizados por Faulds, Herchel y Galton durante el siglo XVIII. Estos métodos se basan en la comparación de unas características extraídas de las huellas llamadas minucias. Una huella dactilar está formada por un conjunto de crestas y valles característicos para cada individuo. Las minucias son el resultado del análisis de las crestas y valles que conforman la huella. Por tanto, para cada individuo solo se almacena su patrón de puntos de minucias. Previa a la identificación de usuarios se realiza una preclasificación de las huellas con el fin de reducir el número de búsquedas exhaustivas (diferente clasificación dependiendo del país, todas ellas recogidas en [F. M. Rodríguez et al, 2004]). La Figura 10 ofrece de manera abreviada el proceso de extracción de características.

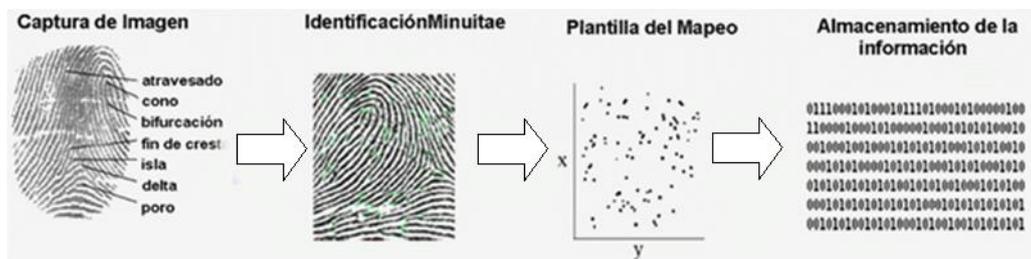


Figura 10: Extracción de Características de la Huella Dactilar

La extracción de características se basa en el estudio de una imagen en escala de grises. Para este estudio existen dos procedimientos: Por un lado está el descrito en [K. C. Chan et al, 2003] para sistemas basados en correlación, donde se superponen dos imágenes de la huella dactilar y se calcula la correlación entre ambas para diferentes posiciones espaciales mediante rotación y translación de las mismas [Anil K. Jain et al, 2000] [Koichi Ito et al, 2004]. Por otro lado, se encuentra el descrito en [N. Yager et al, 2004] para los sistemas basados en la extracción de puntos característicos. En ellos se segmenta la imagen (principalmente por filtros Gabor descritos en [F. M. Rodríguez et al, 2004] y DFT descrita en [Koichi Ito et al, 2005]) y se localizan las minucias (bifurcaciones, fines de cresta...) cuya disposición espacial ofrece una elevada capacidad de diferenciación. En ambos métodos, según [J. González et al, 2004], la decisión se toma mediante aproximaciones Bayesianas.

## 2.4.4 Iris

El iris es la región anular del ojo que se encuentra limitada por la pupila en la parte interior y la esclera o esclerótica por el exterior. Al igual que la huella, sus patrones se forman durante el desarrollo del feto y se estabilizan a los dos años. La compleja textura del iris ofrece una información muy discriminativa de cada individuo. Aunque es una tecnología moderna [J. Daugman, 1992] ya existen varias aproximaciones al problema del reconocimiento por iris. Todas ellas se basan en la extracción de características globales del iris o bien de ciertas regiones locales [Unsang Park et al, 2009].

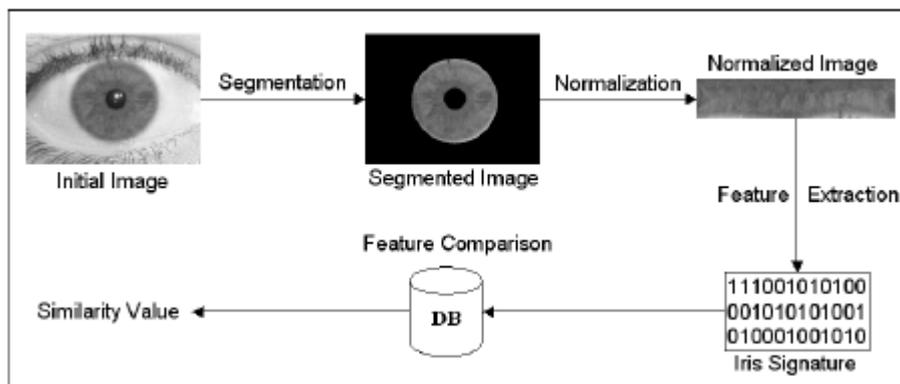


Figura 11: Fases del Proceso de Reconocimiento de Iris

El reconocimiento biométrico a través del iris según [Shrikanth Mohan, 2004] se realiza mediante una serie de pasos: la adquisición de la imagen, el pre-procesado de la misma, la localización del iris, normalización de la imagen, extracción de características y la comparación (Figura 11). Algunos sistemas, según [Hasan Demirel et al, 2009] utilizan los métodos de ajuste basados en modelos de histogramas. Estos procesan la imagen del iris mediante filtros paso-banda de dos dimensiones y utilizan el discriminante lineal de Fisher para realizar el proceso de comparación. Otros sistemas emplean la transformada de Hough para localizar el iris y mediante la transformada de Hilbert [C. L. Tisse, 2003] o una Laplaciana piramidal [R. P. Wildes, 1994] obtienen las características de éste. Otros sistemas realizan un cambio a coordenadas polares y mediante el algoritmo del gradiente obtienen información discriminante de los límites del iris (radio y centroides) [Amjad Zaim, 2007]. El último tipo de sistemas utilizan transformadas de Hough circulares para localizar el iris y filtros de Gabor para extraer el conjunto de características y realizan la comparación mediante el cálculo de la distancia de Hamming (HD) [C. L. Tisse et al, 2002] [Amjad Zaim, 2007].

### 2.4.5 DNA

El DNA (DeoxyriboNucleic Acid) podría ser el código único definitivo para cada individuo si no existiera la posibilidad de gemelos idénticos, los cuales poseen idénticos patrones de DNA. Es una característica unidimensional y es mayoritariamente usada en el contexto forense. Según [Anil K. Jain et al, 2004a] existen problemas que limitan el uso de esta biometría de aplicación. Por un lado, es muy fácil robar una muestra de DNA de un sujeto concreto y suplantar así su identidad. Para realizar una autenticación se deben de realizar complicados métodos químicos limitando su identificación en tiempo real. Por otro lado, no conviene que el código genético sea de dominio público.

### 2.4.6 Huella de la palma de la mano

La palma de la mano contiene una gran cantidad de crestas y valles al igual que las huellas dactilares. La zona de la palma de la mano es mucho mayor que la de los dedos por lo que puede ofrecer mayor potencial discriminativo. El reconocimiento biométrico basado en la huella de la palma de la mano es muy similar al de huella dactilar. Se apoya igualmente en la información local o global representada por los valles y crestas de la superficie de la palma, como bien recoge [Jane You et al, 2004] [José García-Hernández et al, 2005]. La Figura 12 recoge de manera global la extracción de características descrito anteriormente. Por otro lado, también se puede obtener la información biométrica mediante el método utilizado por [Guangming Lu et al, 2002] llamado Eigen Palms y es similar al Eigen Faces (véase sección 2.4.2). Este método consiste en la representación de una imagen como combinación lineal de otras.

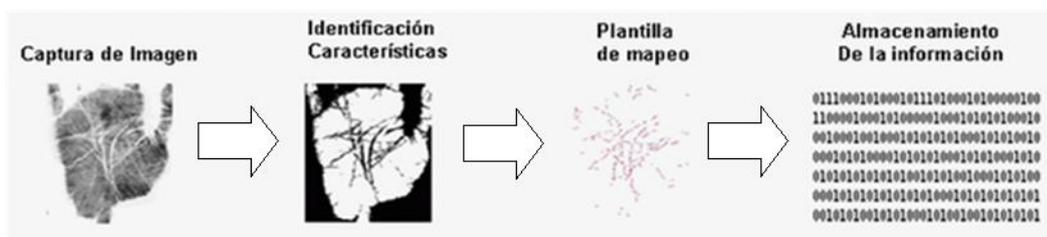


Figura 12: Extracción de Características de la Palma de la Mano

### 2.4.7 Dinámica del Teclado

Es un tipo de rasgo biométrico conductual empleado en la verificación de la identidad de un individuo mediante su cadencia de escritura en un teclado [Alen Peacock et al, 2005]. Esta tecnología se sostiene sobre la premisa de que cada individuo exhibe un patrón y una cadencia distintivos cuando pulsa una combinación de teclas [F. Monroe et al, 2000].

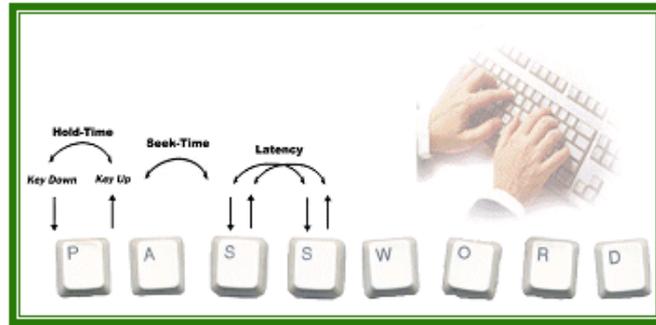


Figura 13: Características de la Dinámica del Teclado

En la mayoría de los sistemas se emplea la latencia (tiempo transcurrido entre la pulsación de dos teclas) entre pulsaciones como característica; sin embargo, como dice [O. Coltell et al, 2006] otros utilizan también el tiempo que permanecen la teclas presionadas. Esta tecnología no requiere hardware o dispositivos adicionales ya que se soporta sobre software de captura. Los principales sistemas utilizan clasificadores bayesianos, redes neuronales y lógica difusa.

### 2.4.8 Movimiento (caminar)

Es un tipo de rasgo biométrico conductual. Se emplea en la verificación de un individuo mediante un examen espacio temporal de su patrón de marcha (Figura 14). Los sistemas se pueden clasificar estáticos (basados en la silueta humana [Jeff P. Foster et al, 2003]) o dinámicos (basados en el movimiento del individuo [Liang Wang et al, 2003]).



Figura 14: Extracción de Características de la Silueta

### 2.4.9 Voz

El reconocimiento biométrico mediante voz se basa en unas características tanto físicas del individuo como de comportamiento del mismo [Elizabeth Shriberg et al, 2004]. Utilizando la clasificación de [D. A. Reynolds et al, 1995] existen dos tipos de reconocimiento de locutor: unos, basados en el conocimiento de una frase pre-

establecida (text-dependent). Otros, no imponen ningún tipo de restricción de texto (text-independent). La Figura 15 muestra una visión general del proceso de extracción de características.

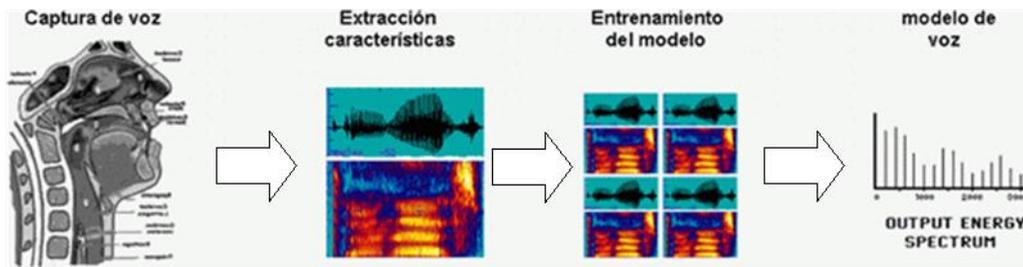


Figura 15: Extracción de las Características de la Voz

A lo largo de los años se ha ido mejorando la técnica para realizar el reconocimiento de locutores mediante la introducción de diferentes algoritmos [Dessimoz et al, 2005]. Se han ido utilizando los métodos probabilísticos para modelar el comportamiento de la señal de voz, entre ellos se encuentran el HMM [A. Poritz, 1982] y el GMM [D. A. Reynold et al, 1995]. Otros métodos de clasificación utilizados para el reconocimiento de locutor son NN [J. Oglesby et al, 1988] [Anil K. Jain et al, 2004b], el SVM [M. Schmidt et al, 1996], DTW [M. Pandit et al, 1998] y el perceptrón multicapa (MLP) [C. E. Vivaracho et al, 2006]. O combinación de clasificadores y modeladores en la etapa de entrenamiento [Todor Ganchev et al. 2002].

## 2.5 BIOMETRÍA MULTIMODAL

Según [Anil K. Jain et al, 2004a] los sistemas biométricos que trabajan con un sólo rasgo biométrico (unimodales, véase sección 2.4) poseen las siguientes limitaciones:

- ✓ *Ruido*: La información capturada puede incorporar ruido o algún tipo de distorsión. Estos hechos son el resultado de la utilización de sensores defectuosos y de condiciones ambientales adversas.
- ✓ *Variación intra-clase*: Los datos capturados de un individuo en la operación de autenticación pueden ser muy diferentes de los empleados en la generación de su modelo durante la operación de alistamiento.
- ✓ *Variación inter-clase*: La escasa variación de un rasgo biométrico entre los sujetos de una población limita su capacidad de discriminación.
- ✓ *Universalidad*: Es posible que un determinado subconjunto de usuarios no posea un rasgo biométrico particular.

- ✓ *Ataques de impostores*: Un usuario ilegítimo puede falsificar el rasgo biométrico de otro legítimo para eludir el sistema.

Las limitaciones impuestas por los sistemas unimodales pueden atenuarse usando múltiples modalidades biométricas [L. Puente et al, 2008]. Por ejemplo, el requisito de universalidad (véase sección 2.1) ya no es imprescindible. El uso de múltiples modalidades biométricas también puede conseguir un aumento de la seguridad del sistema ante ataques de impostores [B. Yanikoglu et al, 2004]. En definitiva, tal como describe [Hong et al, 1999] los sistemas biométricos multimodales utilizan varios rasgos biométricos para establecer el reconocimiento de individuos.

Según [Anil K. Jain et al, 2004a] los sistemas biométricos multimodales pueden operar en tres diferentes modos:

- ✓ *Serial Mode (Modo Consecutivo)*: La información de un rasgo biométrico es utilizado para reducir el número de posibles identidades antes de la entrada del nuevo rasgo biométrico.
- ✓ *Parallel Mode (Modo Paralelo)*: La información de múltiples rasgos biométricos es utilizada para realizar el reconocimiento.
- ✓ *Hierarchical Mode (Modo Jerárquico)*: En este modo, se combinan clasificadores individuales en una estructura de árbol.

Estos sistemas biométricos presentan información tomada de múltiples indicadores biométricos. Para conseguir el funcionamiento del sistema en los modos descritos anteriormente se deben combinar los sistemas biométricos (o fusionar). Existen tres niveles de fusión [Anil K. Jain et al, 2004a]:

- ✓ *Fusión a nivel de extracción de características*: En este nivel se combinan las características extraídas de los rasgos biométricos en un único vector.
- ✓ *Fusión a nivel de puntuaciones*: En este nivel se combinan las puntuaciones que describen las similitudes entre los datos adquiridos y los modelos de cada sistema. Este modo de fusión requiere que las puntuaciones de los subsistemas se encuentren normalizadas en un dominio común [Anil K. Jain et al, 2005]. La fusión a este nivel es la preferida debido a la facilidad de acceso y de combinación que ofrecen las puntuaciones.
- ✓ *Fusión a nivel de decisiones*: En este nivel se combinan las decisiones tomadas por cada sistema con el fin de obtener la decisión final.

Los sistemas biométricos multimodales están diseñados para trabajar de cinco formas diferentes (Figura 16) y aunque algunas de ellas no implican múltiples modalidades, implican la fusión de información en algunos puntos [Anil K. Jain et al, 2004a]. Los esquemas de operación son los siguientes:

- ✓ *Un rasgo, múltiples sensores:* Se adquiere un único rasgo mediante diferentes sensores y se combina para mejorar el proceso de reconocimiento.
- ✓ *Múltiples rasgos:* Se obtienen y combinan diferentes características biométricas de una persona.
- ✓ *Un rasgo, múltiples unidades:* Se adquieren y combinan diferentes unidades del mismo rasgo.
- ✓ *Un rasgo, múltiples representaciones:* Se adquieren múltiples tomas de un mismo rasgo por el mismo sensor.
- ✓ *Un rasgo, múltiples clasificadores:* Se obtiene un rasgo a través de un único sensor y mediante diferentes técnicas de extracción de características y/o de comparación de patrones se realiza el proceso de reconocimiento.

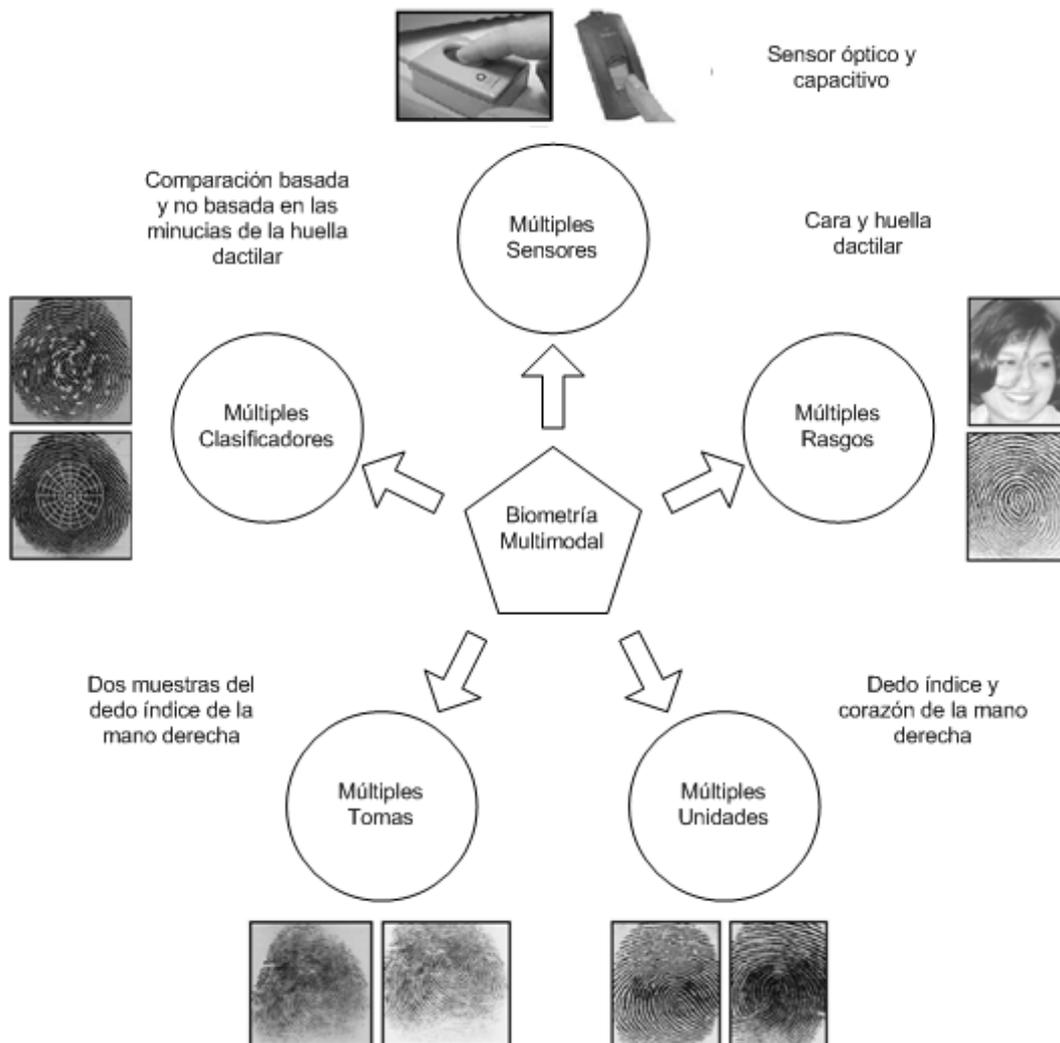


Figura 16: Varios Escenarios de un Sistema Biométrico Multimodal (extraído de [Anil K. Jain et al, 2004a])



---

# CAPÍTULO 3

---

## RECONOCIMIENTO DE LOCUTOR

En este capítulo, se presenta un acercamiento al lector en el proceso de reconocimiento de locutor y la manera de resolverlo en la actualidad. Por disponibilidad de bases de datos se centró en la tecnología biométrica de reconocimiento de locutor y se creyó conveniente centrarlo en el proceso de identificación.

### 3.1 INTRODUCCIÓN

La voz humana se caracteriza por englobar diferentes tipos de información. Por un lado, se tiene el significado de la palabra (lo que se intenta transmitir al oyente). Por otro lado, se tiene la información que incluye el habla (como las emociones del hablante, el género, el lenguaje, la identidad del hablante). La identificación biométrica de locutor extrae, clasifica y reconoce la información sobre la identidad del hablante contenida en una muestra de su voz [D. A. Reynolds, 2002].

El reconocimiento biométrico de locutor puede realizarse de dos maneras dependiendo del contexto de actuación, estas son: verificación de locutor e identificación de locutor (véase, sección 2.2.1). En la verificación se corrobora la identidad pretendida por una persona mediante su voz. En la identificación no hay identidad pretendida y el sistema decide de qué identidad se trata de todas las registradas en el sistema.

La identificación biométrica puede ser sobre conjunto cerrado (closed-set) ó conjunto abierto (open-set). La identificación en closed-set decide cual de los usuarios registrados podría ser el autor de la muestra biométrica. Mientras que en la identificación en open-set se decide a cuál de los usuarios registrados podría pertenecer la muestra o bien se concluye que la muestra es desconocida. Según [D. A. Reynold et al, 1995] dependiendo del algoritmo usado en la identificación la tarea se puede dividir en: identificación dependiente del texto (text-dependent) o independiente del texto (text-independent). En el primer caso, el sistema reconoce al locutor dependiendo del texto (normalmente un texto pregrabado). Mientras que en el segundo caso, el sistema debe ser capaz de identificar al locutor independientemente del texto [D. T. Toledano et al, 2008]. La Figura 17 recoge un esquema del proceso biométrico descrito.

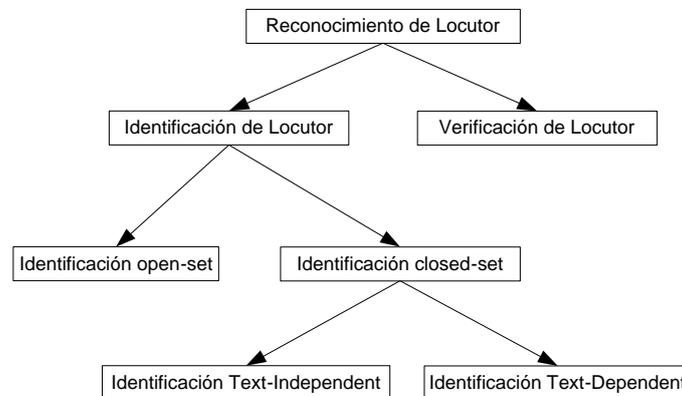


Figura 17: Diagrama del Reconocimiento de Locutor

El proceso de identificación se divide en dos fases. En la primera fase (fase de alistamiento) se obtiene la muestra de los hablantes para crear sus modelos. Estos modelos se almacenan en el sistema. En la segunda fase (fase de identificación) se extrae la muestra del hablante y se compara con los modelos almacenados en el sistema (1:N comparaciones). La Figura 18 recoge gráficamente ambas fases.

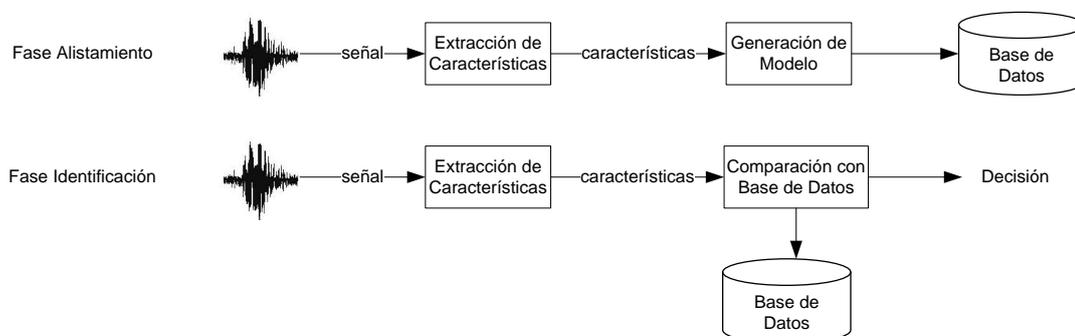


Figura 18: Fase Entrenamiento/Identificación

## 3.2 EXTRACCIÓN DE CARACTERÍSTICAS

Las propiedades acústicas de la señal de voz contienen información diversa sobre el locutor. Como se puede ver en [J. M. Naik, 1990] esta información contiene propiedades de alto nivel tales como el contexto, dialecto, el estilo de habla y el estado emocional entre otras. Y propiedades de bajo nivel como el pitch, la intensidad, el formato de la frecuencia y su ancho de banda, la correlación espectral, el espectro a corto plazo entre otros [B. S. Atal, 1976]. Estas últimas propiedades son cuantificables y son utilizadas para extraer características discriminantes de la señal de voz.

Las características que se suelen extraer en el reconocimiento de voz están relacionadas con el espectro instantáneo de la señal de voz o la forma instantánea del tracto bucal. El proceso de extracción de características consigue una representación mensurable de la señal de voz en parámetros discriminantes de la misma [Frederic Bimbot et al. 2004]. Esta transformación permite obtener una representación más compacta, menos redundante y que puede ser comparada con un *score*. Según [B. S. Atal, 1976] la elección de los parámetros es un factor crucial por lo que es aconsejable que sean discriminativos entre locutores, fácilmente mensurables, estables con el tiempo, que ocurran de forma natural y frecuente en el habla, que cambien muy poco independientemente del entorno... siendo la precisión de reconocimiento altamente dependiente del tipo y número de parámetros usados.

### 3.2.1 Coeficientes Cepstrales

La señal de voz varía muy poco en intervalos cortos de tiempo. Se dice que es una onda quasi-estacionaria. Por tanto, cuando es examinada en un periodo de tiempo (20-30 milisegundos) suficientemente corto se consiguen unas características acústicas estables [J. R. Deller et al, 2000].

La señal de audio puede representarse con una rápida variación en el torrente de datos convolucionada con una variación lenta representada por el tracto vocal (Figura 19a). Convendría que estas dos características se pudieran separar y la única forma de hacerlo es mediante una representación cepstral. La representación cepstral realiza la separación de los componentes anteriores en una suma de partes (Figura 19b), mediante la DFT inversa del logaritmo de la magnitud del espectro de cada frame [J. R. Deller et al, 2000].

Para el cálculo de los componentes cepstrales se preestablece una ventana de longitud (20-30 milisegundos) y se va moviendo a lo largo de la señal produciendo solapamiento entre ventanas cercanas (análisis short-term). A continuación se procesa el trozo de señal residente en cada desplazamiento de la ventana que dará lugar a los cepstrum. Existen múltiples tipos de ventana, las más usadas en voz son la Hamming y la Hanning descritas en [S. Furui, 2001].

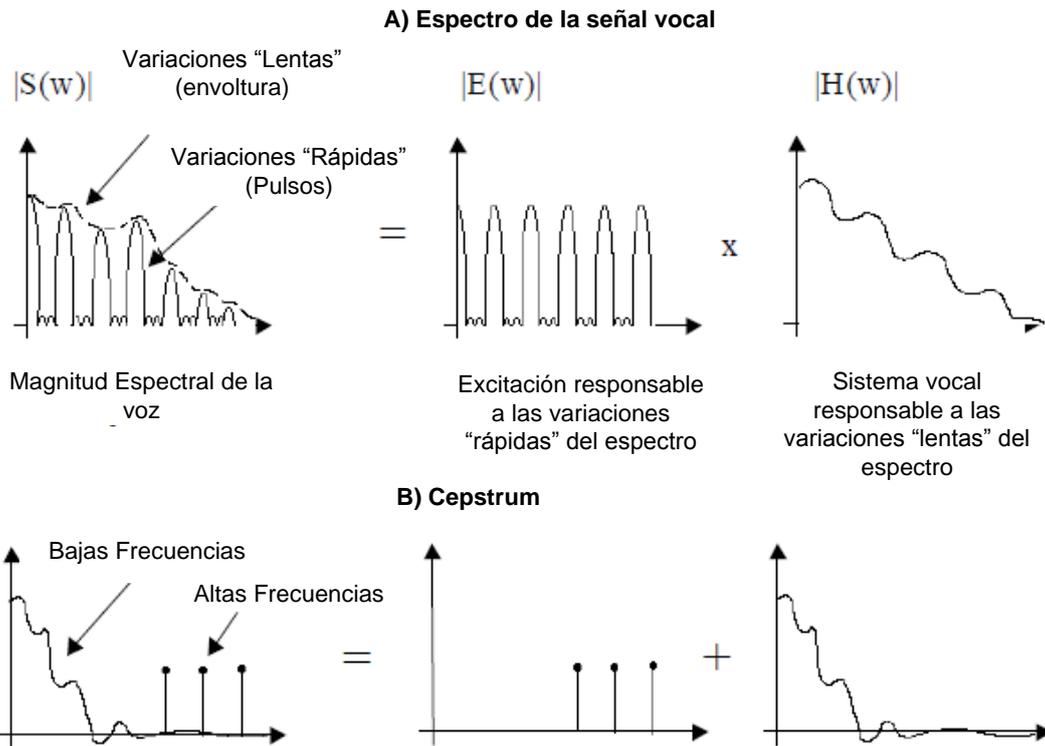


Figura 19: Espectro de la señal vocal/Cepstrum

Existen diferentes maneras de obtener los coeficientes cepstrales. Por un lado, se realiza mediante bancos de filtros triangulares (coeficientes mel-cepstrales). Por otro lado mediante predicción (LPC) de los parámetros.

**Coficientes cepstrales basados en la frecuencia Mel o Mel-Cepstrales**

La información que posee la señal de voz en bajas frecuencias es fonéticamente más importante para los humanos que la existente en las altas frecuencias. Para realizar un mayor hincapié en las bajas frecuencias se debe introducir un paso previo a la DFT inversa llamado escalado mel [J. R. Deller et al, 2000]. El mel es una unidad especial de medida o escala de percepción del pitch del tono. La escala de mel se basa en el estudio psicológico de la percepción humana para diferentes frecuencias de sonido.

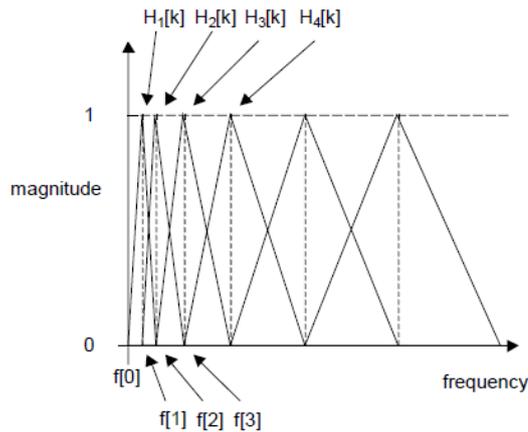


Figura 20: Filtros Triangulares para obtener Mel-Cepstrum

Se puede crear un espectro de mel utilizando un banco de filtros (suele ser filtros triangulares), uno para cada frecuencia de la escala mel (Figura 20). Como paso final se realiza la DFT inversa o bien, se puede sustituir por una DCT (Transformada Discreta del Coseno) para mejorar el rendimiento computacional. El número de coeficientes mel cepstrales es relativamente bajo, ronda los 12 ó 20. Normalmente los coeficientes mel cepstrales no son relativamente importantes en la identificación [Z. Bin et al, 2001] debido a su elevado coste computacional. La Figura 21, extraída de [Frederic Bimbot et al, 2004], recoge el proceso de extracción de los coeficientes mel-cepstrales.

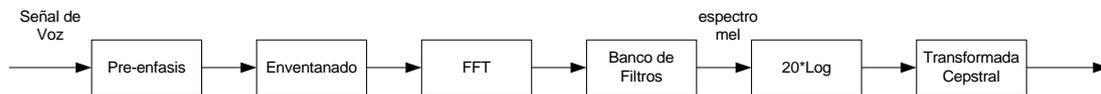


Figura 21: Proceso de Extracción de los Coeficientes mel-cepstrales

### Coefficientes cepstrales basados en LPC (Linear Predictive Coding)

El análisis LPC se basa en el modelo lineal de la producción de voz. Este modelo utiliza ARMA (Auto Regresive Moving Average) que se simplifica en AR (Auto Regresive) [G. Fant, 1970].

El aparato fonador humano se suele describir como la combinación de cuatro módulos: el sonido glotal, el tracto vocal, el tracto nasal y los labios. Cada módulo puede ser representado por un filtro según [Frederic Bimbot et al, 2004]. Por tanto, el aparato fonador se puede representar por un filtro ARMA. La caracterización de la señal de voz (normalmente una porción enventanada, véase sección 3.2.1) es equivalente a determinar los coeficientes del filtro.

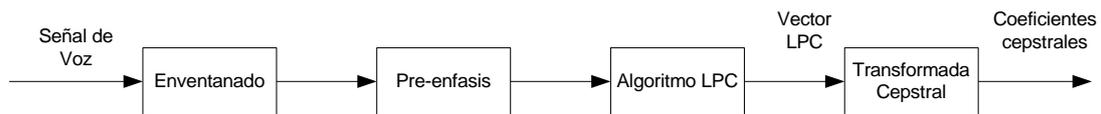


Figura 22: Proceso de Extracción de los Coeficientes cepstrales mediante LPC

La idea principal del análisis LPC es estimar los parámetros de un filtro AR enventanando una porción de la señal. A continuación, se mueve la ventana y se calcula una nueva estimación. Para cada ventana, se estiman un conjunto de coeficientes (llamados coeficientes predictivos o coeficientes LPC) y pueden ser usados como un vector de parámetros. Finalmente, el espectro contenido en la ventana puede estimarse a través de los coeficientes LPC. A partir estos coeficientes es posible calcular directamente los coeficientes cepstrales de la manera descrita en [D. Petrovska-Delacrétaz et al, 2000]. Con este método se obtiene una representación de la señal de voz como una combinación lineal de muestras pasadas [L. Rabiner et al, 1993]. La Figura 22 muestra la extracción de los coeficientes cepstrales mediante el análisis LPC.

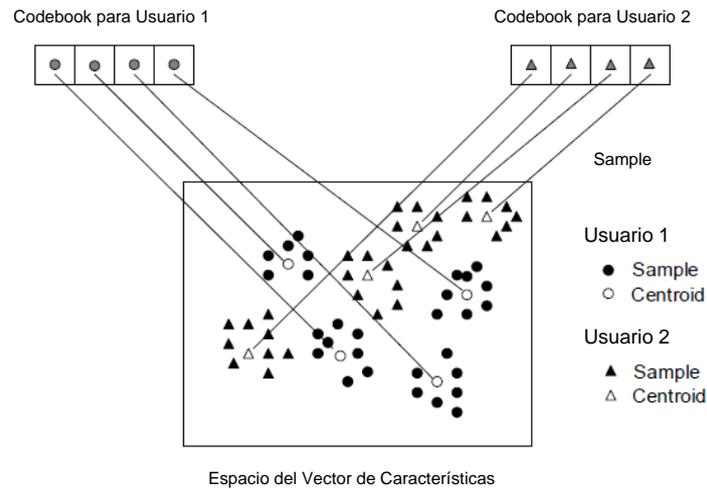
### 3.3 ALGORITMOS: CLASIFICACIÓN Y MODELADO DE CARACTERÍSTICAS

El objetivo primordial del modelado de características es obtener una representación medible y comprable del usuario. Una vez obtenido una representación medible se puede clasificar mediante algún algoritmo de clasificación. Estos algoritmos obtienen una puntuación que determina la similitud de la muestra desconocida con respecto a los modelos almacenados.

Los algoritmos más conocidos para modelado de características son el GMM, HMM y el VQ (Vector Quantization), también llamados generativos. También se suele usar el algoritmo GA (Genetic Algorithm) detallado en [D. E. Goldberg, 1989] para establecer la arquitectura óptima de HMM y GMM. Para clasificación (o comparación), los más conocidos son las NN, SVM y VQ, también llamados discriminativos.

#### 3.3.1 Vector Quantization (VQ)

Este algoritmo mapea vectores de un espacio vectorial a un número finito de regiones en ese espacio. Esas regiones se llaman *clusters* y son representadas por su vector central llamado *centroide*. A un conjunto de centroides (que representan el completo espacio vectorial) se le llama *codebook*. Como resultado, cada usuario queda representado por el *codebook*. La Figura 23 recoge la formación del *codebook*.



**Figura 23: Representación Espacial del Codebook**

Actualmente, existen discrepancias entre autores con respecto a este tipo de aproximación. Unos autores defienden que el algoritmo VQ ignora las variaciones temporales de la señal, ya que usa simplemente aproximaciones globales (*centroides*), entre ellos se encuentra [J.P. Campbell, 1997]. Otros autores consideran que es un método probabilístico o estocástico, porque utiliza los *centroides* para estimar la forma de la distribución de probabilidad, entre ellos se encuentran [H. Gish, 1994] [S. Ong et al, 1996].

A la hora testear utilizando el VQ, se compara el vector de características de la voz desconocida con las distancias a los diferentes *codebook*. Se establece como identidad aquel modelo cuya distancia acumulada sea mínima. Las distancias más utilizadas son la Euclídea, Mahalanobis y Manhattan [Evgeny Karpov, 2003]. El algoritmo más conocido que utiliza esta técnica es el *k-means*.

### 3.3.2 Hidden Markov Model (HMM)

Es el modelo mayoritariamente usado en reconocimiento de locutor y está basado en el uso de modelos estocásticos [L. R. Rabiner, 1999]. Cada palabra de un locutor determinado es generada por un modelo de Markov. Este modelo es una serie finita de estados interconectados por probabilidades de transición (Figura 24). Por un lado, cada uno de los vectores de características, generados por HMM, tiene cierta probabilidad de mantenerse en el estado actual o avanzar al siguiente. Por otro lado, cada uno de los estados tiene una probabilidad de presenciar una cierta observación (observar un vector de características). En este caso, solo son visibles las observaciones, desconociéndose la secuencia de estados.

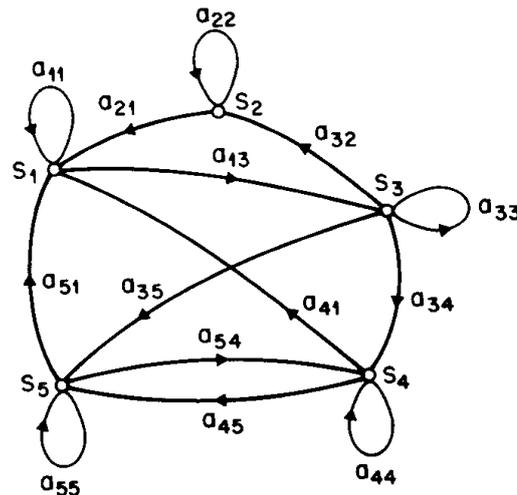


Figura 24: Transiciones de estado del Modelo de Markov de 5 Estados

A la hora de testear, según [Unathi Mahola et al. 2004], la probabilidad del vector de características generado por HMM, se compara usando la transición de probabilidades entre estados y la observación de probabilidades de un vector de características en un estado dado. Este método es muy utilizado para el reconociendo de locutor dependiente del texto (véase sección 3.1). Esto se debe a la existencia de algoritmos elegantes y eficientes tanto en entrenamiento como en reconocimiento.

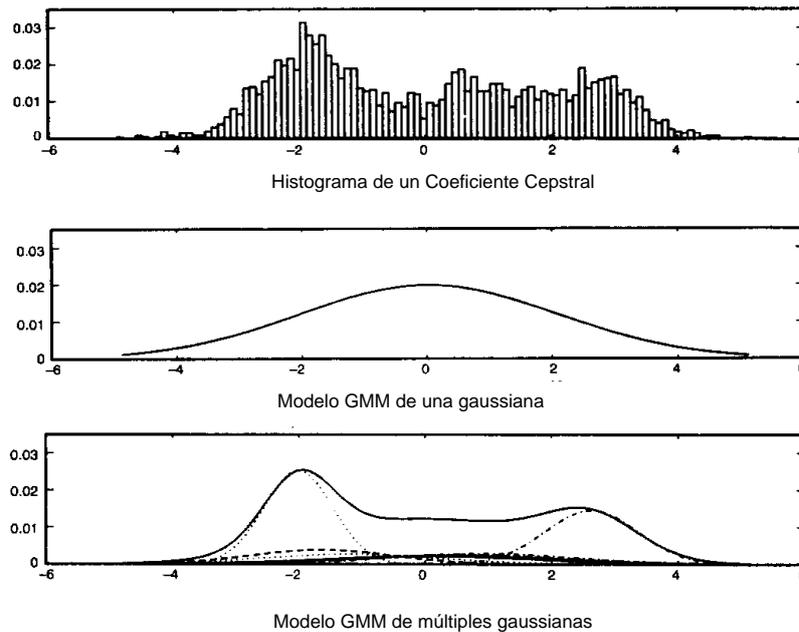
Estos modelos se utilizan en dos niveles. Por un lado, para modelar palabras mediante la utilización de modelos acústicos. Por otro, para modelar el lenguaje, definiendo una red de reconocimiento en la que se imponen las restricciones gramaticales, a partir de las cuales se realiza la decodificación de la secuencia de palabras más probable utilizando la información del nivel acústico. En cualquier caso, la utilización de los *HMM* en el reconocimiento de voz se aplica teniendo en cuenta dos suposiciones:

- ✓ La voz se puede dividir en segmentos (estados) en los que la señal de voz puede considerarse prácticamente estacionaria.
- ✓ La probabilidad de observación de que un vector de características se genere sólo depende del segmento actual, y no de símbolos anteriores o posteriores. Esto recibe el nombre de supuesto de independencia.

### 3.3.3 Gaussian Mixture Models (GMM)

Este método, al igual que el HMM, también se basa en el uso de modelos estocásticos. Realiza un modelado de la señal basándose en las variaciones estadísticas de las características. Por tanto, genera una representación estadística del hablante mediante mezcla de gaussianas [D. A. Reynold et al, 1995].

El conjunto de coeficientes cepstrales de un locutor se modela con múltiples gaussianas. A mayor número de gaussianas mayor resolución (Figura 25). Existen diversos algoritmos para la generación de estas gaussianas, entre los más populares se encuentran el Expectation-Maximization (EM) definido en [D. A. Reynolds et al, 1995] y el Maximum Likelihood estimation (ML) [A. Dempster et al, 1977] (esta maximización no es posible hacerla directamente, por lo que se usa EM para obtener los parámetros). En ellos, iterativamente se refina la ubicación, anchura y máximo de las gaussianas hasta un umbral de convergencia [H. Gish et al, 1994].



**Figura 25: Ejemplo de Modelado GMM**

En la fase de test, se evalúa cada vector de características con la distribución del modelo. La puntuación obtenida será la suma de todas estas evaluaciones. A mayor cercanía entre los vectores (de características) cepstrales de la voz desconocida y el modelo, la evaluación caerá en los valores más altos de las gaussianas respectivas. Por lo tanto, su puntuación será mayor. Para dar con el umbral (puntuación) óptimo para la discriminación bastará con algunas sesiones de entrenamiento como bien describen [H. Gish et al, 1994] [D. A. Reynolds et al, 1995] [J.P. Campbell, 1997] [D. A. Reynolds, 2000].

En el caso particular del proyecto se utiliza para inicializar el GMM el algoritmo k-means (VQ). Aplicando el k-mean sobre los archivos de entrenamiento se obtienen los centroides y el número de vectores pertenecientes a dichos centroides. Estos centroides se utilizan para la asignación de las medias y los pesos de cada una de las gaussianas que forman cada modelo.

### 3.3.4 Support Vector Machines (SVM)

Este método es utilizado para la resolución de problemas de clasificación, regresión y estimación. Está basado en la teoría del aprendizaje descrita en [V. Vapnik, 1995a] [V. Vapnik, 1995b]. Según [T. Marwala et al, 2005] el objetivo principal es transformar los vectores de entrada (los vectores de características) en otros vectores de dimensión más alta en los que el problema pueda solucionarse linealmente.

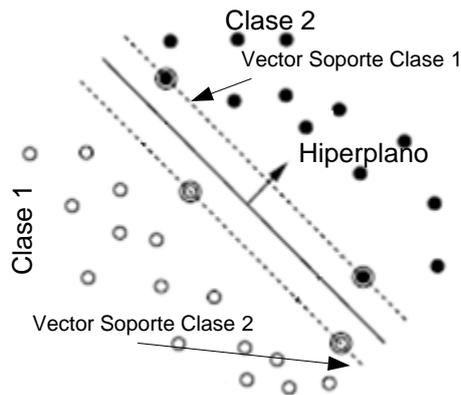


Figura 26: Hiperplano Separador de Clases

La frontera de decisión lineal corresponde a un hiperplano/s [Pai-Hsuen Chen et al, 2003]) separador de la clase/s óptimo/s ([Pai-Hsuen Chen et al, 2003]) [C. Burges, 1998]. El hiperplano elegido es el que maximiza la distancia euclídea a los vectores de características más cercanos de cada clase (Figura 26). Estos vectores se denominan vectores soporte. Cuando las clases no son linealmente separables, se busca el hiperplano que minimice los errores a través de una función de coste [C. Burges, 1998]. La extensión a límites no lineales se realiza a través de un núcleo (kernel) que satisfaga las condiciones de Mercer recogidas en [J. Mercer, 1909]. Existen distintos tipos de kernel (gaussiano, polinomial, perceptrón, sigmoide)<sup>89</sup> dependiendo de la aproximación que se quiera realizar, en el proyecto se utilizó concretamente uno del tipo gaussiano, que a su vez es el más utilizado según [Baudat G. et al, 2001]. Básicamente, cada vector de características es mapeado a un espacio de alta dimensionalidad en que los datos son separables linealmente. El producto escalar de los vectores transformados se puede escribir como el núcleo, por lo que se necesita trabajar en el espacio extendido.

SVM trabaja en bloque, por lo tanto, un cambio en los datos supone un nuevo entrenamiento. En la etapa de test, SVM obtiene una puntuación que indica cuantos vectores de características de la muestra desconocida pertenecen a una clase y cuantos pertenecen a la otra. También puede devolver directamente la clase a la que pertenece.

<sup>8</sup> [www.kernel-machines.org](http://www.kernel-machines.org)

<sup>9</sup> [www.support-vector.net](http://www.support-vector.net)

### 3.3.5 Neural Networks (NN)

Este clasificador se encuentra dentro de los clasificadores supervisados. La esencia principal del aprendizaje supervisado responde a la existencia de una serie de muestras de entrenamiento, véase [K. Fukushima, 1988]. Estas muestras atraviesan sucesivamente la red neuronal, donde el vector de coeficientes cepstrales se multiplica por los respectivos pesos y los resultados se suman y se evalúan mediante una función de activación del tipo sigmoideal, escalón u otra [A. V. Oppenheim et al, 1998]. La salida de la red neuronal simboliza la pertenencia a una clase u otra. Los valores tomados a la salida pueden variar dependiendo de si es binario (1, -1) o multiclase (1, 0, -1...).

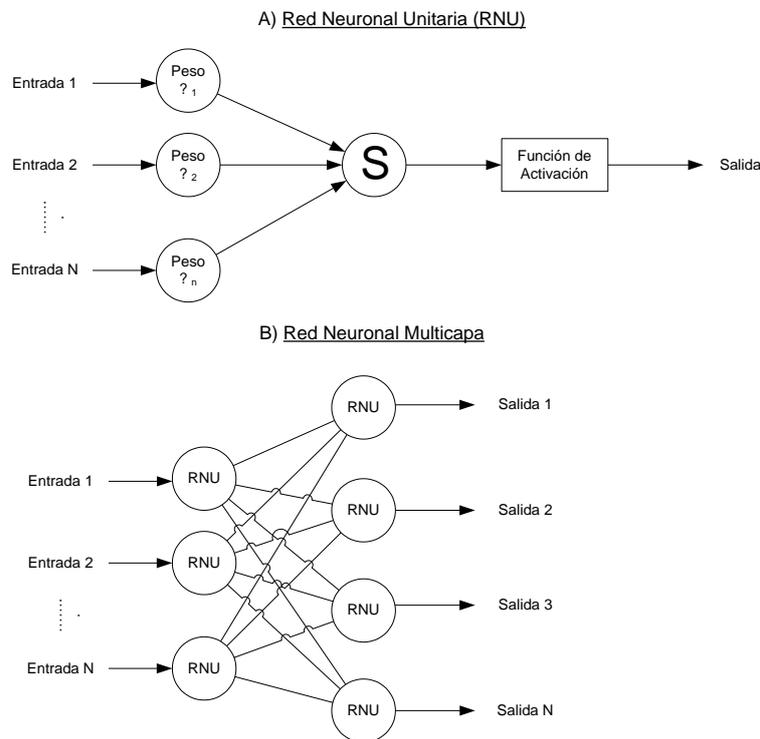


Figura 27: Diagrama de Red Neuronal Unitaria y Multicapa

Las muestras de entrenamiento, son pasadas por la red neuronal forzando a un valor de salida. En el caso de una red unitaria, si se obtiene en la salida un valor diferente a lo esperado se vuelven a calcular los pesos y se vuelve a entrenar (Figura 27a). Por el contrario, si la red es multicapa se vuelven a calcular los pesos de los nodos de la red neuronal teniendo en cuenta el error del nodo anterior (Figura 27b). Las redes unitarias están formadas por un único perceptrón (MLP), mientras que las redes multicapa se forman con múltiples perceptrones anidados (la salida de uno es la entrada de otro).

La red neuronal multicapa modela una línea de decisión de las clases entrenadas, lo que conlleva que para incorporar una nueva clase se deba entrenar nuevamente toda la red. En el test, el vector de características de la voz desconocida se somete al paso de la

red ya entrenada (con los pesos determinados) y a la salida se asociará con alguna de las clases.

### 3.4 DECISIÓN

Una vez obtenidas todas las puntuaciones para cada una de las muestras de entrenamiento del sistema, el último paso a proceder es asignar la correcta clasificación para la muestra introducida. Este proceso depende del tipo de comparación elegida y los algoritmos de comparación o clasificación elegidos. Por un lado, si se elige métodos basados en la comparación de modelos, la clasificación se basa en el cálculo y evaluación de distancias [R. Bajcsy et al, 1989] [Anil K. Jain et al. 2000]. Como consecuencia, se elegiría aquel modelo cuya distancia fuera menor con respecto a la muestra desconocida. Por otro lado, si se elige métodos estocásticos, la clasificación se basan en el cálculo y evaluación de probabilidades [L. Devroye et al, 1996] [Frederic Bimbot et al. 2004]. En este caso se elegiría aquel modelo cuya probabilidad con respecto a la muestra desconocida fuera mayor. La Figura 28 recoge de manera general el proceso de decisión.

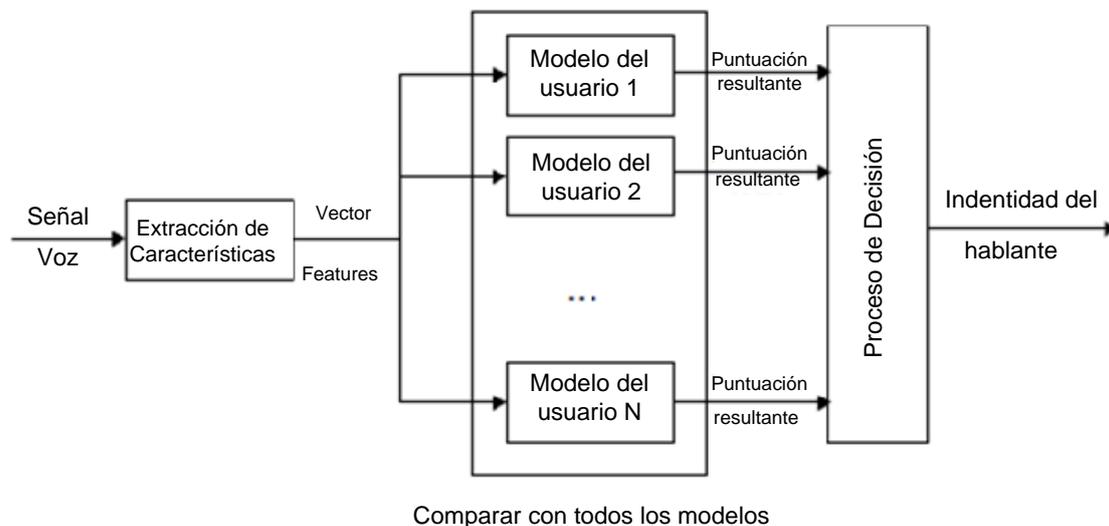


Figura 28: Diagrama General del Proceso de Decisión

---

# CAPÍTULO 4

---

## ESTADO DEL ARTE

Una vez introducido el concepto de biometría, sistemas biométricos y reconocimiento de locutor, este capítulo se centra en describir el estado del arte actual de la identificación biométrica y la manera de solventarlo en la actualidad.

### 4.1 INTRODUCCIÓN

La identificación de una persona, objeto o animal entre un conjunto de semejantes, se realiza constantemente en la vida cotidiana. Las personas, mediante sensores que posee su cuerpo extrae las características necesarias, y el cerebro humano se encarga de interpretar lo obtenido y suministrar una identidad al desconocido. Cuando se intenta imitar este proceso con una máquina se hace uso de un conjunto de herramientas que simulen el sistema humano. Entre ellas, sensores para extraer las características biométricas, algoritmos para la extracción de las mismas, una base de datos de individuos y finalmente, un proceso de toma de decisión. El conjunto de todo lo anterior forma un sistema biométrico.

Para realizar la identificación de un individuo, el sistema biométrico compara la muestra desconocida con los patrones o modelos biométricos existentes en el sistema [Anil K. Jain et al. 2004a] (véase sección 2.2.2). A partir de los resultados obtenidos en las sucesivas comparaciones se tomará una decisión (véase, para más información,

sección 3.4). En definitiva, el proceso de identificación pretende resolver la formulación de las preguntas: *¿A quién pertenecen estos datos biométricos? ¿Quién es esa persona?*

El proceso de identificación puede realizarse de manera activa o pasiva. En la primera, hay participación del individuo a identificar. La segunda se realiza sin colaboración del individuo (reconocimiento en casinos, aeropuertos...). Evidentemente, para que el sistema adjudique una identidad debe haber sido almacenado previamente su modelo. Surgiendo la necesidad de una base de datos de modelos de usuarios. Debe contener tanto los modelos de usuarios que pertenecen al sistema, como la identidad asociada a cada modelo.

A la identificación descrita anteriormente se le llama búsqueda exhaustiva (ES) [Feng Hao et al, 2008]. La toma de decisión en la búsqueda exhaustiva se puede realizar mediante un único umbral, un umbral por usuario o bien mediante la elección de la mayor verosimilitud. A continuación se describen más detalladamente cada uno de los métodos.

### **Búsqueda Exhaustiva (ES)**

En este método se establece como identidad aquella cuya verosimilitud resultante de comparar la muestra con el modelo sea mayor, siempre y cuando sobrepase un umbral previamente establecido [Feng Hao et al, 2008]. Para ello, se realiza un entrenamiento previo (véase, sección 2.2.2) con el fin de establecer un umbral de confianza óptimo. A continuación, se compara la muestra contra todos modelos de la base y se almacenan todas las puntuaciones obtenidas, junto con la identidad del usuario que obtuvo dicha puntuación. De todas las puntuaciones almacenadas se elige aquella con mayor valor para ser comparada con el umbral de confianza óptimo. Si la puntuación supera el umbral se sugiere como identidad, en caso contrario se establece como identidad desconocida.

Una desventaja de este método es la necesidad de unos elevados tiempos de cómputo para realizar una identificación [Nalini K. Ratha et al, 1996] [Xianchao Qin et al, 2007]. Puesto que para cada muestra introducida se debe recorrer toda la base de datos para conseguir una identificación. Repercutiendo así en un elevado coste computacional. La Figura 29 muestra el proceso descrito.

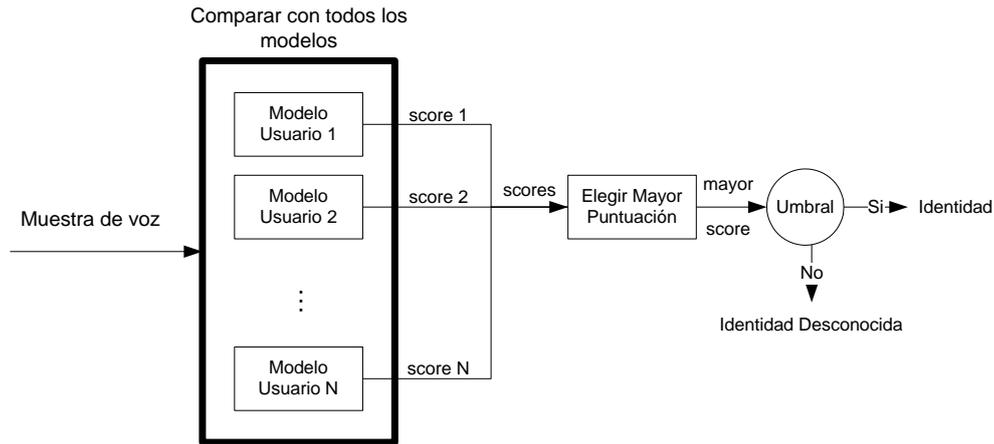


Figura 29: Diagrama de la Búsqueda Exhaustiva (ES)

### Búsqueda Exhaustiva con umbral general (ES-GT)

En este método se establece como identidad aquella que sobrepase un umbral preestablecido [Feng Hao et al, 2008]. Para establecer este umbral general se realiza un exhaustivo entrenamiento. A continuación, se compara la muestra desconocida con los modelos de la base de datos consiguiendo una puntuación por modelo. La primera puntuación, o score, obtenida que supere el umbral se sugiere como identidad. Si ninguna puntuación consigue superar el umbral se establece como identidad desconocida. La Figura 30 detalla el proceso descrito anteriormente.

En esta alternativa, sólo se realizará un elevado número de comparaciones cuando el modelo de la identidad sugerida sea almacenado en la última posición. Como contrapartida si existen dos modelos cuyas puntuaciones de *matching* superen el umbral, el sistema establece como identidad aquella de las dos que se encuentre en primer lugar, independientemente de si es la verdadera o no.

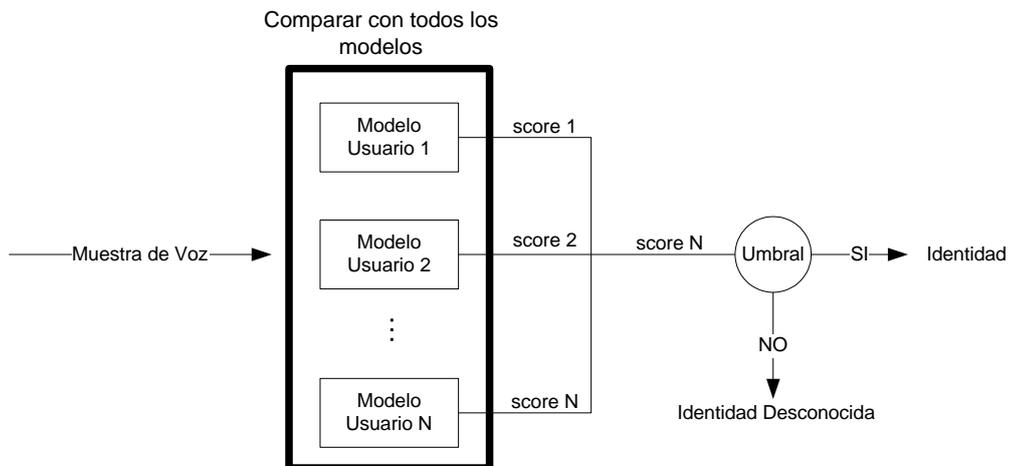


Figura 30: Diagrama de la Búsqueda Exhaustiva con umbral general (ES-GT)

### Búsqueda Exhaustiva con umbral por usuario (ES-TPU)

Este último método se basa en el establecimiento de un umbral para cada modelo de usuario almacenado en la base de datos. Se sugiere como identidad aquella cuya verosimilitud resultante de la comparación supere el umbral establecido para dicho modelo (Figura 31). El cálculo de estos umbrales se realiza previamente mediante un exhaustivo entrenamiento del sistema. En esta alternativa, la muestra desconocida se introduce en el sistema y se compara con el modelo del primer usuario almacenado en la DB. Si la puntuación obtenida supera el umbral pre-establecido para dicho usuario se sugiere como identidad. En caso contrario, se repetiría el mismo proceso con el siguiente modelo de usuario almacenado. Si se llegara al último usuario almacenado y no se pudo sugerir ninguna identidad, se sugiere como identidad desconocida. La Figura 24 muestra el proceso descrito.

Al igual que la alternativa anterior, puede ocurrir que una muestra consiga pasar los umbrales de dos modelos de usuarios diferentes. Al ser un proceso secuencial, se sugeriría como identidad la encontrada en primer lugar. Por otro lado, al igual que la alternativa anterior, se evita comparar con toda la base de datos reduciendo así el tiempo de cómputo y la capacidad de cálculo.

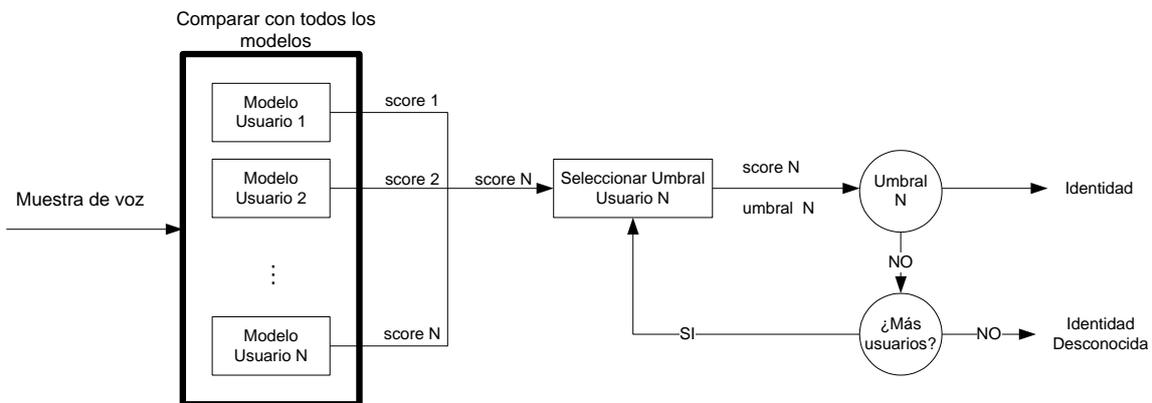


Figura 31: Diagrama de la Búsqueda Exhaustiva con Umbral por Usuario (ES-TPU)

La base de datos esta en continuo crecimiento afectando a la probabilidad de falsas comparaciones, la cual incrementa a medida que aumenta el número de usuarios alistados en la DB [R. Mukherjee, 2007]. Recapitulando, al sistema biométrico se le introduce una muestra biométrica de un individuo desconocido y se compara con todos los modelos almacenados en la base de datos. Esta comparación consume mucho tiempo de cómputo, especialmente cuando la base de datos es grande. Por tanto, cuanto mayor sea el número de comparaciones mayor será el tiempo necesario para realizar la identificación.

## 4.2 PROBLEMA DE LA IDENTIFICACIÓN

La identificación biométrica que se describió en la sección anterior posee un problema: “los elevados tiempos de cómputo”. El problema referido como elevados tiempos de cómputo es originado por el elevado número de comparaciones que se deben realizar para llegar a una identificación. Por tanto, cuanto mayor sea el número de comparaciones mayor será el tiempo necesario para realizar la identificación.

Para una base de datos de  $N$  usuarios registrados el tiempo necesario para realizar una comparación es  $m$  segundos. Por lo tanto, el tiempo mínimo requerido para una identificación será de  $N \cdot m$  segundos. En el entorno de uso de la identificación biométrica se requiere obtener una identidad en el menor tiempo posible. Si tenemos en cuenta este ámbito de uso (casinos, aeropuertos, forenses...) estos tiempos de cómputo no son óptimos.

## 4.3 SOLUCIÓN ACTUAL DEL PROBLEMA

Esta sección describe las diferentes maneras de abordar el problema en la actualidad, así como parte de las soluciones utilizadas. Para cada modalidad biométrica existen soluciones específicas que hacen uso del conocimiento previo del proceso de extracción de características. Por tanto, las soluciones que se describen a continuación están basadas en la semántica asociada a los procesos de tratamiento de las respectivas muestras biométricas.

### 4.3.1 Soluciones Adoptadas en Huella Dactilar

En esta modalidad se realiza una preclasificación de los modelos con el fin de reducir el número de operaciones a realizar [F. M. Rodríguez et al, 2004]. El número y las clases utilizadas en la preclasificación no son siempre las mismas, pero siempre se basan en los puntos singulares de la huellas (núcleo y deltas). Los algoritmos de clasificación juzgan un papel importante a la hora de clasificar las huellas en las diferentes. El funcionamiento de todos ellos es parecido: primero se extrae la muestra, a continuación los puntos singulares (minucias) y por último se clasifica dentro de las clases predefinidas [Nalini K. Ratha et al, 1996]. Existen distintos tipos de clasificación de las huellas:

- *Clasificación en 5 clases:* Fue la primera clasificación de huellas dactilares. Esta clasificación fue realizada por Henry [H. C. Lee et al, 1991]. Clasificó las huellas en 5 tipos o clases (Figura 32): arco (arch), arco de carpa (tented arch),

lazo izquierdo (left loop), lazo derecho (right loop) y anillo de cresta (whorl) [R. C. Veltkamp et al, 2000] [F. M. Rodríguez et al, 2004].

- *Clasificación en 4 clases<sup>10</sup>*: Se basa en el estudio realizado por Federico Olóriz. Cabe resaltar que es la utilizada en España. Se basa en una reclasificación de las clases de Henry. Federico concluyó que las huellas dactilares se podrían clasificar en 4 tipos (en vez de 5 como propuso Henry): adelta (no tiene delta, se trata de un arco o arco de carpa), dextrodelta (delta a la derecha y lazo izquierda), sinistrodelta (delta a la izquierda y lazo derecha) y bidelta (dos deltas a ambos lados, en estos casos el núcleo hace un remolino: anillo de cresta).
- *Clasificación en 3 clases*: Se basa en una reestructuración del estudio de Henry. Al igual que sus predecesoras se concluyó que las huellas dactilares se pueden clasificar en 3 tipos: espiral (formada por espiral y doble lazo), lazo (formado por lazo izquierdo y lazo derecho) y arco (formado por arco y arco de carpa) [Anil K. Jain et al, 2000] [K. C. Chan et al, 2003].



Figura 32: Tipos de huella (dactiloscopia de Henry)

De acuerdo con lo anterior, si se puede clasificar una huella en una de las clases disjuntas mencionadas anteriormente sólo habrá que compararla con las huellas pertenecientes a esa clase. Por lo tanto, se realizarían menos comparaciones repercutiendo en un aumento de la velocidad de identificación. Reduciéndose a un 20%, 25% y 33% de comparaciones, respectivamente.

Como se puede observar es una solución no genérica sólo aplicable a la identificación biométrica basada en la huella dactilar. Esto se debe a la existencia de diferencias considerables entre los distintos tipos de huellas que permiten una preclasificación previa.

<sup>10</sup> <http://www.biografiasyvidas.com/biografia/o/oloriz.htm>  
[http://servicios.ideal.es/granadinos/federico\\_oloriz.html](http://servicios.ideal.es/granadinos/federico_oloriz.html)

### 4.3.2 Soluciones Adoptadas en Iris

En este apartado se describen las diferentes soluciones adoptadas en el proceso de identificación de iris. La comparación entre dos muestras de iris es realmente sencilla, se cuenta el número de bits que difieren entre dos vectores binarios (distancia de Hamming [Feng Hao et al, 2008]).

#### Basadas en “Partitioning Principle”

La mayoría de las técnicas NNS (Nearest Neighbour Search) están basadas en “partitioning principle” [P. Zezula et al, 2006]. La idea de estas técnicas es dividir el espacio de búsqueda en regiones, así al identificar una muestra sólo se busca en algunas regiones. Generalmente, el primer paso de estas técnicas es la elección de pivotes (puntos de referencia que dividen el espacio). En las técnicas “ball partitioning”, la pelota (ball) central es un pivote. Las pelotas dividen el espacio en mitades (dentro de la pelota y fuera de ella). Esta división esférica se ejecuta recursivamente y como consecuencia se obtiene un árbol binario formado por pivotes usados como nodos y los datos como hojas. Un método alternativo es el llamado “generalized hyperplane partitioning”, el cual separa el conjunto de datos en dos, basándose en la distancia relativa de los dos puntos de referencia (pivotes).

Las técnicas “partitioning principle” generan varias estructuras de árbol (*m tree*, *vp tree*, *fq tree*... [Feng Hao et al, 2008]). Estas técnicas difieren de las anteriores en cómo se genera el árbol y la manera de recorrerle. Seleccionar los correctos pivotes es un paso esencial y no trivial. Debido a la complejidad de este paso, muchas técnicas los eligen de manera aleatoria. Algunas técnicas emplean pre-computación para reducir la carga computacional, pero requiere una gran cantidad de memoria [B. Bustos et al, 2003].

Desafortunadamente las estructuras de árbol sucumben a la “maldición de la dimensionalidad”. Muchas técnicas funcionan bien en 2-D y 3-D pero al aumentar la dimensión de los datos, el tiempo y la base de datos perciben un crecimiento exponencial y no obtienen mejores resultados que la fuerza bruta (búsqueda exhaustiva) como bien se describe en [M. Datar et al, 2004].

#### Basadas en Locality Sensitive Hashing (LSH)

La parte central de esta técnica es la definición de una función hash familiar (llamada  $G$ ), de la cual se eligen un número  $l$  de funciones (llamadas buckets) de manera uniforme y aleatoria, como se describe en [P. Indyk et al, 1998]. Durante el pre-procesado, cada punto  $p \in P$  se almacena en cada uno de los  $l$  buckets anteriores, identificado por  $g_i(p)$ . Teniendo así un bucket formado por un conjunto de puntos. A la hora de procesar una muestra  $q$ , LSH busca exhaustivamente en todos los buckets ( $g_1(q)$ ,  $g_2(q)$ ...) la puntuación más cercana. Como premisa se asume con alta probabilidad que la puntuación más cercana existe en uno de los buckets. Para evitar la

repetición de puntos en la comparación, se guarda una marca en memoria de los puntos ya comparados, ignorando así dichos puntos en siguientes iteraciones.

La mayor limitación que posee el LSH es que a menudo la búsqueda necesita un porcentaje de la base de datos. La posible solución adoptada es la creación de un umbral e interrumpir la búsqueda cuando haya suficientes puntos en el bucket [P. Indyk et al, 1998].

Existe una técnica similar que divide una muestra de 1024 byte en 256 submuestras de 32 bits [J. Haitzma et al., 2002]. Se presupone que para una muestra levemente degradada, al menos existe una submuestra libre de errores. Así, esta submuestra libre de errores puede servir como un punto de mira para posibles búsquedas. El problema es que se utilizaría una tabla de 32 bits que requeriría una memoria mínima de 4GB.

### **Basadas en “Single Collision Principle”**

Este principio asume que si existen dos muestras similares, poden tener relativamente una elevada posibilidad de tener al menos un segmento colindante idéntico (o compartir una porción de la muestra), obtenido para cada función hash o muestra directa.

A menudo los algoritmos LSH y Single Collision Principle ignoran las posibles múltiples ocurrencias de la colisión. Esas posibles ocurrencias dieron lugar a los algoritmos basados en “Multiple Colliding Segments Principle”.

### **Basadas en “Multiple Colliding Segments Principle”**

El algoritmo más conocido que usa esta técnica se llama Beacon-Guided Search (BGS). Una beacon es una colección de ID de código de iris que poseen la misma definición de características. Esta difiere de un bucket en que es más que una unidad de almacenamiento, son marcas de referencia. La forma de trabajo de estos algoritmos se divide en dos etapas: preprocesado y búsqueda. Durante el preprocesado todas las muestras de iris existentes en la base de datos son índices para crear una estructura beacon de referencia. Esta estructura está formada por 128 espacios de beacon con  $2^m$  beacons en cada espacio. Los  $m$  bits del índice de la beacon  $b_i$  identifica unívocamente a cada beacon del  $i$ -ésimo espacio. De esta manera el índice de la beacon y el ID de la muestra de iris solo ocupan 32-bits en la beacon por espacio de beacon [Feng Hao et al, 2008]. En la etapa de búsqueda BGS selecciona para cada espacio una beacon y recupera el identificador (ID) almacenado en cada beacon. A continuación se almacena en un array las ocurrencias producidas por dicha beacon (las veces que se la pidió recuperar el identificador). Si una determinada ID ha sido encontrada  $c$  veces ( $c$  es un valor que se debe establecer), el algoritmo BGS carga el archivo completo de la muestra de iris utilizando la ID almacenada en la beacon. Como último se realiza la comparación de la misma manera que en ES (Exhaustive Searching).

### **Basadas en clasificación de categorías**

El principal propósito de estas técnicas es encontrar características globales de las imágenes de iris con el fin de poder clasificarlas en varias categorías, y reducir el número de muestras necesarias para realizar una búsqueda.

El algoritmo visual-dictionary-learning, en primer lugar compacta y discrimina las características visuales, llamadas Iris-Texton, automáticamente de un conjunto de imágenes de entrenamiento. Posteriormente el histograma del Iris-Texton se usa para representar la característica global de la imagen del iris. Finalmente, el clasificador k-means clasifica una imagen de iris de entrada en una de las cinco categorías.

Otras técnicas descritas en [K. Karu et al, 1996] utilizan la información tal como números, tipos y la ubicación de puntos singulares para clasificar las huellas dactilares. Por otro lado en [X. Wu et al, 2004] usan detectores lineales direccionales para extraer las principales líneas, para luego clasificarlas en seis categorías de acuerdo con el número de las principales líneas y el número de sus intersecciones. Mientras en [L. Yu et al, 2005] se utiliza un método llamado “box-counting” que estima las dimensiones fractales del iris para la posterior clasificación de la imagen en una de las cuatro categorías de acuerdo con la dimensión fractal. Para finalizar en [J. Fu et al, 2005] se utiliza un filtro de color artificial para detectar información de color en las imágenes de iris, esta información se utiliza para mejorar la precisión del reconocimiento y disminuir el tiempo de búsqueda.

Sin embargo, no ha sido propuesto ningún algoritmo que aborde el problema careciendo de información sobre la semántica de los datos almacenados en los vectores de características que parametrizan la muestra biométrica



# CAPÍTULO 5

## METODOLOGÍA DE TRABAJO

En este capítulo se describe la metodología de trabajo adoptada para resolver el problema elevados tiempos de cómputo. En capítulos posteriores se detallarán cada uno de los pasos, así como todo el material necesario y las evaluaciones realizadas.

La consecución del proyecto es una ardua tarea que puede amenizarse con el establecimiento de unas directrices a seguir. El conjunto de directrices o pasos que se utilizaron en este proyecto formaron la metodología de trabajo. La (Figura 33) describe de manera esquematizada la metodología de trabajo que se ha seguido.

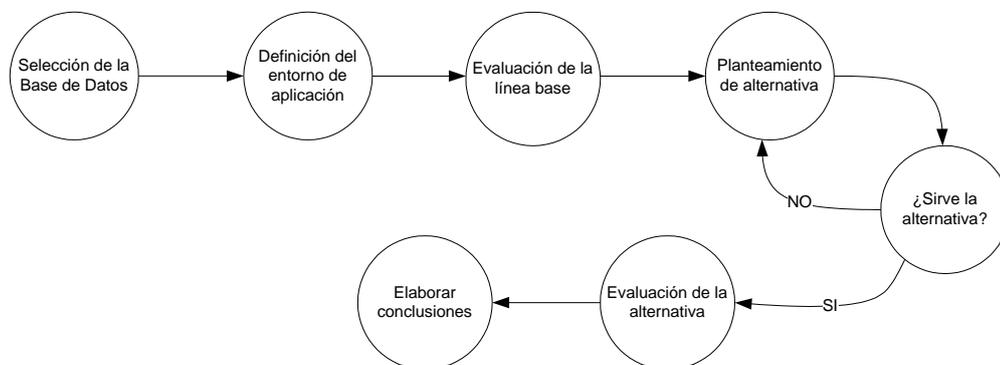


Figura 33: Diagrama de Procesos de la Metodología de Trabajo

## 5.1 METODOLOGÍA DE TRABAJO

1. **Primero:** *Selección de la base de datos:* Para poder comenzar a trabajar se necesita elegir una base de datos con la que realizar pruebas.
2. **Segundo:** *Definición del entorno de trabajo:* Se elige el lenguaje de programación, sistema operativo, maquina con que trabajar.
3. **Tercero:** *Evaluación de la línea base:* Será la referencia necesaria para la evaluación de los resultados y valorar así la culminación de los objetivos. Se medirán los tiempos de cómputo de la identificación con la ayuda de un software.
4. **Cuarto paso:** *Planteamiento de una alternativa a la línea base:* Se elabora una alternativa que mejore solución del paso anterior.
5. **Quinto paso:** *Evaluar la alternativa:* Se evalúa la solución del paso anterior. En caso de no alcanzar los requisitos pre-establecidos se vuelve al paso cuarto.
6. **Sexto paso:** *Elaboración de conclusiones:* Se estudiarán las ventajas y desventajas de las alternativas creadas.

---

# CAPÍTULO 6

---

## BASE DE DATOS

A la hora de evaluar un experimento biométrico una parte importante del mismo es la elección de la base de datos. Esta debe de ser lo suficientemente extensa y variada como para obtener una buena representación de los individuos. Debe poseer varias muestras de alta y baja calidad de cada individuo con el fin de evaluar la robustez de la solución. La base de datos fue suministrada en Access. Access ofrece una particularidad: con una sola query se puede acceder a todos los valores de las columnas de la base de datos. De esta forma se ahorra tiempo de cómputo evitando abrir múltiples queries.

A lo largo de este capítulo se describen las diferentes bases de datos que disponía el departamento y las características de las mismas. También se describirá la elección de una de ellas en detrimento de las demás.

### 6.1 BANCA

BANCA es una base de datos nacida de un proyecto Europeo con el fin de establecer e implementar sistemas de seguridad con identificación, verificación y acceso de control para aplicaciones de Internet [E. Bailly-Baillire et al, 2003].

Para el proyecto BANCA se necesitaba una base de datos multimodal que incluiría varios pero realistas escenarios donde grabar (de prueba), usando diferentes tipos de materiales y en diferentes lenguas Europeas.

A pesar de cumplir los requisitos necesarios para una base de datos se optó por descartarla por su inferior número de sujetos (52) y de muestras (208).

### **Sistema de Adquisición**

Para la grabación de la base de datos se utilizaron dos tipos de cámaras diferentes, una webcam analógica y una cámara digital de alta calidad. Para la duración del grabado, las cámaras se pusieron en modo automático. En paralelo, se usaron dos micrófonos, uno de alta calidad y otro de calidad pobre. La base de datos se grabó en formato PAL DV (sistema propietario cuyas capturas de video poseen una resolución de color de 4:2:0). El audio de ambos micrófonos fue capturado con 12bit y 16bit y una frecuencia de 32kHz. El video fue comprimido con una tasa de 5:1.

### **Especificaciones**

La base de datos BANCA fue diseñada en orden de un test multimodal IV con varios sensores de adquisición (2 cámaras y 2 micrófonos) y bajo múltiples escenarios (controlado, degradado y adverso). Para 4 diferentes lenguajes (inglés, francés, italiano y español), el video y la voz fueron tomados de 52 sujetos diferentes (26 hombres y 26 mujeres), por ejemplo un total de 208 sujetos.

Cada sujeto realizó 12 sesiones de grabado, cada una de ellas contenía dos grabaciones: 1 verdadera (el cliente accede) y 1 informal (ataque de un impostor). Las 12 sesiones fueron separadas en 3 diferentes escenarios:

- *controlado* para las sesiones 1-4
- *degradado* para las sesiones 5-8
- *adverso* para las sesiones 9-12

La webcam se utilizó en el escenario degradado, mientras que la cámara cara se usó en los restantes. Los micrófonos fueron utilizados simultáneamente en cada uno de los tres escenarios con la salida siendo grabada en diferentes líneas del DV tape.

Durante cada grabación, el sujeto tuvo que decir un número aleatorio de 12 dígitos, su nombre, su dirección y su fecha de cumpleaños. Cada grabación tuvo una duración de 20 seg.

## 6.2 BIOSEC BASELINE CORPUS

Esta base de datos fue creada con el fin de superar la ausencia de rasgos, sensores e informes de simulaciones falsificadas en las bases de datos multimodales existentes. BioSec fue un proyecto Europeo para crear una elevada base de datos multimodal que incluyera huellas dactilares, iris, voz y rostro, de alta y baja calidad. A esta base de datos se terminó denominando BioSec baseline corpus [J. Fierrez-Aguilar et al, 2006].

### Sistema de Adquisición

El escenario de adquisición fue una oficina, con un ancho escritorio para el hardware de adquisición y dos sillas para el donante y el supervisor de la adquisición. Las condiciones medioambientales no fueron controladas para simular una situación real. El hardware de adquisición incluye un ordenador personal y un número de sensores biométricos comerciales.

Los sensores utilizados son: Authentec AES4000 (Huella dactilar), Atmel FCDEMO4 (Huella dactilar), Biometrika FX2000 (Huella dactilar), Philips ToUcam PRO II (Rostro y voz), Plantronics DSP-400 (voz), LG IrisAccess EOU (Iris).

### Especificaciones

Biosec baseline corpus contiene rasgos biométricos multimodales de 200 sujetos diferentes. Cada sujeto participo en dos sesiones de adquisición separadas entre ellas de una a cuatro semanas. Para cada sujeto se obtuvieron los siguientes rasgos:

**Rostro.** Se tomaron 4 imágenes de la cara en una pose neutral a una distancia de 30 cm de la cámara (2 al comienzo y 2 al final de la sesión). En cada imagen se fueron cambiando la expresión del rostro para evitar samples idénticos del rostro. Haciendo un total de 1600 imágenes.

**Voz.** Se tomaron 4 muestras para cada número específico de usuario de ocho dígitos (2 al comienzo y 2 al final) y otras 3 más de los otros números de usuarios con el fin de simular ataques impostores. Los 8 dígitos fueron pronunciados uno a uno en una continua y fluida secuencia. Estos fueron grabados en español e inglés. El número total de expresiones de voz es  $2 \times 200 \times (4+3) \times 2$  sensores  $\times$  2 idiomas.

**Iris.** 4 imágenes del iris de cada ojo, cambiando de ojo entre exposiciones para un total de 3200 imágenes.

**Huella dactilar.** 4 capturas con cada uno de los 3 sensores de las huellas de los 4 dedos (derecha e izquierda del índice y del anular), intercambiando dedos entre consecutivas exposiciones. Hasta un total de imágenes de  $2 \times 200 \times 4 \times 3$  sensores  $\times$  4 dedos.

Aunque esta base de datos es bastante más completa que la anterior se optó por no trabajar con ella por el mismo motivo que la anterior (pocas muestras por usuario).

### 6.3 MICROAES

Se trata de una extensa base de datos de habla española. Es una de las bases de datos más completas existentes en idioma español con un amplio número de muestras por sujeto de alta y baja calidad [A. R. Rincón et al, 2003].

#### **Sistema de Adquisición**

El equipo de grabado está basado en un portátil que usa su PCMCIA como un interfaz de equipamiento de audio. Toda la grabación fue bajo el sistema operativo Windows 2000. Un software de grabado llamado Nanny Record creado por UPC. También se usaron cuatro micrófonos que grababan simultáneamente. Amplificadores, periféricos mensurables y cualquier otra especificación se puede consultar en [1]. Los filtros paso altos no se usaron.

Todas las grabaciones tuvieron lugar en un entorno de oficina donde la gente trabajaba en sus escritorios, probablemente con un ordenador. No hubo discusiones o disputas durante la grabación. El hardware utilizado fueron: micrófonos Sennheiser ME 104 (distancia corta), Nokia Lavalier HDC-6D (distancia corta), Sennheiser ME64 (distancia media) y Mikrofonbau Haun MBNM-550 E-L (larga distancia).

Los sujetos fueron reclutados desde diferentes universidades españolas, clubs de deporte y ATLAS. El equipo de audio se desplazó a cada área dialéctica para realizar las grabaciones. La edad de los usuarios reclutados está comprendida en la Tabla 2: Edades de los donantes. Las muestras consistían en la lectura de textos y a la respuesta de una batería de preguntas.

Age	Masculine	Feminine	Total users	%
-15	6	8	14	4,67%
16-30	63	62	125	41,67%
31-45	31	36	67	22,33%
46-60	34	34	68	22,67%
61-	13	13	26	8,67%
<b>TOTAL</b>	147	153	300	100,00%

Tabla 2: Edades de los donantes

### Especificaciones

La muestra fue almacenada en un fichero con formato raw, sin cabeceras en el archivo señal. Cada uno de los cuatro canales se grabó a 16 kHz con 16 bit de cuantificación con el último byte significativo primero como enteros. Una descripción de la tasa de la muestra, la cuantificación y el orden de bytes usados está especificada en cada etiqueta SAM del archivo al que corresponde cada muestra. Esta etiqueta contiene información sobre la calidad de la señal del hablante.

La base de datos está formada por 300 sujetos. El corpus está formado por 30 sets de 15 párrafos dando un total de 450 párrafos. Para grabar 300 sujetos se generaron 300 secuencias. Cada sentencia fue repetida 10 veces. Dando un total de 20388 muestras, unas 64 por usuario. Para más información sobre esta base de datos léase el documento [1].

### SELECCIÓN DE LA BASE DE DATOS

Se decidió trabajar con la base de datos MICROAES debido al elevado número de muestras de alta y baja calidad que ofrece por usuarios. Pudiendo conseguir así una mejor resolución del sistema y unos mejores resultados en la evaluación del proyecto.



---

# CAPÍTULO 7

---

## ENTORNO DE TRABAJO

Para la realización del proyecto fue necesaria la elección de un lenguaje de programación. En el cual plasmar los algoritmos elaborados, y poder evaluar las soluciones y la línea base de manera automática.

Se buscaba un lenguaje de programación sencillo, que no complicara en demasía la implementación en código de la solución. Las opciones que se barajaron fueron Java y C++.Net. Ambos son lenguajes orientados a objetos, lo cual ofrece una facilidad para la creación de la algoritmia necesaria y la orientación a objetos ofrecía una sencillez añadida. El elegido finalmente fue el C++.Net. La motivación que originó su elección fue la reutilización de librerías ya implementadas, lo cual simplifica la creación de código de manera radical y evitando así la re-codificación de algoritmos ya implementados.

La aplicación del lenguaje fue a través de Visual Studio 2005. Microsoft Visual Studio<sup>11</sup> es un entorno de desarrollo integrado para sistemas basados en Windows. Soporta lenguajes de programación como Visual C++, Visual C#, Visual J#, ASP.NET y Visual Basic .NET y actualmente se están desarrollando extensiones para muchos otros lenguajes.

Visual Studio permite crear aplicaciones, sitios y aplicaciones web, así como servicios web en cualquier entorno que soporte la plataforma. Visual Studio se ha

---

<sup>11</sup> <http://www.microsoft.com/spanish/msdn/vs2005/default.aspx>

utilizado para codificar las diferentes aplicaciones y algoritmos necesarios para la evaluación y creación del motor de Base de Datos.

Las pruebas se realizaron con un ordenador de características técnicas: Dual-Core AMD (Opteron) 1,80 GHz, 4GB de RAM. Este ordenador se utiliza como servidor del departamento de biometría de la universidad.

---

# CAPÍTULO 8

---

## LINEA BASE

En este capítulo se describe la línea base que sirvió como punto de partida para el proyecto. Esta línea base ha sido utilizada para comparar los resultados obtenidos con los sucesivos experimentos, así como servirá de referencia.

Se tomo como línea base la identificación definida en el capítulo 4. Como se pudo observar en el capítulo mencionado, existen tres tipos diferentes de identificación bruta o identificación exhaustiva (ES, ES-GT, ES-TPU). Se implementaron cada uno de ellos en el entorno de aplicación seleccionado y se evaluaron los resultados obtenidos.

Para el establecimiento de los umbrales óptimos de cada una de las técnicas identificativas se fue buscando el punto EER. Para ello se estableció un umbral inicial comprendido entre los valores 0 y 1 (en este caso se ajusto a 0.5) y se fueron ajustando dichos extremos dependiendo de los valores de las tasas FMR y FNMR. Si la diferencia de extremos era menor que un valor pre-establecido (criterio de parada) se establecía como umbral óptimo la media aritmética de los extremos. El proceso es el mismo tanto para generar un umbral general, como para la generación de un umbral por usuario.

	COMPARATIVA		
	ES	ES-GT	ES-TPU
Tasa de Falso No Reconocimiento (FNMR)	34,52%	71,61%	54,77%
Tasa de Falso Reconocimiento (FMR)	17,90%	60,13%	22,40%
Halt Total Error Rate (HTER)	26,21%	65,87%	38,58%
Tiempo Utilizado (UT) (min) (25 muestras)	2204,68	685,55	2204,56
Tiempo por Muestra (MUT) (min)	1,32	0,41	1,31

Tabla 3: Comparativa de las Diferentes Variantes de Identificación Exhaustiva

Como se puede observar en la Tabla 3 la búsqueda exhaustiva sin modificaciones (ES), también llamada identificación bruta, es la que mejores tasas de FNMR, FMR y HTER ha obtenido. Como contrapartida la identificación ES-GT es la que peores resultados consiguió. Y por último la técnica ES-TPU fue la que obtuvo resultados intermedios.

Conociendo el funcionamiento de los procesos de identificación ES, ES-GT, ES-TPU los resultados obtenidos se podría decir que son esperados, no tanto en los ratios FNMR, FMR y HTER conseguidos pero si en el orden de ratios esperados para cada técnica. La identificación bruta ES fue la que mejores resultados obtuvo pero a su vez requería más tiempo en realizar una comparación. La identificación con umbral general (ES-GT) obtuvo unos paupérrimos resultados tal y como cabía esperar, en contraposición consiguió unos tiempos de cómputo menores. Aún así no es muy aconsejable esta técnica. Por último, la identificación biométrica basada en umbral por usuario (ES-TPU) fue la que obtuvo un mejor resultado después de ES y al igual que ella empleo un mayor tiempo en realizar una comparación.

El tiempo en realizar una identificación depende del número de comparaciones. A continuación se ofrece una estimación del número de comparaciones necesarias para cada una de las técnicas utilizadas en identificación bruta:

$$C_{ES} = N$$

$$C_{ES-GT} = \frac{7}{2}N$$

$$C_{ES-TPU} = N$$

Donde  $N$  es el número de modelos existentes en la base de datos y  $C$  es el número de comparaciones.

---

# CAPÍTULO 9

---

## PLANTEAMIENTO Y EVOLUCIÓN DEL GMT

Una vez introducido el concepto biometría y sistemas biométricos, describiendo entre tanto el proceso de reconocimiento e identificación biométrica para concluir con el problema que este último conlleva. En este capítulo se detalla la mejora adoptada para paliar el problema referido como “elevados tiempos de cómputo”. También recoge las soluciones intermedias creadas hasta la obtención de la definitiva y las sucesivas mejoras de la misma.

Como se ha visto a lo largo de la memoria, en el proceso de identificación biométrica se necesita realizar tantas comparaciones como usuarios existentes en el sistema. Estas comparaciones hacen que el proceso se ralentice a medida que aumenta el número de usuarios registrados. Todo esto repercute en la velocidad de identificación, aumentándola considerablemente y directamente en el coste computacional del sistema de la misma forma.

La solución inmediata es reducir la base de datos de tal manera que no se limite el número de usuarios que puedan acceder al sistema, así como tampoco sea necesario eliminar usuarios del mismo.

Este PFC propone una agrupación de los modelos de usuarios en una estructura de árbol binario, bautizada como GMT (Group Model Tree).

## 9.1 PLANTEAMIENTO

Dado un sistema biométrico con una base de datos de tamaño  $N$ , para la identificación de un individuo se necesitarían un número  $N$  de comparaciones. Si se realiza una preclasificación de los modelos del sistema en subgrupos de tamaño  $m$  (donde  $m < N$ ) el número de comparaciones queda reducido a  $m+1$  comparaciones. Cada subgrupo queda representado por un modelo de grupo que identifica al conjunto de usuarios en su haber. Por lo tanto, la nueva base de datos estará formada por los modelos de los subgrupos en vez de los modelos de los usuarios. La Figura 34 describe de manera gráfica la metodología propuesta.

Con la creación del GMT se reduciría el número de comparaciones sin disminuir el tamaño de la base de datos. Aliviando a su vez el coste computacional necesario para realizar una identificación. Todo esto se consigue mediante la agrupación de los patrones biométricos (o modelos de usuarios) almacenados en la base de datos, en subgrupos de individuos que compartan algún tipo de característica. Con esta premisa, se pretende reducir el número de comparaciones de  $N$  a  $\log(N)$ .

El funcionamiento de un sistema biométrico que utilice dicha técnica no difiere mucho del tradicional. Cuando al sistema se le introduce una muestra biométrica se compara con los modelos de subgrupos devolviendo una puntuación (score). A la salida del sistema se devuelve una lista de los posibles candidatos que ofrecen mayor semejanza. El sistema evalúa el resultado e interpreta que el individuo que buscamos es muy posible que se encuentre en el subgrupo del modelo que ofrece una mayor coincidencia. La comparación de la muestra biométrica con los modelos almacenados se realiza sólo sobre aquellos que pertenecen al subgrupo candidato ( $m+1$  comparaciones,  $m < N$ ).

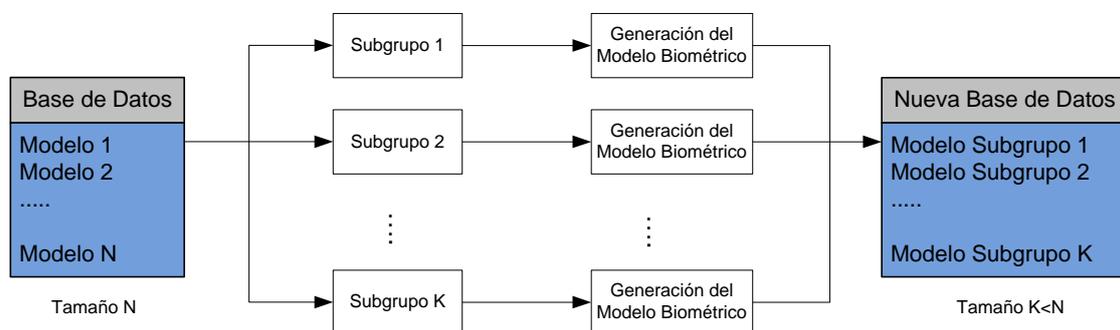


Figura 34: Metodología de Agrupación Propuesta

Aplicando esta técnica podría ocurrir que el tamaño de los subgrupos formados sea aún lo suficientemente grande repercutiendo en una velocidad de identificación elevada. Para paliar esta posible complicación, se puede aplicar esta técnica continuamente a

cada uno de los subgrupos generados hasta llegar a un criterio de parada preestablecido o a una solución final. La Figura 35 muestra el funcionamiento de la metodología de preclasificación con criterio de parada.

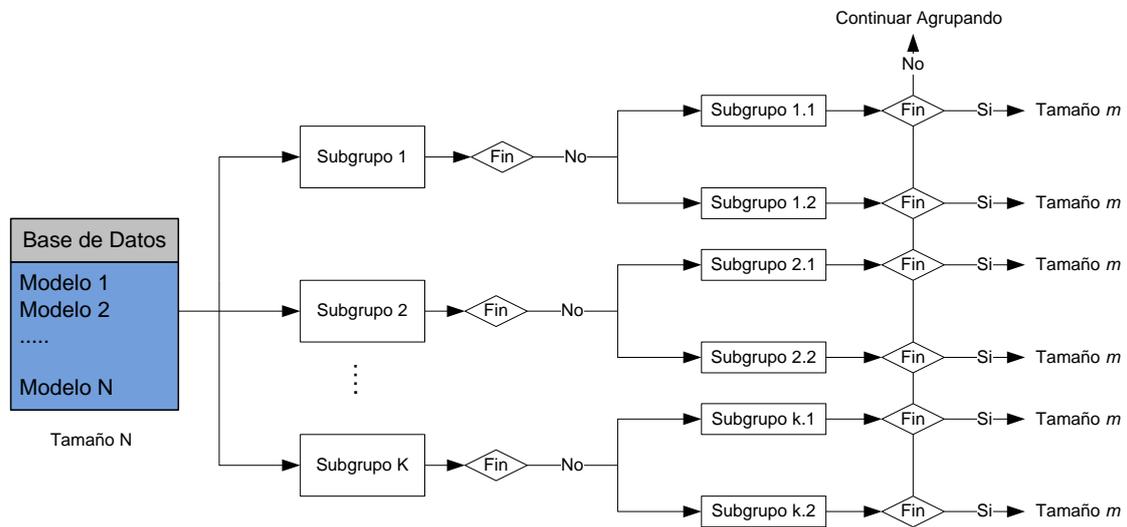
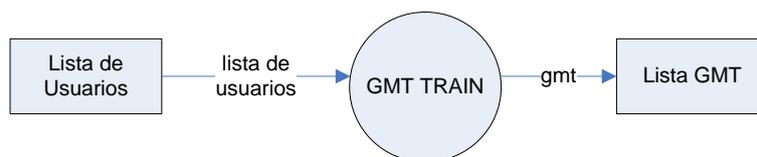


Figura 35: Metodología de Agrupación con Criterio de Parada

En definitiva, la solución propuesta consiste en transformar la lista de usuarios del sistema en una lista de grupos de usuarios, bautizada como GMT (Group Model Tree). Esta nueva lista es menor en tamaño que la original pero manteniendo el mismo número de usuarios, por tanto la misma capacidad identificativa. Otra diferencia que posee, es que la nueva base de datos no estará formada por modelos de usuarios sino por modelos de grupos.

Se pensó que la creación del GMT se hiciera en un módulo independiente debido a que pretende ser una técnica independiente de la tecnología biométrica utilizada. Este se encargará de transformar la lista de usuario y convertirla en el GMT. Como se observa en la (Figura 36) este módulo se encarga de la creación de la lista GMT.

**Módulo GMT TRAIN**



**Diagrama de Contexto**

Figura 36: Módulo para la Creación del GMT

## 9.2 GMT: EVOLUCIÓN

La creación del módulo anterior y de la metodología de agrupación puede ser implementada de diferentes formas. El GMT es una lista de grupos formada por todos los nodos que forman el árbol de decisión. Cada usuario forma un grupo unitario. Cada grupo está formado por uno o dos grupos y a cada nodo se denomina grupo (Figura 37).

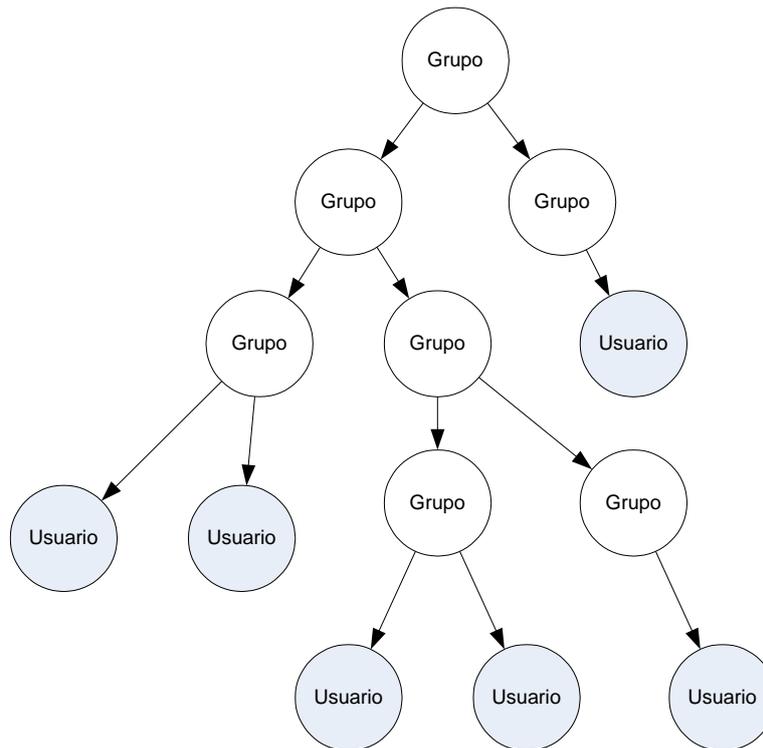


Figura 37: Árbol GMT

En este PFC se decidió adaptar la metodología a una estructura de árbol de decisión [J.R. Quinlan, 1986] con solución final (árbol binario ó árbol-B). Cada grupo posee como atributos un identificador de grupo (Figura 38), un identificador del grupo al que pertenece (denominado grupo padre), la ruta donde está almacenado el modelo del grupo y un valor que indica la rama del árbol en la que se encuentra (-1 izquierda; 1 derecha). De aquí en adelante se utilizará el término grupo para referirse tanto a usuarios como a nodos.

Esta pre-clasificación reduce el tiempo necesario para realizar una correcta identificación. Si se tiene una base de datos de 7 usuarios (Figura 38), con el método tradicional para identificar un individuo se necesitarían siete comparaciones. Por el contrario, si se genera el árbol GMT sobre la base de datos, el número de comparaciones se verá reducido de siete a tres comparaciones. De esta manera se realizan un 57.15% menos de comparaciones.

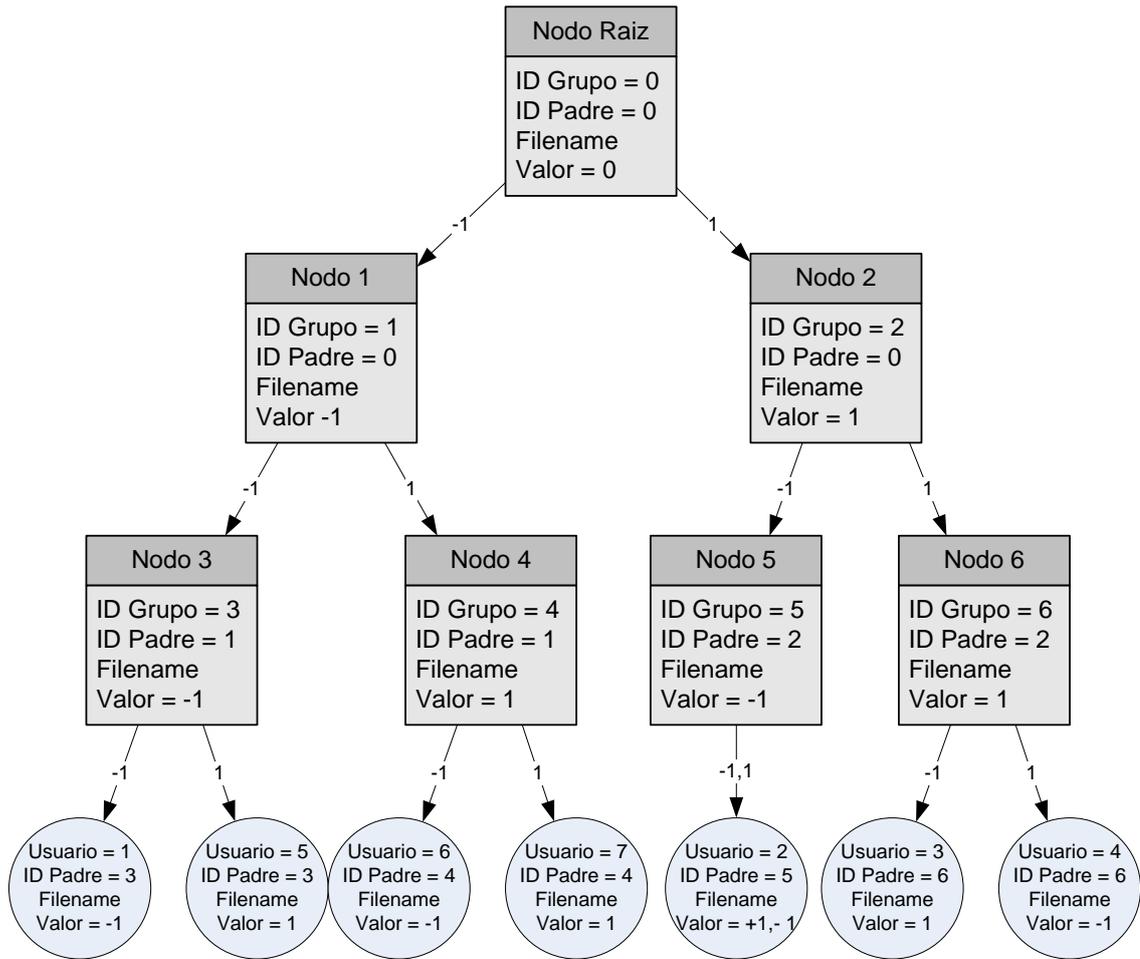


Figura 38: Esquema de un GMT con siete usuarios

El árbol descrito en la Figura 38 al construirlo en forma de lista GMT quedaría de la siguiente manera:

ID Grupo	ID Padre	Filename	Valor
1	1000002	E:\Filename\01.train.bin	1
2	1000001	E:\Filename\02.train.bin	1
2	1000001	E:\Filename\02.train.bin	-1
3	1000002	E:\Filename\03.train.bin	-1
1000000	0	E:\Filename\1000000.model.bin	0
1000001	1000000	E:\Filename\1000001.model.bin	-1
1000002	1000000	E:\Filename\1000002.model.bin	1

Tabla 4: Forma de la Lista que Constituye el GMT

El significado de los valores de la tabla son: El campo Valor indica la rama del árbol en la que se encuentra el grupo, el valor 0 en el campo ID Padre y Valor denota al nodo raíz. ID Grupo se refiere a cada nodo formado, también hace referencia a los usuarios finales. ID Padre hace referencia al nodo padre del que cuelga dicho grupo. Cuando un grupo unitario se replica en la lista GMT indicando que ambas direcciones



1. Se inicializa la lista destructiva a la lista de usuarios del sistema. Se elige un usuario al azar y se calculan sus distancias con respecto al resto.
2. El usuario cuya distancia entre modelos dio un valor menor formará un nuevo grupo con el usuario elegido al azar. Como un mismo usuario no puede estar en dos grupos, se eliminan de la lista destructiva aquellos que formaron un grupo.
3. Se comprueba si la lista destructiva está vacía, en ese caso se intercambian las listas. En caso contrario se comprueba cuantos grupos quedan aún en la lista. En caso de quedar uno, este pasaría a formar un grupo unitario. En caso de quedar dos usuarios, ambos pasan a formar un nuevo grupo. Si por el contrario quedan más de dos usuarios se vuelve al paso 1 hasta haber vaciado la lista destructiva.
4. Se comprueba si en la lista hay sólo dos grupos, en ese caso se crea el nodo raíz y se finaliza el algoritmo. En caso contrario, se vuelve al paso 1.

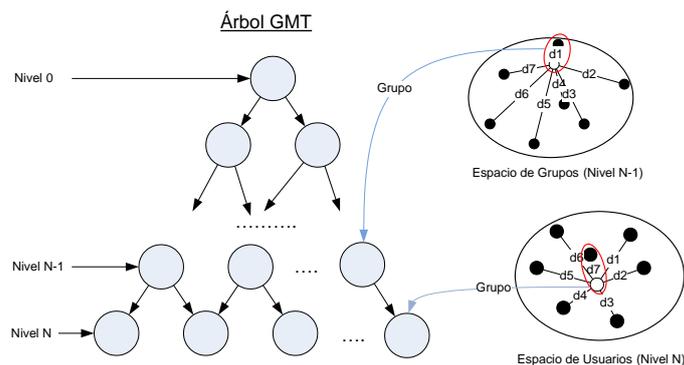


Figura 40: Esquema de Creación del B-GMT

Para generar los modelos de usuarios se utilizó GMM (véase, sección 3.3.3) de  $N$  gaussianas, con el fin de reducir el cálculo de distancias a un simple cálculo de distancias entre gaussianas. De entre las múltiples maneras de calcular dichas distancias se describen a continuación las utilizadas en el PFC.

### ***Distancia basada en la puntuación del GMM***

Al clasificador GMM se le introduce una muestra desconocida y devuelve un *score* que indica la probabilidad de que la muestra pertenezca al modelo con que se comparó. Por tanto, cuanto mayor sea el valor del *score* mayor es la semejanza entre la muestra y el modelo usado para comparar. Partiendo de la premisa anterior se estableció el siguiente supuesto: “Cuanto mayor sea el valor del *score*, mayor será la semejanza entre usuarios y más cercanos en el espacio se encontrarán”. Teniendo en cuenta todo lo

anterior, si se encuentran más cercanos es posible que esos usuarios posean características similares.

Para calcular esta distancia se debía calcular previamente los modelos GMM de todos los usuarios. A continuación se comparaba el modelo GMM del usuario elegido con los ficheros de entrenamiento de los restantes usuarios. El usuario cuya muestra obtenía un mayor *score* pasaba a formar grupo con el usuario elegido.

### ***Distancia basada en la puntuación del GMM normalizada***

Se realiza una normalización de las distancias calculadas en el apartado anterior, con respecto al usuario elegido. Para ello, se testea el modelo del usuario elegido con respecto a él mismo y la verosimilitud obtenida se utiliza para la normalización.

La normalización se realiza dividiendo el *score* obtenido en la comparación de la muestra con el modelo del usuario elegido, entre la verosimilitud resultante de testear el modelo del usuario elegido con el mismo. La fórmula utilizada es:

$$\text{Distancia Normalizada} = \frac{\text{distancia sin normalizar}}{\text{verosimilitud usuario}} \quad (\text{Ecuación 1})$$

Donde:

- *distancia sin normalizar* = Valor obtenido de testear el fichero de entrenamiento del usuario de la lista con el modelo del usuario elegido.
- *verosimilitud usuario* = Verosimilitud obtenida al testear el modelo del usuario elegido con él mismo.
- *Distancia Normalizada* = Es la distancia normalizada.

### ***Distancia de Mahalanobis***

La distancia de Mahalanobis se puede definir con una distancia Euclídea ponderada, donde la ponderación determinan el rango de variabilidad del punto tomado como referencia, expresados por la matriz de covarianzas [M. Wölfel et al, 2005]. En definitiva, la distancia de Mahalanobis calcula la distancia existente entre un punto determinado a una gaussiana, teniendo en cuenta la anchura de esta. La fórmula que rige esta distancia es:

$$\text{Distancia Mahalanobis} = D_M = (\underline{x} - \underline{m})^T V_{xy}^{-1} (\underline{x} - \underline{m}) \quad (\text{Ecuación 2})$$

En el caso del PFC, cada modelo GMM está formado por diez Gaussianas multidimensionales (27 dimensiones) y cada una de las gaussianas posee diferente importancia (peso) dentro del modelo. Al trabajar con múltiples gaussianas multidimensionales, se ideó compactar todas las gaussianas en una única gaussiana multidimensional con el fin de conseguir un único punto de referencia (Figura 41, punto

verde). Lo que se hizo fue crear una gaussiana media de todas las gaussianas con una única media y única matriz de covarianzas. La media y matriz de covarianzas única se calcularon de la siguiente manera:

$$\mu_g = \text{Media general} = \sum_i m_i \omega_i = \underline{m} \omega \quad (\text{Ecuación 3})$$

$$V_g = \text{Covarianza general} = \sum_i V_{xx}^i \omega_i = \underline{\underline{V_{xx}}} \omega \quad (\text{Ecuación 4})$$

Donde  $\underline{m}$  es el vector de medias de cada una de las gaussianas,  $\underline{\underline{V_{xx}}}$  la matriz de covarianzas de cada una de las gaussianas y  $\omega$  el vector de pesos. Consiguiendo de esta manera una gaussiana  $G(\mu_g, V_g)$ , de media  $\mu_g$  y varianza  $V_g$ . Esta nueva media ( $\mu_g$ ) se utilizará como punto de referencia quedando la (Ecuación 2) quedará de la siguiente forma:

$$\text{Distancia Mahalanobis} = D_M = (\mu_g - \underline{m})^T V_{xx}^{-1} (\mu_g - \underline{m}) \quad (\text{Ecuación 5})$$

Donde  $\mu_g$  es el punto medio con respecto al que se miden las distancias,  $\underline{m}$  es el vector de medias de la gaussiana a la que se mide la distancia y  $V_{xx}^{-1}$  es la inversa de la matriz de covarianza de la gaussiana.

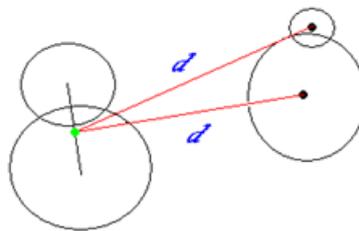
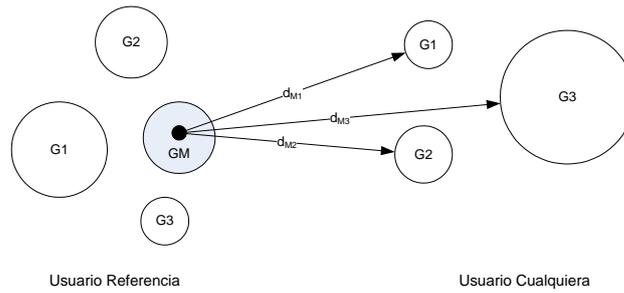


Figura 41: Distancia Punto Medio a Gaussianas

En esta versión, la distancia de Mahalanobis se aplicó de dos maneras diferentes. En la primera manera, se obtiene del usuario elegido el punto medio de referencia y se calculan las distancias con respecto a él de cada una de las gaussianas de los modelos de usuarios. En la segunda forma, por un lado se obtiene del usuario elegido el punto medio de referencia (GM, Figura 42) y se calculan las distancias con respecto a él de cada una de las gaussianas de los modelos de usuarios (Figura 42). Por otro lado se obtiene del usuario de la lista el punto medio de referencia y se calcula la distancia con respecto a las gaussianas que forman el modelo del usuario elegido. La distancia resultante es la media aritmética de ambas distancias. Para calcular esta distancia se partió de la premisa de que “no existe la misma distancia desde un punto de una gaussiana A hasta una gaussiana B, que desde un punto de la gaussiana B hasta la gaussiana A”. La Figura 24 muestra la manera descrita anteriormente, el usuario de referencia puede ser tanto el usuario elegido como el usuario de la lista.



**Figura 42: Forma Ideada para Cálculo de Distancias (caso particular de 3 Gaussianas/Usuario)**

La fórmula que se utilizó para el cálculo de la distancia de Mahalanobis es la siguiente:

$$D_M = \frac{1}{K} \sum_{i=0}^K \left[ (\underline{\mu}_g - \underline{m}^i \omega^i) V_{xx_k}^{-1} (\underline{\mu}_g - \underline{m}^i \omega^i)^T \right] = \frac{1}{K} \left[ \left( \underline{\mu}_g - \sum_{i=0}^K \underline{m}^i \omega^i \right) \sum_{i=0}^K V_{xx_k}^{-1} \left( \underline{\mu}_g - \sum_{i=0}^K \underline{m}^i \omega^i \right)^T \right]$$

(Fórmula 1)

### ***Distancia Euclídea***

Es la medida que más comúnmente se utiliza, se aplica en multitud de algoritmos (por ejemplo, VQ [M. Filippone et al, 2007]). Al igual que la distancia de Mahalanobis, también es utilizada para medir la distancia o similitud entre dos objetos [Rui Xu et al, 2005]. Concretamente, la distancia Euclídea mide la distancia existente entre dos puntos del espacio euclídeo.

$$Distancia\ Euclídea = D_E = \sqrt{(\underline{x} - \underline{m})^T (\underline{x} - \underline{m})} \quad (\text{Ecuación 6})$$

Se decidió aplicar esta medida de similitud como la distancia existente entre las medias de las gaussianas. La manera de aplicarlo es muy similar a la descrita en el apartado anterior como segunda forma, la única diferencia reside en que no se necesita el cálculo de la covarianza media. Por un lado se obtiene del usuario elegido el punto medio de referencia (Ecuación 6) y se calculan las distancias con respecto a él de cada una de las medias de las gaussianas de los modelos de usuarios. Por otro lado se obtiene del usuario de la lista el punto medio de referencia (Ecuación 6) y se calcula la distancia con respecto a las medias de las gaussianas que forman el modelo del usuario elegido. La distancia resultante es la media aritmética de ambas distancias. Se partió también de la premisa “no existe la misma distancia desde un punto de una gaussiana A hasta una gaussiana B, que desde un punto de la gaussiana B hasta la gaussiana A”. El cálculo es similar al que se muestra en la Figura 42 pero sólo utilizando el vector de medias tanto de la gaussiana media como de las que forman un usuario cualquiera.

La fórmula utilizada para el cálculo de la distancia Euclídea es la descrita a continuación:

$$D_E = \sqrt{\frac{1}{K} \sum_{i=0}^K [(\underline{\mu}_g - m^i \omega^i)(\underline{\mu}_g - m^i \omega^i)^T]} = \sqrt{\frac{1}{K} \left[ \left( \underline{\mu}_g - \sum_{i=0}^K m^i \omega^i \right) \left( \underline{\mu}_g - \sum_{i=0}^K m^i \omega^i \right)^T \right]}$$

(Fórmula 2)

### Distancia de Manhattan o City-Block

La distancia de Manhattan o city-block es preferida a la Euclidiana debido a su simplicidad de cálculo. Por el contrario, el error en la aproximación es mayor que en la Euclídea.

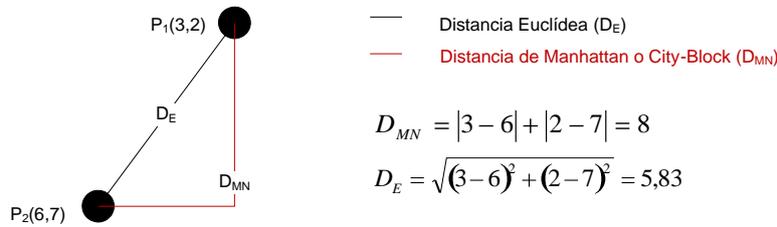


Figura 43: Distancia Euclídea Vs Manhattan

Para esta métrica, la distancia entre dos puntos cualesquiera es la medida a lo largo de los ejes de coordenadas (y por lo tanto en ángulo recto) (Figura 43), siendo fácilmente generalizable para cualquier dimensión del espacio de medida [A. Zilinskas et al, 2006]. La ecuación de esta medida es:

$$\text{Distancia de Manhattan} = D_{MN} = \sum_{i=0}^N |x_{ia} - x_{ib}| \quad (\text{Ecuación 7})$$

Donde  $\underline{x}_a$  y  $\underline{x}_b$  son en este caso los vector de medias de las gaussianas A y B. En este caso, sólo se calcularía la distancia existente entre los vectores de medias de las gaussianas. Es decir, distancia entre la media de la gaussiana 0 del usuario de referencia con respecto a la media de la gaussiana 0 del usuario de la lista, distancia entre la media de la gaussiana 1 del usuario de referencia con respecto a la media de la gaussiana 1 del usuario de la lista... y así hasta terminar con el número de gaussianas (Figura 44). Obteniendo así una distancia media global entre las gaussianas multidimensionales que forman los modelos de los usuarios. La fórmula que describe este procedimiento es:

$$D_{MN} = \frac{1}{N} \sum_{i=0}^N \sum_{j=0}^K |m_A^i[j] - m_B^i[j]|$$

(Fórmula 3)

Donde  $\underline{m}_A^i$  y  $\underline{m}_B^i$  son los vector de medias de la gaussiana  $i$  de los usuarios A y B respectivamente.

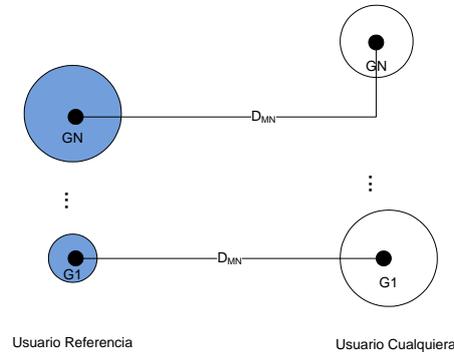


Figura 44: Cálculo de Distancia City-Block (Bottom-Up)

Tanto la distancia de Manhattan como la distancia Euclídea son un caso particular de la distancia de Minkowski [Rui Xu et al, 2005] [Arslan Brömme, 2005]:

$$Distancia\ de\ Minkowski = D_p = \left[ \sum_{i=1}^M (x_i - y_i)^p \right]^{\frac{1}{p}}$$

(Fórmula 4)

Casos particulares:

$$p = 1; \quad Distancia\ de\ Manhattan = D_{MN} = \sum_{i=1}^M |x_i - y_i|$$

$$p = 2; \quad Distancia\ Euclidea = D_E = \sqrt{\sum_{i=1}^M (x_i - y_i)^2}$$

## 9.2.2 Conclusiones

Después de elaborar, implementar y evaluar todo lo descrito anteriormente, se optó por desechar esta primera versión por los malos resultados obtenidos. La versión B-GMT poseía unos ratios del 100% de falso rechazo. Estos malos resultados pueden ser debidos a:

1. Una representación pobre del modelado de usuarios. Podría necesitarse un mayor número de gaussianas para representar a un usuario. Como contrapartida se podría sobre-ajustar.
2. Una representación sencilla de la verosimilitud obtenida por el GMM como semejanza, sin tener en cuenta la coincidencia de la distribución completa de las gaussianas.

3. Una selección de vectores poco representativa a la hora de crear el fichero de entrenamiento del grupo. Por tanto, la generación del modelo SVM realiza una pobre generalización.
4. Una mala elección en el modelado de la distancia entre funciones de distribución, optando por un modelo demasiado sencillo que podría eliminar información relevante.
5. Un mal planteamiento de la solución.

### 9.2.3 T-GMT (Top-Down)

A la vista de los resultados anteriores se decidió cambiar la línea de desarrollo inicial y abordar la creación del GMT desde otra perspectiva. El nuevo planteamiento proponía la creación del GMT desde arriba hacia abajo (Top-Down) (desde la raíz hasta las hojas).

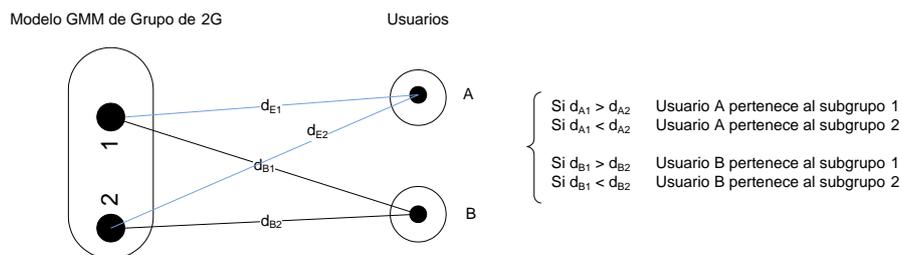


Figura 45: Esquema de Creación del T-GMT

Para cada nodo, o grupo del GMT, se crea un modelo GMM de dos gaussianas y se miden las distancias euclídea desde el punto medio de cada usuario a cada una de las gaussianas del nodo. La cercanía a cada gaussiana determina el subgrupo al que pertenece cada usuario (Figura 45). Para cada grupo se crea un fichero de entrenamiento formado por 400 vectores de cada usuario que pertenece a dicho grupo. Con los ficheros de entrenamiento de los grupos hijos se genera el modelo SVM del grupo padre (que es el que se utiliza para testear). A continuación se describe el funcionamiento del algoritmo (Figura 46):



8. Se crea el modelo SVM del grupo padre mediante los ficheros de entrenamiento de los grupos hijos.
9. Se comprueba si la lista ha terminado, en caso contrario se vuelve al paso 2.

Para el cálculo de las distancias entre los usuarios y las medias de las gaussianas del modelo de grupo se utilizó la Fórmula 2. Se evaluaron los resultados obtenidos y se decidió no implementar ninguna otra mejora.

#### 9.2.4 Conclusiones

Después de la implementación y evaluación de esta nueva metodología, se optó por desecharla debido a los pobres resultados obtenidos. Aún así mejoraban considerablemente con respecto a la versión anterior ( $FRR = 66\%$ ), no eran suficientes para utilizarlos como línea base. Estos pobres resultados pudieron ser debidos a:

1. Una mala elección en el modelado de la distancia entre funciones de distribución, optando por un modelo demasiado sencillo que podría eliminar información relevante.
2. Una selección de vectores poco representativa para crear el fichero de entrenamiento del grupo que representa un espacio de usuarios elevado. Por tanto, mala generalización por parte del clasificador GMM.
3. Una insuficiente resolución del modelo GMM de cada grupo, pudiéndose necesitar un mayor número de gaussianas.
4. Una mala generalización del espacio de usuarios optando por un modelo sencillo y aglomerativo de agrupación.
5. Una mala clasificación de vectores lo cual no es característico (se eligen vectores de muchos usuarios y se modelan) de cada usuario, perdiendo resolución discriminativa por usuario.

Como conclusión final se puede decir que se eligió un buen camino de investigación pero la línea de desarrollo no fue la adecuada.

#### 9.2.5 T-NGMT (Top-Down)

A la vista de los resultados alentadores de la versión anterior se decidió continuar con la misma línea de investigación (top-down) pero variando en la manera de generar el árbol. En la versión anterior se modelaba vectores pertenecientes a un conjunto de usuarios, lo que se resumía en una clasificación de vectores. Para esta nueva versión se

decidió ajustar los usuarios al algoritmo de tal manera que se clasificarían usuarios y no vectores.

La idea se basa en elegir los usuarios que se encuentren más alejados entre ellos, con el fin de obtener una frontera de decisión lo suficientemente discriminativa para poder realizar una clasificación mediante SVM (Figura 47). Ambos usuarios (los dos más alejados) formarían el modelo SVM del grupo. Los usuarios se clasificarían en los diferentes subgrupos en función de la probabilidad obtenida de testarlos con el clasificador SVM.

Para aumentar el poder discriminativo del algoritmo, se aplicó una técnica replicante de los usuarios dudosos. Para ello se establecieron umbrales; sabiendo que el valor de la verosimilitud varía entre 0-1, se estableció una partición en tres de ese espacio quedando así:  $[0 - 0.33]$   $[0.33 - 0.66]$   $[0.66 - 1]$ . Los usuarios cuya verosimilitud este comprendida en el rango  $[0.33 - 0.66]$  se replican y pasa a formar parte de ambos subgrupos. Se supuso que los usuarios pertenecientes a ese rango no se asemejan en demasía a ninguno de los dos subgrupos.

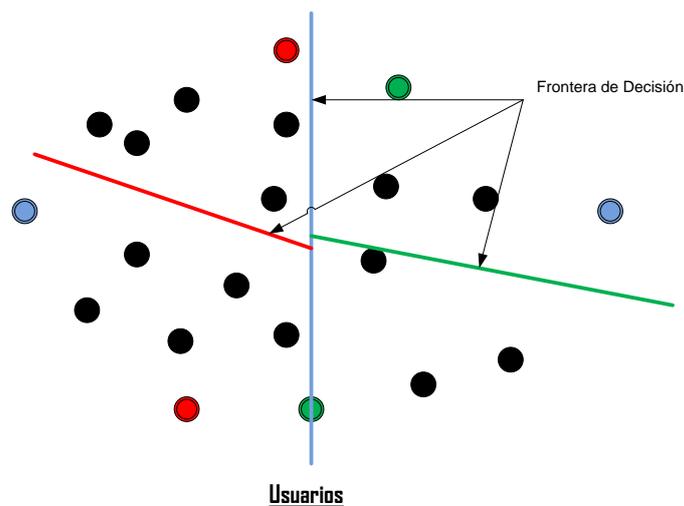


Figura 47: Esquema de Creación del T-NGMT

Aplicar la técnica replicante influye en el proceso con un aumento del número de iteraciones necesarias para converger, repercutiendo en el número de niveles del árbol y por tanto, en el tiempo de identificación. Pero aún con este pequeño aumento se reduce de manera drástica el tiempo de identificación con respecto al método tradicional. A continuación se describe el funcionamiento del algoritmo (Figura 48):

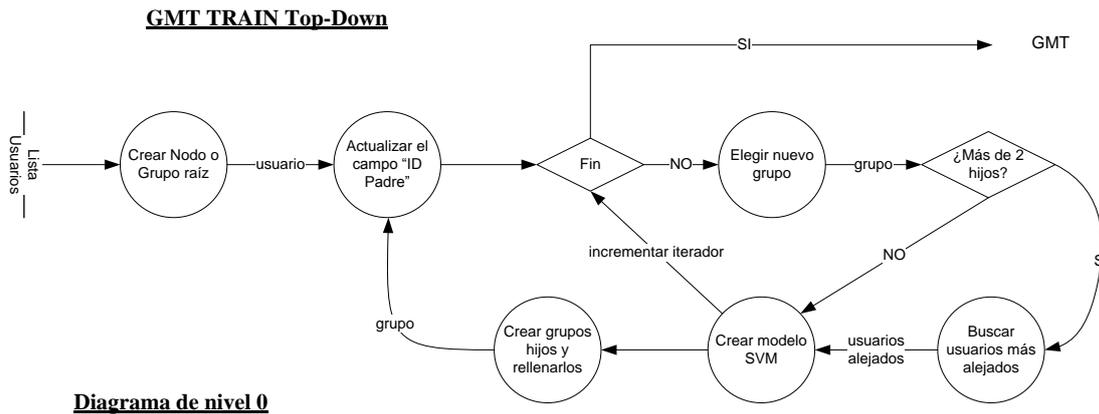


Figura 48: Diagrama de Procesos T-NGMT

1. Se crea el grupo o nodo raíz. Se actualiza el campo “ID Padre” de todos los usuarios al identificador del grupo raíz.
2. Se elige grupo y se comprueba si posee más de dos usuarios. En ese caso, se buscan los dos usuarios más lejanos. En caso contrario, si hay dos usuarios se forma el modelo SVM del grupo con los ficheros de entrenamiento de los usuarios y se elige el siguiente grupo. Si hay un usuario se forma el modelo SVM del grupo con el fichero de entrenamiento del usuario y se elige el siguiente grupo.
3. Se crea el modelo SVM del grupo con los dos usuarios más alejados del mismo, utilizando sus ficheros de entrenamiento.
4. Se crean los subgrupos hijos que pertenecen al grupo.
5. Se testean los usuarios del grupo con el fichero SVM del mismo. Dependiendo de la verosimilitud obtenida pertenecen a uno u otro subgrupo ó ambos.
6. Se comprueba si es el final de la lista. En ese caso, se termina el algoritmo. En caso contrario, se vuelve al paso 2.

Con la implementación de este algoritmo se obtuvieron unos resultados alentadores. Dichos resultados mejoraban en un 70% la primera versión y un 40% la segunda versión, por lo que se decidió tomarla como línea base de investigación.

Esta nueva versión generalizaba mejor a nivel de usuario en vez de a nivel de vectores. Consiguiendo ajustar la información de cada usuario al algoritmo de pre-clasificación. También se mantenía la resolución óptima de gaussianas para cada modelo GMM evitando el fallo del algoritmo por falta de resolución discriminativa de los modelos.

Mediante la técnica de replicación se conseguía un aumento en la probabilidad de acierto de usuarios dudosos (debido a que se replicarían en todos los nodos en que siembren duda). Por tanto se decidió continuar con esta línea de investigación y se fueron introduciendo posibles mejoras para el proceso.

### 9.3 MEJORAS PARA EL T-NGMT

Una vez elegida la línea base sobre la que continuar con la investigación, se decidió realizar posibles mejoras con el fin de mejorar y optimizar el algoritmo. En este apartado se describen las posibles mejoras desarrolladas para el T-NGMT. Estas mejoras se dividen en dos grupos: Distinta forma de calcular las distancias entre usuarios y distinta manera de selección de usuarios.

#### 9.3.1 Distintas maneras de calcular distancias entre usuarios

En esta sección se describen las diferentes alternativas utilizadas para establecer que usuarios se encuentran más alejados entre sí en el espacio de usuarios. Las medidas de similitud o distancia utilizadas son la Euclídea, la de Mahalanobis y la de Manhattan o city block. A continuación se describirá las versiones y fórmulas utilizadas para calcular la distancia entre dos usuarios (A y B) cualesquiera del sistema.

##### *Distance Between Middles 1*

Para esta primera alternativa, se crea una gaussiana promedia (mediante (Ecuación 3) (Ecuación 4)) para cada usuario. A continuación se calcula la distancia (Euclídea, Mahalanobis, Manhattan) existente entre las gaussianas promedias o los vectores de medias de las gaussianas promedias de cada usuario. Se puede observar dicho procedimiento en la Figura 49 donde las circunferencias de color azul representan a las gaussianas promedio.

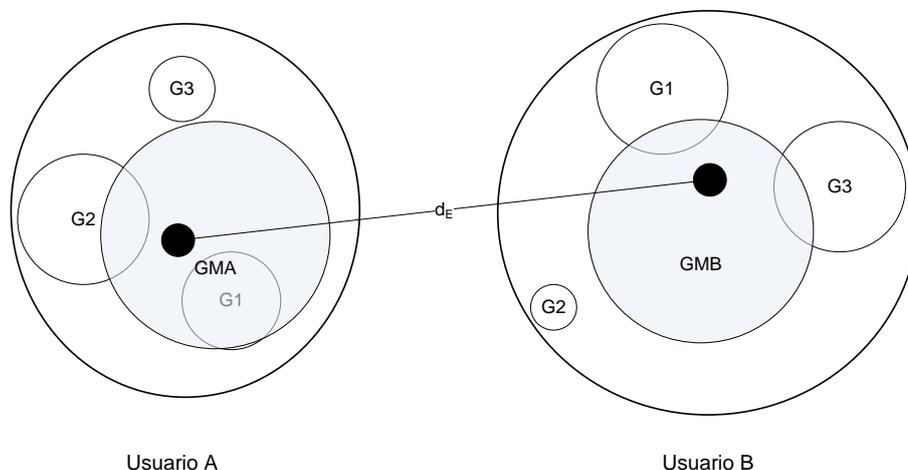


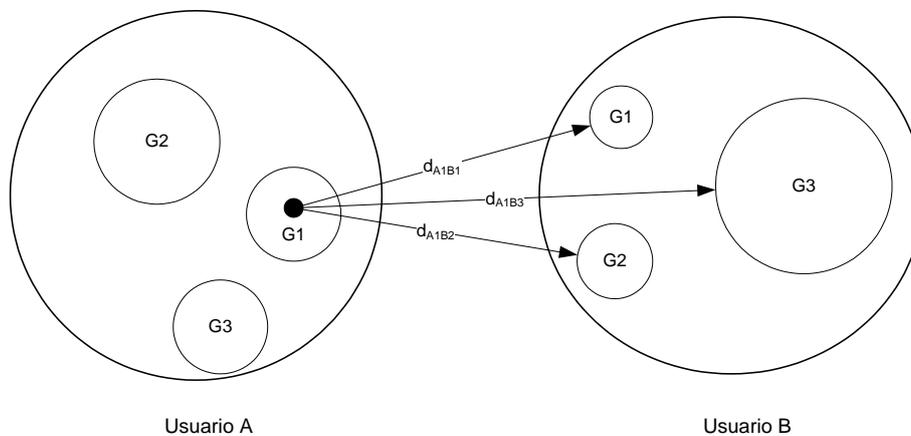
Figura 49: Representación Gráfica de la Distance Between Middles 1

***Distance Between Middles 2***

Esta alternativa es una variante de la anterior, se basa en la premisa de que la distancia existente entre dos gaussianas es diferente dependiendo de la gaussiana tomada como referencia. Se calcula la distancia existente entre la gaussiana promedio, o el vector de medias de las gaussianas promedio, del usuario A hasta el B. También se calcula la distancia existente entre la gaussiana promedio, o el vector de medias de las gaussianas promedio, del usuario B hasta el A. Por tanto, se toma como distancia la media aritmética de ambas distancias. Esta alternativa sólo es utilizada en la distancia de Mahalanobis ya que tiene en cuenta la anchura de las gaussianas.

***Distance Between Gaussians 1***

Para esta nueva alternativa se calcula la distancia como la media del promedio de la distancia existente entre cada una de las gaussianas que forman un usuario contra todas las gaussianas del otro. Supóngase dos usuarios A y B cuyos modelos están formados por 10 gaussianas. Se calcula un promedio entre las distancias de la gaussiana 1 del usuario A, a todas las gaussianas del usuario B; se continúa calculando el promedio de distancias entre la gaussiana 2 del usuario A, a todas las gaussianas del usuario B (Figura 50)...se continúa así hasta haber terminado con todas las gaussianas del usuario A. Se utiliza como distancia la media de todos los promedios calculados.



**Figura 50: Representación Gráfica de la Distance Between Gaussians 1**

***Distance Between Gaussians 2***

Es una variante de la anterior. Parte de la misma premisa que Distance between middles 2. Se realiza el proceso descrito en Distance between gaussians 1 pero tomando también como referencia el usuario B. Se toma como distancia la media aritmética de ambas.

### *Distance Middle to Gaussians 1*

En primer lugar, se calcula una gaussiana promedio del usuario A (mediante Ecuación 3) (Ecuación 4)). A continuación, se calcula la distancia entre la gaussiana promedio del primer usuario (A) a cada una de las gaussianas que forman al segundo usuario (Figura 51). El promedio de todas las distancias se usa como distancia.

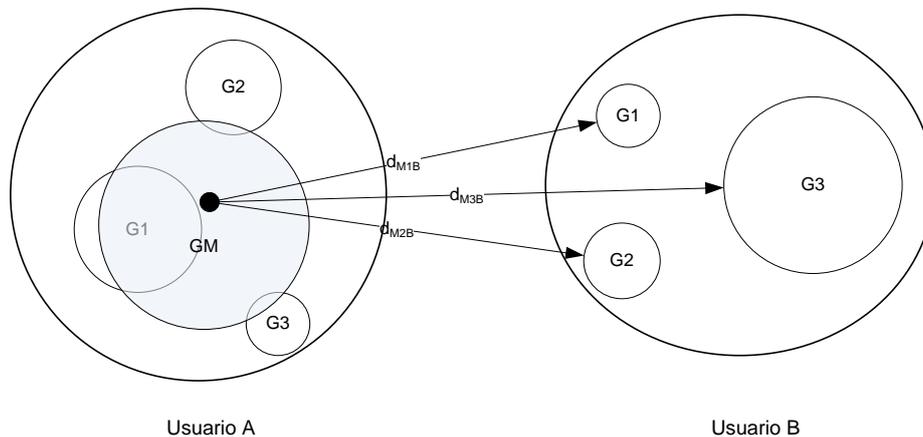


Figura 51: Representación Gráfica de Distance Middle to Gaussians 1

### *Distance Middle to Gaussians 2*

Es una variante de la anterior. Parte de la misma premisa que en Distance between middles 2. Se realiza el proceso descrito en Distance middle to gaussians 1, pero tomando también como referencia el usuario B. Por tanto se obtienen dos distancias ( $D_{A-B}$ ,  $D_{B-A}$ ). Se toma como distancia la media aritmética de ambas.

## 9.3.2 Implementación Matemática

En este apartado se describen las fórmulas matemáticas desarrolladas para adaptar las distancias elegidas (Euclídea, Mahalanobis, Manhattan o City Block) a las alternativas de cálculo ideadas.

### **DISTANCIA EUCLÍDEA**

En este apartado se describe los distintos algoritmos creados para calcular los usuarios más alejados aplicando la distancia Euclídea (Ecuación 6) a las alternativas del apartado 9.3.1.

### **Distancia Euclídea versión 1 (v1)**

Para este primer cálculo, se aplica la distancia euclídea a Distance between middles 1. La fórmula desarrollada para dicho cálculo es:

$$D_{E1} = \sqrt{(\underline{\mu}_j^A - \underline{\mu}_j^B)(\underline{\mu}_j^A - \underline{\mu}_j^B)^T} \quad \text{donde} \quad \underline{\mu} = \sum_{k=0}^M m_k \omega_k = \underline{m} \omega$$

(Fórmula 5)

Donde  $\underline{\mu}$  es la media de la gaussiana promedio (Ecuación 3),  $\underline{m}$  y  $\underline{\omega}$  los vectores de medias y pesos de las gaussianas que forman al usuario.

### **Distancia Euclídea versión 2 (v2)**

En este caso se aplica la distancia euclídea a Distance between gaussians 1. La fórmula que implementa dicho cálculo es:

$$D_{E2} = \sqrt{\frac{1}{N} \sum_{i=0}^N \frac{1}{K} \sum_{j=0}^K (\omega_i^A \underline{m}_i^A - \omega_i^B \underline{m}_i^B)(\omega_i^A \underline{m}_i^A - \omega_i^B \underline{m}_i^B)^T}$$

(Fórmula 6)

Donde  $\underline{m}_i$  y  $\underline{m}_j$  son los vectores de medias de cada una de las gaussianas que forman el modelo de usuario,  $\omega_i$  y  $\omega_j$  los pesos respectivos para cada una de ellas.

### **Distancia Euclídea versión 3.0 (v3.0)**

En este caso se aplica la distancia euclídea a Distance middle to gaussians 2. Las fórmulas que implementan dicho cálculo son:

$$D_{E3} = \frac{D_A + D_B}{2}$$

(Fórmula 7)

Donde  $D_A$  y  $D_B$  son:

$$D_A = \sqrt{\frac{1}{K} \sum_{j=0}^K (\underline{\mu}_A - \omega_j^B \underline{m}_j^B)(\underline{\mu}_A - \omega_j^B \underline{m}_j^B)^T} \quad \text{donde} \quad \underline{\mu}_A = \sum_{k=0}^M \omega_k^A \underline{m}_k^A = \underline{m}^A \omega^A$$

(Fórmula 8)

$$D_B = \sqrt{\frac{1}{K} \sum_{j=0}^K (\underline{\mu}_B - \omega_j^A \underline{m}_j^A) (\underline{\mu}_B - \omega_j^A \underline{m}_j^A)^T} \quad \text{donde} \quad \underline{\mu}_B = \sum_{k=0}^M \omega_k^B \underline{m}_k^B = \underline{m}^B \underline{\omega}^B$$

(Fórmula 9)

Donde  $\underline{m}_i$  y  $\underline{m}_j$  son los vectores de medias de cada una de las gaussianas que forman el modelo de usuario,  $\omega_i$  y  $\omega_j$  los pesos respectivos para cada una de ellas.

### **Distancia Euclídea versión 3.1 (v3.1)**

En este caso se aplica la distancia euclídea a Distance middle to gaussians 1. Las fórmulas que implementan dicho cálculo son la (Fórmula 8) ó (Fórmula 9) dependiendo del usuario que se tome como referencia.

### **DISTANCIA DE MAHALANOBIS**

En este apartado se describe los distintos algoritmos creados para calcular los usuarios más alejados aplicando la distancia Mahalanobis (Ecuación 2) a las alternativas del apartado 9.3.1.

### **Distancia Mahalanobis versión 1.0 (v1.0)**

Para este primer cálculo, se aplica la distancia de Mahalanobis a Distance between middles 2. Las fórmulas desarrolladas para dicho cálculo son:

$$D_{M10} = \frac{D_A + D_B}{2}$$

(Fórmula 10)

Donde  $D_A$  y  $D_B$  son:

$$D_A = (\underline{\mu}^A - \underline{\mu}^B) V_B^{-1} (\underline{\mu}^A - \underline{\mu}^B)^T$$

(Fórmula 11)

$$D_B = (\underline{\mu}^B - \underline{\mu}^A) V_A^{-1} (\underline{\mu}^B - \underline{\mu}^A)^T$$

(Fórmula 12)

Y el valor de  $\underline{\mu}_A$  y  $\underline{\mu}_B$  es (Ecuación 3):

$$\underline{\mu}_A = \sum_{k=0}^M \omega_k^A \underline{m}_k^A = \underline{m}^A \underline{\omega}^A$$

$$\underline{\mu}_B = \sum_{k=0}^M \omega_k^B \underline{m}_k^B = \underline{m}^B \underline{\omega}^B$$

Donde  $\underline{m}$  y  $\underline{\omega}$  los vectores de medias y pesos de las gaussianas que forman al usuario.

### **Distancia Mahalanobis versión 1.1 (v1.1)**

Para este primer cálculo, se aplica la distancia de Mahalanobis a Distance between middles 1. Las fórmulas que implementan dicho cálculo son la (Fórmula 11) ó (Fórmula 12) dependiendo del usuario que se tome como referencia.

### **Distancia Mahalanobis versión 2.0 (v2.0)**

Para este primer cálculo, se aplica la distancia de Mahalanobis a Distance middle to gaussians 2. Las fórmulas desarrolladas para dicho cálculo son:

$$D_{M20} = \frac{D_A + D_B}{2}$$

**(Fórmula 13)**

Donde  $D_A$  y  $D_B$  son:

$$D_A = \frac{1}{K} \sum_{j=0}^K \left( \underline{\mu}_A - \omega_j^B \underline{m}_j^B \right) V_j^{-1} \left( \underline{\mu}_A - \omega_j^B \underline{m}_j^B \right)^T \quad \text{donde} \quad \underline{\mu}_A = \sum_{k=0}^M \omega_k^A \underline{m}_k^A = \underline{m}^A \underline{\omega}^A$$

**(Fórmula 14)**

$$D_B = \frac{1}{K} \sum_{j=0}^K \left( \underline{\mu}_B - \omega_j^A \underline{m}_j^A \right) V_j^{-1} \left( \underline{\mu}_B - \omega_j^A \underline{m}_j^A \right)^T \quad \text{donde} \quad \underline{\mu}_B = \sum_{k=0}^M \omega_k^B \underline{m}_k^B = \underline{m}^B \underline{\omega}^B$$

**(Fórmula 15)**

Donde  $\underline{m}_i$  y  $\underline{m}_j$  son los vectores de medias de cada una de las gaussianas que forman el modelo de usuario,  $\omega_i$  y  $\omega_j$  los pesos respectivos para cada una de ellas.

### **Distancia Mahalanobis versión 2.1 (v2.1)**

Para este segundo cálculo, se aplica la distancia de Mahalanobis a Distance middle to gaussians 1. Las fórmulas que implementan dicho cálculo son la (Fórmula 14) ó (Fórmula 15) dependiendo del usuario que se tome como referencia.

### **Distancia Mahalanobis versión 3.0 (v3.0)**

Para este primer cálculo, se aplica la distancia de Mahalanobis a Distance between gaussians 2. Las fórmulas desarrolladas para dicho cálculo son:

$$D_{M20} = \frac{D_A + D_B}{2}$$

(Fórmula 16)

Donde  $D_A$  y  $D_B$  son:

$$D_A = \frac{1}{N} \sum_{i=0}^N \frac{1}{K} \sum_{j=0}^K \left( \omega_i^A \underline{m}_i^A - \omega_i^B \underline{m}_i^B \right) V_j^{B^{-1}} \left( \omega_i^A \underline{m}_i^A - \omega_i^B \underline{m}_i^B \right)^T$$

(Fórmula 17)

$$D_B = \frac{1}{N} \sum_{i=0}^N \frac{1}{K} \sum_{j=0}^K \left( \omega_i^B \underline{m}_i^B - \omega_i^A \underline{m}_i^A \right) V_j^{A^{-1}} \left( \omega_i^B \underline{m}_i^B - \omega_i^A \underline{m}_i^A \right)^T$$

(Fórmula 18)

Donde  $\underline{m}_i$  y  $\underline{m}_j$  son los vectores de medias de cada una de las gaussianas que forman el modelo de usuario,  $\omega_i$  y  $\omega_j$  los pesos respectivos para cada una de ellas.

### **Distancia Mahalanobis versión 3.1 (v3.1)**

Para este segundo cálculo, se aplica la distancia de Mahalanobis a Distance between gaussians 1. Las fórmulas que implementan dicho cálculo son la (Fórmula 17) ó (Fórmula 18) dependiendo del usuario que se tome como referencia.

### **DISTANCIA DE MANHATTAN O CITY-BLOCK**

En este apartado se describe los distintos algoritmos creados para calcular los usuarios más alejados aplicando la distancia Manhattan (Ecuación 9) a las alternativas del apartado 9.3.1.

### **Distancia Manhattan versión 1 (v1)**

Para este cálculo, se aplica la distancia de Manhattan a Distance between middles 1. Las fórmulas desarrolladas para dicho cálculo son:

$$D_{MN1} = \sum_{j=0}^K |\underline{\mu}^A[j] - \underline{\mu}^B[j]|$$

**(Fórmula 19)**

Y el valor de  $\underline{\mu}_A$  y  $\underline{\mu}_B$  es (Ecuación 3):

$$\underline{\mu}_A = \sum_{k=0}^M \omega_k^A \underline{m}_k^A = \underline{m}^A \omega^A$$

$$\underline{\mu}_B = \sum_{k=0}^M \omega_k^B \underline{m}_k^B = \underline{m}^B \omega^B$$

Donde  $\underline{m}$  y  $\omega$  los vectores de medias y pesos de las gaussianas que forman al usuario respectivamente.

### **Distancia Manhattan versión 2 (v2)**

Para este cálculo, se aplica la distancia de Manhattan a Distance between middles 1. Las fórmulas desarrolladas para dicho cálculo son:

$$D_{MN2} = \frac{1}{M} \sum_{k=0}^M \frac{1}{N} \sum_{i=0}^N \sum_{j=0}^K |\underline{m}_A^k[j] - \underline{m}_B^i[j]| = \frac{1}{MN} \sum_{k=0}^M \sum_{i=0}^N \sum_{j=0}^K |\underline{m}_A^k[j] - \underline{m}_B^i[j]|$$

**(Fórmula 20)**

Donde  $\underline{m}_A^k$  y  $\underline{m}_B^i$  son los vectores de medias de las gaussianas  $k$  e  $i$  ( $k \neq i$ ) de los usuarios A y B respectivamente.

### **Distancia Manhattan versión 3.0 (v3.0)**

Para este cálculo, se aplica la distancia de Manhattan a Distance middle to gaussians 1. Las fórmulas desarrolladas para dicho cálculo son:

$$D_{MN30} = \frac{1}{N} \sum_{i=0}^N \sum_{j=0}^K |\underline{\mu}^A[j] - \underline{m}_i^B[j]| \quad \text{donde} \quad \underline{\mu}_A = \sum_{k=0}^M \omega_k^A \underline{m}_k^A = \underline{m}^A \omega^A$$

**(Fórmula 21)**

Donde  $\underline{m}_i^B$  y  $\underline{\mu}^A$  son los vectores de medias de cada gaussiana  $i$  del usuario B y el vector de medias de la gaussiana promedio del usuario A, respectivamente.

### **Distancia Manhattan versión 3.1 (v3.1)**

Para este cálculo, se aplica la distancia de Manhattan a Distance middle to gaussians  
2. Las fórmulas desarrolladas para dicho cálculo son:

$$D_{MN31} = \frac{D_A + D_B}{2}$$

(Fórmula 22)

Donde  $D_A$  y  $D_B$  son:

$$D_A = \frac{1}{N} \sum_{i=0}^N \sum_{j=0}^K \left| \underline{\mu}^A[j] - \underline{m}_i^B[j] \right| \quad \text{donde} \quad \underline{\mu}_A = \sum_{k=0}^M \omega_k^A \underline{m}_k^A = \underline{m}^A \omega^A$$

(Fórmula 23)

$$D_B = \frac{1}{N} \sum_{i=0}^N \sum_{j=0}^K \left| \underline{\mu}^B[j] - \underline{m}_i^A[j] \right| \quad \text{donde} \quad \underline{\mu}_B = \sum_{k=0}^M \omega_k^B \underline{m}_k^B = \underline{m}^B \omega^B$$

(Fórmula 24)

Donde  $\underline{\mu}_A$  y  $\underline{\mu}_B$  son el vector de medias de la gaussiana promedio del usuario A y B,  $\underline{m}_k^A$  y  $\underline{m}_k^B$  el vector de medias de la gaussiana  $k$  que forman al usuario A y B, y  $\omega_k^A$  y  $\omega_k^B$  el peso de la gaussiana  $k$  que forma parte del usuario A y B, respectivamente.

### **9.3.3 Evaluación de los Resultados**

Se realizaron un total de 14 pruebas. Una por cada forma ideada para el cálculo de las distancias. Las tablas que se muestran a continuación presentan la tasa FNMR, el tiempo promedio de comparación por muestra y el tiempo total en realizar la comparación total.

<b>TASA DE FALSO NO RECONOCIMIENTO (FNMR)</b>			
	<b>Euclídea</b>	<b>Mahalanobis</b>	<b>Manhattan o City-Block</b>
1.0	41,17%	42,75%	40,66%
1.1		41,25%	
2.0	38,79%	47,22%	38,33%
2.1		43,58%	
3.0	42,84%	42,90%	44,18%
3.1	40,99%	42,51%	46,03%

Tabla 5: Comparativa de FNMR del Primer Experimento

<b>TASA DE FALSO RECONOCIMIENTO (FMR)</b>			
	<b>Euclídea</b>	<b>Mahalanobis</b>	<b>Manhattan o City-Block</b>
1.0	21,17%	30,16%	21,08%
1.1		41,39%	
2.0	20,10%	24,48%	19,87%
2.1		38,59%	
3.0	28,21%	22,24%	22,90%
3.1	21,25%	22,04%	23,86%

Tabla 6: Comparativa de FMR del Primer Experimento

<b>HALF TOTAL ERROR RATE (HTER)</b>			
	<b>Euclídea</b>	<b>Mahalanobis</b>	<b>Manhattan o City-Block</b>
1.0	31,17%	36,45%	30,87%
1.1		41,32%	
2.0	29,44%	35,85%	29,10%
2.1		41,08%	
3.0	35,47%	32,57%	33,54%
3.1	31,12%	32,27%	34,94%

Tabla 7: Comparativa de HTER del Primer Experimento

Los resultados de la Tabla 5 y Tabla 6 no son tan alentadores como se esperaban. Se observa que en las versiones 2.0 de las distancias Euclídea y Manhattan se consiguieron unos mejores resultados. Esta mejoría quizás fue debida a que se tuvieron en cuenta las medias de todas las gaussianas que forman el modelo, en función de su importancia en el mismo. De esta forma, no se pierde resolución en el modelado del usuario consiguiendo una discriminación más específica. Esto mismo no ocurre con Mahalanobis en su misma versión. Esto puede ser debido al desplazamiento que sufren las gaussianas al multiplicarlas por su peso (importancia en el modelo). Todo ello repercute en un desplazamiento del usuario en el espacio, obteniendo así una aproximación no muy característica del propio usuario.

<b>TIEMPO UTILIZADO (min)</b>			
	<b>Euclídea</b>	<b>Mahalanobis</b>	<b>Manhattan o City-Block</b>
1.0	767,77	732,96	783,98
1.1		713,49	
2.0	747,78	568,53	682,79
2.1		600,85	
3.0	650,09	661,49	606,71
3.1	677,44	610,69	637,26

Tabla 8: Comparativa Tiempo Total del Primer Experimento

La Tabla 8 muestra el tiempo utilizado para realizar el experimento completo. En este caso el menor tiempo lo consigue la distancia de Mahalanobis en su versión 2.0 y 2.1. Al mismo tiempo, ambas técnicas ofrecen un menor tiempo promedio por muestra comparada (Tabla 9). Pero aumentando un 7% más de FNMR que las distancias de Manhattan y Euclídea (Tabla 5). También con un incremento variable de la tasa FMR comprendido entre 4% - 19% (Tabla 6). Y como se puede ver en (Tabla 7) se produce también un incremento del 6% de la tasa HTER con respecto a las distancias mencionadas anteriormente.

TIEMPO POR MUESTRA (min)			
	Euclídea	Mahalanobis	Manhattan o City-Block
1.0	0,46	0,44	0,47
1.1		0,42	
2.0	0,45	0,34	0,41
2.1		0,36	
3.0	0,39	0,39	0,36
3.1	0,40	0,36	0,38

Tabla 9: Comparativa de Tiempo Promedio por Identificación en el Primer Experimento

### 9.3.4 Conclusiones

A la vista de los resultados obtenidos con las diferentes formas de calcular las distancias entre usuarios, se puede concluir que los malos resultados pueden ser debidos a la presencia de datos espurios (*outliers*). También puede ser debido a una mala elección de las distancias y quizás otras distancias diferentes hubieran obtenidos una mayor generalización.

### 9.3.5 Nueva forma de elección de usuarios

La primera forma de selección de usuarios no tuvo en cuenta la posible existencia de *outliers*. Estos podrían desplazar la frontera de decisión a su antojo repercutiendo en la generalización de la misma (Figura 52). Como consecuencia, se obtendría una frontera de decisión no tan representativa del espacio de usuarios como se deseaba. Se ideó una manera más robusta frente *outliers* y evitar que la posible existencia de ellos perjudicara en la representación del espacio de usuarios. Este nuevo algoritmo, para elegir los usuarios más alejados, se bautizó como *bubble-choice*.

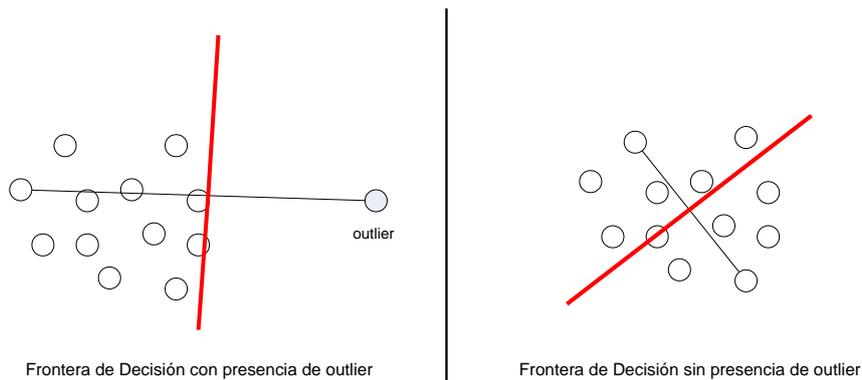


Figura 52: Frontera de Decisión Vs Outliers

### ***ALGORITMO BUBBLE-CHOICE***

En este algoritmo, la elección de los usuarios más alejados se realiza entre aquellos que pertenecen a una burbuja. De esta manera, se evita elegir aquellos que puedan representar *outliers*, centrando la frontera de decisión en el conglomerado espacial de usuarios.

El funcionamiento del algoritmo está basado en la creación de una burbuja. Imagine que quiere dibujar una circunferencia con la ayuda de un compás. Para realizar esta proeza con él, se busca un punto de apoyo donde apoyarlo y a continuación girarle consiguiendo así la circunferencia. Por lo tanto, para calcular dicha circunferencia (en este caso la burbuja) se necesita de un punto intermedio (punto de apoyo) que pertenezca al espacio de usuarios. Este punto medio se decidió que fuera una gaussiana media de todos los usuarios. Por tanto, se calcula una gaussiana promedio de cada usuario ((Ecuación 3) (Ecuación 4)) y a continuación una gaussiana media de todas las gaussianas promedio anteriores. Obteniendo de esta manera el punto intermedio y centro de la circunferencia.

Una vez conseguido el punto intermedio se calculan las distancias de todos los usuarios (recordad, de cada gaussiana promedio) a ese punto. Todas las distancias calculadas se utilizan para establecer la anchura de la burbuja que vendrá dada por el 90% de las muestras. Para ello se almacenan las distancias de los usuarios al centro de la burbuja. A continuación se ordenan en orden ascendente y se elige como umbral aquel usuario que deja a un lado el 90% de ellos

Una vez calculado la forma y anchura de la burbuja, sólo queda ver que usuarios pertenecen a la misma. Para ello se busca aquellos usuarios cuyas distancias al centro (punto intermedio) están comprendidas dentro de la anchura de la burbuja. A

continuación, de todos los usuarios pertenecientes a la burbuja se buscan aquellos dos que se encuentran más separados entre sí y repetir el proceso anterior.

### 9.3.6 Evaluación de los Resultados

La eliminación de los datos espurios mejora ligeramente el cálculo de distancias euclídeas. También mejora los resultados para Mahalanobis y Manhattan para la técnica *distance middle to gaussians 1*. Consiguiendo las tasas FNMR más bajas para la distancia Euclídea en versiones 2.0 y 3.1, y para Mahalanobis en la versión 3.1 (Tabla 10). También se obtienen las tasas FMR más bajas para las mismas distancias (Tabla 11). El ligero aumento de las tasas FNMR para las restantes técnicas de cálculo de distancias puede ser debido a la existencia de pocos (o ningún) datos espurios (outliers). Lo cual deja de manifiesto que las versiones desde la 1.0 hasta la 2.1 de la distancia Euclídea, Mahalanobis y Manhattan trabajan mejor con mayor número de usuarios. Al contrario que en las versiones 3.0 y 3.1 cuyo mayor rendimiento se obtiene acotando la base de datos.

TASA DE FALSO NO RECONOCIMIENTO (FNMR)			
	Euclídea	Mahalanobis	Manhattan o City-Block
1.0	44,33%	41,73%	49,07%
1.1		44,12%	
2.0	39,26%	43,10%	41,13%
2.1		43,70%	
3.0	42,18%	40,72%	44,66%
3.1	39,56%	38,98%	39,88%

Tabla 10: Comparativa de FNMR del Segundo Experimento

TASA DE FALSO RECONOCIMIENTO (FMR)			
	Euclídea	Mahalanobis	Manhattan o City-Block
1.0	32,98%	23,73%	51,44%
1.1		40,87%	
2.0	20,36%	26,35%	31,32%
2.1		23,66%	
3.0	29,87%	30,21%	24,15%
3.1	20,55%	20,21%	20,67%

Tabla 11: Comparativa de FMR del Segundo Experimento

HALF TOTAL ERROR RATE (HTER)			
	Euclídea	Mahalanobis	Manhattan o City-Block
1.0	38,65%	32,73%	50,25%
1.1		42,49%	
2.0	29,81%	34,72%	36,22%
2.1		33,68%	
3.0	35,77%	35,46%	34,40%
3.1	30,05%	29,60%	30,27%

Tabla 12: Comparativa de HTER del Segundo Experimento

Al igual que en los resultados anteriores la técnica basada en la distancia de Mahalanobis fue la que obtuvo un menor tiempo de ejecución para el proceso de test (Tabla 13). Seguida por la de Manhattan. Al mismo tiempo, ambas técnicas ofrecen un menor tiempo promedio por muestra identificada (Tabla 14). Pero ofrecen a su vez un 5% más de FNMR que las distancias de Mahalanobis, Manhattan y Euclídea en sus versiones 3.1 y Euclídea en su versión 2.0 (Tabla 10). En la (Tabla 12) se puede observar que las mejores tasas de HTER coinciden con las mejores tasas de FMR y FNMR conseguidas para la versión 2.0 de la distancia Euclídea y las versiones 3.1 de las distancias Euclídea, Manhattan y Mahalanobis.

TIEMPO UTILIZADO (min)			
	Euclídea	Mahalanobis	Manhattan o City-Block
1.0	637,09	790,89	975,41
1.1		554,57	
2.0	2359,62	626,04	674,58
2.1		608,14	
3.0	704,43	703,53	567,44
3.1	712,00	714,01	624,27

Tabla 13: Comparativa Tiempo Total del Segundo Experimento

TIEMPO POR MUESTRA (min)			
	Euclídea	Mahalanobis	Manhattan o City-Block
1.0	0,38	0,47	0,58
1.1		0,33	
2.0	1,41	0,37	0,40
2.1		0,36	
3.0	0,42	0,42	0,34
3.1	0,42	0,43	0,37

Tabla 14: Comparativa de Tiempo Promedio por Identificación en el Segundo Experimento

El las tablas 13 y 14 para la técnica 2.0 basada en la distancia Euclídea se obtuvieron unos elevados tiempos de cómputo. Esto fue debido al elevado uso de la CPU en el momento del experimento.

### 9.3.7 Conclusiones

Los resultados que se obtuvieron con esta nueva técnica de elección de usuarios no mejoraron los resultados anteriores de sobremanera. Dando a entender que el espacio de los datos de usuarios estaba compacto desde un principio y libre de datos espurios. Por tanto, al acotar el espacio de usuarios se puede producir una mala generalización del mismo y forzar a una clasificación global no tan discriminativa y con existencia de errores.

## 9.4 CONCLUSIONES GENERALES

Los ratios FNMR conseguidos con esta técnica se comprenden entre 38,32% - 38,98%. Estos resultados a simple vista no son nada alentadores, ya que una tasa de falsa aceptación que ronde esos valores no es óptima en la vida real. Pero si se tiene en cuenta los resultados obtenidos para la línea base (capítulo 8) la mejora propuesta sólo difiere del ratio de identificación bruta (ES) para esta misma base de datos en un 3,81%. Por el contrario se encuentra una mejoría del 16,45 % con respecto a ES-TPU y del 34,29 % con respecto a ES-GT.

Por otro lado, los ratios HTER conseguidos con esta técnica están comprendidos entre 29,10% - 30,10%. Para esta misma base de datos, nuevamente esta mejora sólo difiere en un 2,89% del ratio obtenido por ES. Por el contrario mejora en un 36,77% con respecto ES-GT y un 9,48% con respecto a ES-TPU.

El objetivo primordial del proyecto reside en reducir los tiempos de cómputo necesarios para obtener un identificación biométrica, cabe resaltar que este objetivo si ha sido conseguido. Anteriormente, con las técnicas definidas en el estado del arte (capítulo 4) para realizar un identificación se requerían unos tiempos de cómputo de 1,32 minutos para ES, 1,32 minutos para ES-TPU y 0,41 minutos para ES-GT. Con la generación del GMT se han conseguido reducir estos tiempos a valores comprendidos entre 0,33 - 0,36 minutos. Puede llevar a pensar que los valores se aproximan a la identificación ES-GT, es cierto que se aproximan pero este método genera un 34,39 % más de tasa FNMR lo que le resta credibilidad. Si se comparan los resultados obtenidos en la línea base, con los obtenidos con la mejora que ofrece el GMT se consiguió reducir un 75% los tiempos de cómputo con respecto ES y ES-TPU, y un 19,6 % con respecto ES-GT. La estimación aproximada para el cálculo del tiempo promedio con la técnica GMT es:

$$T_{GMT} = T_{Id} \cdot 0.25; T_{Id} = \text{tiempo necesario en realizar una identificación}$$

También con la metodología de preclasificación se ha disminuido el número de comparaciones necesarias para obtener una identificación. Este era el otro punto fuerte

del proyecto: disminuir la carga computacional necesaria en el proceso de identificación. Con las técnicas de identificación recogidas en el estado del arte el número estimado de comparaciones para cada una es:

$$C_{ES} = N$$

$$C_{ES-GT} = \frac{7}{2}N$$

$$C_{ES-TPU} = N$$

Con la generación del GMT el número de comparaciones viene en función del número de grupos que se forman el árbol, por tanto del número de niveles que lo forman. La fórmula que estima dicho cálculo es:

$$C_{GMT} = \log_2 G; \text{ siendo } G = \text{Número de Grupos formados}$$

Véase como ejemplo; se tiene una base de datos de  $N = 25$  usuarios, para esta base el GMT genera de media unos 1400 grupos. El número de comparaciones obtenidas son:

$$C_{ES} = 25 \text{ comparaciones}$$

$$C_{ES-GT} = 8,3 \text{ comparaciones}$$

$$C_{ES-TPU} = 25 \text{ comparaciones}$$

$$C_{GMT} = 11 \text{ comparaciones}$$

La técnica ES-GT obtiene el menor número de comparaciones pero consiguiendo una mayor tasa de falsa aceptación, como se vio anteriormente. Comparando la metodología con la técnica que menor tasa de falsa aceptación obtuvo (ES) ha reducido el número de comparaciones un 56 %. Se puede decir que la metodología propuesta reduce el número de comparaciones con respecto al estado del arte en un 50 % aproximadamente. La fórmula que estima el número de comparaciones con respecto a ES, ES-GT, ES-TPU es:

$$C_{GMT} = 0,44 \cdot C_{ES(-TPU)} + 0,25 \cdot 0,33 \cdot C_{ES-GT} \cdot Th; \text{ donde } Th = (0,1)$$

Si en el experimento se ha evaluado la identificación con umbral general el valor de la variable Th es 1, en caso contrario 0.

Para concluir, cabe resaltar que la metodología no está basada en la semántica de la tecnología, sino en la información contenida en el archivo de modelos que parametriza a un usuario. Por lo tanto, los posibles malos resultados obtenidos se deben al desarrollo de esta metodología para un ámbito general y válido para cualquier tecnología biométrica.

Se puede deducir de los resultados obtenidos, que dependiendo del tamaño del espacio de usuarios unas, distancias funcionan mejor que otras. Por ejemplo, cuando no hay acotamiento del espacio se obtienen mejores resultados mediante el cálculo de las distancias por los algoritmos de Manhattan y Euclídea, a través de la técnica *distance between middles 2*. Por el contrario, cuando se produce un acotamiento del mismo espacio funcionan en valores similares Euclídea, Manhattan y Mahalanobis, bajo la técnica *distance middle to gaussians 2*.

---

# CAPÍTULO 10

---

## CONCLUSIONES Y TRABAJOS FUTUROS

En este capítulo se presentan los objetivos conseguidos en la consecución del proyecto y los compara con los planteados inicialmente. Así mismo se presentan a su vez las posibles vías futuras de investigación.

### 10.1 CONCLUSIONES

Elaborar unas conclusiones puede ser la tarea más ardua de todas las posibles para un proyecto de investigación. Se debe enfrentar todo lo acontecido a la consecución del proyecto contra lo pretendido conseguir al comienzo del mismo. Este enfrentamiento entre optimismo y realidad debe de quedar plasmado en las siguientes líneas de este apartado.

Todo proyecto se comienza con mucho optimismo en cuanto a los requisitos que se pretenden conseguir. No siempre se consiguen todos ellos pero si se pueden aproximar bastante a los del principio. Para poder establecer el grado de éxito del PFC, se debe de conocer los requisitos conseguidos y el grado de dificultad de los mismos. Los objetivos que se cubrieron con la creación de la metodología de preclasificación de modelos biométricos han sido los siguientes:

- Se ha conseguido establecer un criterio de agrupación, o clasificación de los modelos, que se aproximara a los ratios que ofrecía la línea base.
- Se ha conseguido reducir el número de comparaciones necesarias para una identificación biométrica. Reduciendo así el coste computacional necesario por identificación.
- Se ha conseguido reducir el tiempo en realizar una identificación como consecuencia del objetivo anterior.
- Se ha reestructurado la base de datos en forma de árbol de decisión binario basados en el criterio de agrupación.
- Por último, se ha conseguido establecer una metodología de preclasificación de los modelos biométricos independiente de la semántica de la tecnología biométrica.

Como contrapartida, no se pudieron mejorar los ratios de la identificación bruta pero se consiguieron mantenerlos muy cercanos a los obtenidos por el método tradicional.

El punto clave del proyecto fue establecer la política de agrupación. La manera en que los datos serian agrupados dependientemente de la información contenida en ellos y no de la semántica de la tecnología. Aunque una parte muy importante fue el establecimiento de una metodología de trabajo, que ayudó a disminuir la carga laboral. Sobre todo en el análisis y diseño previo a la implementación del código para la creación del árbol GMT.

Una parte importante para el desarrollo ha sido la reutilización de código ya implementado. Tanto este proyecto como los experimentos desarrollados en el departamento de biometría comparten ciertas partes del código, como el de los clasificadores (GMM, SVM...), modeladores de usuarios, métodos de fusión de archivos.

Aunque desde mi punto de vista, la parte más importante que aprendí con el desarrollo de este proyecto fue inculcarme una metodología de trabajo aplicable a cualquier tipo de proyecto, trabajo, experimento... que pensándolo bien, es de lo que se trata de conseguir con la realización de un PFC.

El cumplimiento de los requisitos planteados al inicio se puede considerar satisfactorio. El desarrollo de la metodología y el criterio de agrupación que describe este PFC alivian los tiempos de cómputo y el coste computacional de cualquier tecnología biométrica con independencia de la semántica. Quizás, aplicándole la semántica de cada tecnología a la metodología establecida se podrían conseguir mejores resultados.

## 10.2 TRABAJOS FUTUROS

En esta sección se presentan algunas propuestas para posibles trabajos futuros para la metodología de preclasificación de modelos:

1. Mejorar el proceso de agrupación o generación de grupos. Pudiéndose utilizar diferentes medidas para calcular distancias y coeficientes de similitud y disimilitud.
2. Podría mejorar los resultados del GMT un ajuste mayor a la información de los usuarios que a los modelos generados con ellos.
3. Evaluar la metodología eligiendo nuevos clasificadores y modeladores de características, en vez de los propuestos en el proyecto.
4. Establecer un criterio de parada en un nivel anterior del árbol GMT.
5. Aumentar el número de usuarios que forman los nodos finales ( $n$ -arios). Para ello, se debería llevar una cuenta de los fallos producidos en los nodos con el fin de establecer el nivel que menor tasa de error ofrezca.
6. Agregar un algoritmo de *backtracking* que retome el nodo en que se produjo el fallo y enmendarlo escogiendo otro camino.
7. Realizar un estudio sobre aquellos usuarios que ofrecen mayor tasa de error con el fin de poder establecer un vínculo que mejore la agrupación.
8. Lanzar un experimento con los usuarios mejor identificados y otro con los que ofrecieron mayor tasa de error para poder establecer algún nexo en las características que mejore la discriminación.



## REFERENCIAS

- [**J. Mercer, 1909**] J. Mercer. “*Functions of Positive and Negative Type and Their Connection with the Theory of Integral Equations*”. Philosophical Transactions of the Royal Society, London, A 209:415–446, 1909.
- [**G. Fant, 1970**] G. Fant, “*Acoustic Theory of Speech Production*”, Mouton, The Hague, The Netherlands, 1970.
- [**Raphael et al, 1974**] D. E. Raphael and J.R. Young: “*Automated Personal Identification*”. SRI International. 1974.
- [**B. S. Atal , 1976**] B. S. Atal, “*Automatic Recognition of Speakers from their Voices*”, *Proceedings of the IEEE*, vol 64, 1976, pp 460 – 475.
- [**A. Dempster et al, 1977**] Dempster, A., Laird, N., and Rubin, D., “*Maximum likelihood from incomplete data via the EM algorithm*”, *J. Roy. Stat. Soc.* **39** (1977), 1–38.
- [**H. Sakoe, 1979**] H. Sakoe, “*Two-level DP Matching – A Dynamic Programming-Based Pattern Matching Algorithm for Connected Word Recognition*”; *IEEE Trans. On Acoustics, Speech and Signal Proc*, ASSP-27(6); pp.588-595; December 1.979.
- [**C.S. Myers et al, 1981**] C.S. Myers, L. R. Rabiner, “*A Level Building Dynamic Time Warping Algorithm for Connected Word Recognition*”, *IEEE Trans. on Acoustics, Speech and Signal Proc*, ASSP-29(2), pp. 284-297, April 1.981.
- [**Y. Sato et al, 1982**] Y. Sato and K. Kogure, “*Online signature verification based on shape, motion, and writing pressure*”, in *Proceedings 6th International Conference on Pattern Recognition*, 1982, pp. 823–826.
- [**A. Poritz, 1982**] A. Poritz, “*Linear predictive hidden markov models and the speech signal*” in *Proceedings IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP’82)*, vol. 7, May 1982, pp. 1291–1294.
- [**Watanabe, 1985**] S. Watanabe, *Pattern Recognition: Human and Mechanical*. New York: Wiley, 1985.
- [**L.R. Rabiner et al, 1986**] L.R. Rabiner, B.H. Huang, “*An Introduction to Hidden Markov Models*”, in *IEEE ASSP Magazine*, pp. 4-16, January 1.986.

- [**J.R. Quinlan, 1986**] J.R. Quinlan, “*Simplifying Decision Trees*”. Massachusetts Institute of Technology Artificial Intelligence Laboratory, AI Memo No. 930. December, 1986.
- [**L. Sirovich et al, 1987**] L. Sirovich and M. Kirby, “*Low-dimensional procedure for the characterization of human faces*” Journal of the Optical Society of America, vol. 4, no. 3, pp. 519–524, 1987.
- [**C.H. Lee et al, 1988**] C.-H. Lee and L.R. Rabiner; “*A Network-Based Frame-Synchronous Level Building Algorithm for connected Word Recognition*”, in Proc. ICASSP 88, vol. II, pp. 410-413, 1.988.
- [**S.J. Cox, 1988**] S.J. Cox, “*Hidden Markov Models for Automatic Speech Recognition: Theory and Application*”, in Br Telecom Technol J, vol. 6, N° 2, pp. 105-115, April 1.988.
- [**J. Oglesby et al, 1988**] J. Oglesby, J. Mason, “*Speaker recognition with a neural classifier*” in Speech 88: Proceedings of the 7th Federation of Acoustical. Societies of Europe (FASE) Symposium, Edinburgh, UK, 1988, pp. 1357–1363.
- [**K. Fukushima, 1988**] K. Fukushima, “*Neocognitron: A Hierarchical Neural Network Capable of Visual Pattern Recognition*”. Neural Networks, Vol.1, pp. 119-130, 1988.
- [**C.H. Lee et al, 1989**] C.H. Lee and L.R. Rabiner, “*A Frame-Synchronous Network Search Algorithm for Connected Word Recognition*”, IEEE Trans. on Acoustics, Speech and Signal Processing, vol. 37, n°11, pp. 1649-1658, November 1.989.
- [**R. Bajcsy et al, 1989**] R. Bajcsy and S. Kovacic, “*Multiresolution Elastic Matching*”, Computer Vision Graphics Image Processing, vol. 46, pp. 1-21, 1989.
- [**D. E. Goldberg, 1989**] D. E. Goldberg, “*Genetic Algorithm in Search, Optimization, and Machine Learning*”, Addison-wesly publishing company, pp 1-27, 1989.
- [**J. M. Naik , 1990**] J. M. Naik, “*Speaker Verification: A Tutorial*”, IEEE Communications Magazine, January 1990, pp.42-48.
- [**H. C. Lee et al, 1991**] H. C. Lee and R. E. Gaensslen, “*Advances in Fingerprint Technology*”. Eds. New York: Elsevier, 1991.
- [**J. Daugman, 1992**] J. Daugman, “*High confidence personal identification by rapid video analysis of iris texture*”, Proc. Of the IEEE, International Carnahan conf. on security technology, 1992.
- [**J.R. Deller Jr et al, 1993**] J.R. Deller Jr, J.G. Proakis, J.H.L. Hausen, “*Discrete-Time Processing of Speech Signals*”. Ed. MacMillan. 1.993.

- [**L. Rabiner et al, 1993**] L. Rabiner and B.-H. Juang, “*Fundamentals of Speech Recognition*”, Englewood Cliffs (N.J.), Prentice Hall Signal Processing Series, 1993.
- [**H.D. Chang et al, 1993**] H.D. Chang , J.F. Wang, and H.-M. Suen, “*Dynamic Handwritten Chinese Signature Verification*,” in Proceedings second IEEE International Conference on Document Analysis and Recognition, 1993, pp. 258–261.
- [**M. Magerman, 1994**] M. Magerman, “*Natural Language Parsing as Statistical Pattern Recognition*”. Doctoral dissertation. Stanford university. 1994.
- [**R. P. Wildes, 1994**] R. P. Wildes, J. C. Asmuth, G.L. Green and S.C. Hsu, “*A system for automated iris recognition* “, IEEE paper, 1994.
- [**H. Gish, 1994**] H. Gish and M. Schmidt, “*Text Independent Speaker Identification*”, IEEE Signal Processing Magazine, Vol. 11, No. 4, 1994, pp. 18-32.
- [**D. A. Reynold et al, 1995**] Douglas A. Reynold y Richard C. Rose, *Robust Text-Independent Speaker Identification Using Gaussian Mixture Speaker Models*”.
- [**W.S. Humphrey, 1995**] W.S. Humphrey: “*A discipline for software engineering*”. Ed. Addison Wesley. 1995.
- [**V. Vapnik, 1995a**] V. Vapnik, “*The Nature of Statistical Learning Theory*”. Springer-Verlag, New York, 1995.
- [**V. Vapnik, 1995b**] V. Vapnik, *Statistical Learning Theory*. John Wiley and Sons, Inc., New York.
- [**Roberto Brunelliand et al, 1995**] Roberto Brunelli y Daniele Falavigna, “*Personal Identification Using Multiple Cues*”.
- [**K. Karu et al, 1996**] K. Kart and A. K. Jain, “*Fingerprints Classification*”. Pattern Recognition 29 (3), 389-404, 1996.
- [**Nalini K. Ratha et al, 1996**] Nalini K. Ratha, Kalle Kart, Shaoyun Chen and Anil K. Jain, “*A Real-Time Matching System for Large Fingerprint Database*”. IEEE Transactions on Patter Analysis and Machine Intelligence. Vol, 18. August 1996.
- [**M. Schmidt et al, 1996**] M. Schmidt and H. Gish, “*Speaker identification via support vector classifiers*” in Proceedings IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP’96), vol. 1, May 1996, pp. 105–108.
- [**L. Devroye et al, 1996**] L. Devroye, L. Gyorf, and G. Lugosi, “*A Probabilistic Theory of Pattern Recognition*”. Berlin: Springer-Verlag, 1996.

- [**R. S. Kashi et al, 1996**] R. S. Kashi, W. T. Turin, and W. L. N. Nelson, “*On-line handwritten signature verification using stroke direction coding*,” *Optical Engineering*, vol. 35, no. 9, pp. 2526–2533, September 1996.
- [**S. Ong et al, 1996**] S. Ong, S. Sridharan, Cheng-Hong Yang, Miles Moody, “Comparison of Four Distance Measures for Long Time Text-Independent Speaker Identification”, *ISSPA*, 1996, pp. 369-372
- [**V. S. Nalwa, 1997**] Vishvjit S. Nalwa, “*Automatic On-line Signature Verification*”.
- [**J.P. Campbell, 1997**] J.P. Campbell, “*Speaker Recognition: A Tutorial*”, *Proc. of the IEEE*, vol. 85, no. 9, Sept 1997, pp. 1437-1462.
- [**Huang et al, 1997**] Kai Huang and Hong Yan: “*Off-line signature verification based on geometric feature extraction and neural network classification*”. *Pattern Recognition*, vol. 30, no. 1, pp. 9-17. 1997.
- [**M. Pandit et al, 1998**] M. Pandit and J. Kittler, “Feature selection for a DTW-based speaker verification system,” in *Proceedings IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP’98)*, vol. 2, 1998, pp. 769–772.
- [**C. Burges, 1998**] C. Burges, “*A Tutorial on Support Vector Machines for Pattern Recognition*”. DM 121-167.
- [**A. V. Oppenheim et al, 1998**] Alan V. Oppenheim and Alan S. Willsky, “*Signals and Systems, Second Edition*”. Prentice Hall, 1998.
- [**P. Indyk et al, 1998**] P. Indyk and Moteani, “*Approximate Nearest Neighbors: Towards Removing the Curse of Dimensionality*”. In *Proc. 30<sup>th</sup> Annu. ACM Symp. Theory of Computing*, 1998, pp. 604-613.
- [**Y. Wu et al, 1999**] Y. Wu, A. Ganapathiraju and J. Picone, “*Baum-Welch Re-estimation of Hidden Markov Models*”. Institute for Signal and Information Processing. University of Mississippi. June, 1999.
- [**Liu Yi et al, 1999**] Liu Yi and Pascale Fung, “*Decision Tree-Based Triphones are Robust and Practical for Mandarin Speech Recognition*”. Human Language Technology Center, Department of Electrical and Electronic Engineering, University of Science and Technology. June, 1999.
- [**Hong et al, 1999**] L. Hong, A.K. Jain and S. Pankanti: “*Can multibiometrics improve performance?*” *Proceedings of AutoID ’99*, pp. 59-64. 1999.
- [**L. R. Rabiner, 1999**] L. R. Rabiner, “*A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition*”.

- [**D. Petrovska-Delacrétaz et al, 2000**] D. Petrovska-Delacrétaz, J. Cernocky, J. Hennebert, and G. Chollet, “*Segmental approaches for automatic speaker verification*”. Digital Signal Processing, vol. 10, no. 1–3, pp. 198–212, 2000.
- [**A. Kavcic et al, 2000**] A. Kavcic and José M. F. Moura, “*The Viterbi Algorithm and Markov Noise Memory*”, IEEE Transactions on Information Theory, vol. 46, NO. 1, January 2000.
- [**J. R. Deller et al, 2000**] J. R. Deller, J. H. L. Hansen, J. G. Proakis, “*Discrete-Time Processing of Speech Signals*”, Piscataway (N.J.), IEEE Press, 2000.
- [**D. A. Reynold, 2000**] Douglas A. Reynold, Thomas F. Quatieri and Robert B. Dunn, *Speaker Verification Using Adapted Gaussian Mixture Models*.
- [**Anil K. Jain et al. 2000**] Anil K. Jain, Robert P.W. Duin and Jianchang Mao, “*Statistical Pattern Recognition: A Review*”. IEEE transaction of pattern analysis and machine intelligence, vol. 22, no. 1, January 2000.
- [**P. J. Phillips et al, 2000a**] P.J Phillips, A. Martin, C.L. Wilson and M. Przybocky: “*An introduction to evaluating biometric systems*”. Computer, vol. 33, no. 2. 2000.
- [**P. J. Phillips et al, 2000b**] P. J. Phillips, H. Moon, S. A. Rizvi and P. J. Rauss, “*The FERET Evaluation Methodology for Face Recognition Algorithms*”. IEEE Transactions on Patterns Analysis and Machine Intelligence, Vol. 22, NO. 10, October 2000.
- [**Anil K. Jain et al, 2000**] Anil K. Jain, S. Prabhakar, Lin Hong and S. Pankanti, “*Filterbank-Based Fingerprint Matching*”.
- [**F. Monrose et al, 2000**] F. Monrose and A.D. Rubin: “*Keystroke dynamics as a biometric for authentication*”. Future Generation Computer Systems, vol. 16, pp. 351-359. Elsevier. 2000.
- [**J. R. Deller et al, 2000**] J. R. Deller, J. H. L. Hansen, J. G. Proakis, *Discrete-Time Processing of Speech Signals*, Piscataway (N.J.), IEEE Press, 2000.
- [**R. C. Veltkamp et al, 2000**] R. C. Veltkamp and M. Tanase, “*Content-Based Image Retrieval System: A survey, Technical Report*”. UU-CS-200\_34, Department of Computing Science, Utrecht University.
- [**S. Furui, 2001**] S. Furui, “*Digital Speech Processing, Synthesis and Recognition*”, New York, Marcel Dekker, 2001.
- [**Baudat G. et al, 2001**] Baudat G. and Anouar F. , “*Kernel-based Methods and Function Approximation*”. MEI, Mars Electronics International. USA 2001.

[**Z. Bin et al, 2001**] Z. Bin, W. Xihong, C. Huisheng, “*On the Importance of Components of the MFCC in Speech and Speaker Recognition*”, Center for Information Science, Peking University, China, 2001.

[**J. L. Wayman, 2001**] J. L. Wayman, *Fundamentals of Biometric Authentication Technologies*.

[**T. Rhee et al, 2001**] T. Rhee, S. Cho, and J. Kim, “*On-line signature verification using modelguided segmentation and discriminative feature selection for skilled forgeries*,” in Proceedings Sixth International Conference on Document Analysis and Recognition, September 2001, pp. 645–649.

[**D. M. Magerman, 2002**] David M. Magerman, “*Statistical Decision-Tree Models for Parsing*”. Bolt Beranek and Newman Inc. April, 2002.

[**D. A. Reynolds, 2002**] Douglas. A. Reynolds, “An Overview of Automatic Speaker Recognition Technology”, *ICASSP 2002*, pp 4072-4075.

[**A. J. Mansfield et al, 2002**] A. J. Mansfield and J. L. Wayman, *Best Practices in Testing and Reporting Performance of Biometric Device*, August 2002.

[**T. Ganchev et al. 2002**] Todor Ganchev, Anastasios Tsopanoglou, Nikos Fakotakis and George Kokkinakis, *Probabilistic Neural Networks Combined with GMMs for Speaker Recognition Over Telephone Channels*.

[**C. L. Tisse et al, 2002**] Christel-loïc TISSE, Lionel MARTIN, Lionel TORRES, Michel ROBERT, “*Person identification technique using human iris recognition* “. Advanced System Technology STMicroelectronics – Université de Montpellier.

[**Guangming Lu et al, 2002**] Guangming Lu, D. Zhang and K. Wang, “*Palmprint Recognition using Eigenpalms Features*”. *Pattern Recognition Letters* 24 (2003) 1463–1467.

[**J. Haitsma et al., 2002**] J. Haitsma and T. Kalker, “*A Highly Robust Audio Fingerprinting System*”, in Proc. Int. Symp. Musical Information Retrieval, 2002, pp. 144-148.

[**A. Cuevas, 2003**] A. Cuevas: “*Gestión del proceso software*”. Centro de Estudios Ramón Areces, S.A. 2003

[**Pai-Hsuen Chen et al, 2003**] Pai-Hsuen Chen, Chih-Jen Lin, and Bernhard Schölkopf, “*A Tutorial on Support Vector Machines*”.

[**E. Bailly-Baillire et al, 2003**] E. Bailly-Baillire, Samy Bengio, Frédéric Bimbot, Miroslav Hamouz, Josef Kittler, Johnny Mariéthoz, Jiri Matas, Kieron Messer, Vlad

Popovici, Fabienne Porée, Belen Ruiz and Jean-Philippe Thiran, “*The BANCA Database and Evaluation Protocol*”.

[**A. R. Rincón et al, 2003**] Antonio Rincon Rincon and Elena Hernandez, “*Atlas Spanish Microphone Database*” collected in Spain. Applied Technologies on Language and Speech. April, 10, 2003.

[**C. L. Tisse, 2003**] C.L. Tisse: “*Contribution à la vérification biométrique de personnes par reconnaissance de l’iris*”. Ph. D, Université de Montpellier. 2003.

[**Evgeny Karpov, 2003**] Evgeny Karpov, “*Real Time Speaker Identification*”. University of Joensuu, Department of Computer Science. Master’s Thesis.

[**Hao Feng et al, 2003**] Hao Feng and Chan Choong Wah, “*Online signature verification using a new extreme points warping technique*”.

[**W. Zhao et al, 2003**] W. Zhao, R. Chellappa, J. Phillips, and A. Rosenfeld: “*Face recognition: A literature survey*”. ACM Computing Surveys, vol. 35, no. 4, pp. 399-458. 2003.

[**Danilo Simón, 2003**] Autor: Danilo Simón, Director: Javier Ortega-García, *Reconocimiento Automático Mediante Patronos Biométricos de Huella Dactilar*.

[**P. J. Phillips et al, 2003**] P.J. Phillips, P. Grother, R.J. Michels, D.M. Blackburn, E. Tabassi and J.M. Bone: “*Facial recognition vendor test 2002, evaluation report*”. 2003. Disponible en: <http://www.frvt.org/FRVT2002>

[**S. Prabhakar et al, 2003**] S. Prabhakar, S. Pankanti and Anil K. Jain. “*Biometric Recognition: Security and Privacy Concerns*”

[**J. Richiardi et al, 2003**] J. Richiardi and A. Drygajlo, “*Gaussian mixture models for on-line signature verification,*” in International Multimedia Conference, Proceedings 2003 ACM SIGMM workshop on Biometrics methods and applications, Berkeley, USA, Nov. 2003, pp. 115–122.

[**K. C. Chan et al, 2003**] K. C. Chan, Y. S. Moon and P. S. Cheng, “*Fast Fingerprint Verification using Sub-region of Fingerprint Images*”.

[**M. Padilla et al, 2003**] M. Padilla and Z. Fan, “*Digital Image Processing Project-Automatic Face Detection Using Color Based Segmentation and Template/Energy Thresholding*”, Department of Electrical Engineering, Stanford University.

[**Jeff P. Foster et al, 2003**] Jeff P. Foster, Mark S. Nixon and Adam Prügel-Bennett, “*Automatic Gait Recognition using Area-Based metrics*”. Department of Electronics and Computer Science, ISIS, University of Southampton, Southampton, SO17 1BJ, UK.

[**Liang Wang et al, 2003**] Liang Wang, Tieniu Tan, Weiming Hu and Huazhong Ning, “*Automatic Gait Recognition Based on Statistical Shape Analysis*”. IEE Transaction of Image Processing, vol. 12, NO. 9, September 2003.

[**B. Bustos et al, 2003**] B. Bustos, G. Navarro and E. Chávez, “*Pivot Selection Techniques for Proximity Searching in Metric Spaces*”. Pattern Recognition. Lett. Vol. 24, no. 14, pp. 2357-2366, 2006.

[**M. Datar et al, 2004**] M. Datar, N. Immorlica, P. Indyk and V. S. Mirrokni, “*Locality-sensitive Hashing Scheme Based on P-Stable Distribution*”. Proc. 12<sup>th</sup> Symp. Computacional Geometry, pp. 253-262, 2004.

[**X. Wu et al, 2004**] X. Wu, D. Zhang, K. Wang and B. Huang, “*Palmprint Classification Using Principal Lines*”. Pattern Recognition 37, 1987-1998, 2004.

[**Blackburn, 2004**] D. Blackburn: “*Biometrics 101*”. Federal Bureau of Investigation. 2004.

[**B. Yanikoglu et al, 2004**] B. Yanikoglu and A. Kholmatov, “*Combining multiple biometrics to protect privacy*” in Proceedings of ICPR-BCTP Workshop, 2004.

[**Elizabeth Shriberg et al, 2004**] Elizabeth Shriberg Andreas Stolcke , “*Direct Modeling of Prosody: An Overview of Applications in Automatic Speech Processing* “. Speech Technology and Research Laboratory SRI International, Menlo Park, CA, USA And International Computer Science Institute Berkeley, CA, USA.

[**Jane You et al, 2004**] Jane You, Wai-Kin Kong, D. Zhang and K. H. Cheung, “*On Hierarchical Palmprint Coding With Multiple Features for Personal Identification in Large Databases*”.

[**Shrikanth Mohan, 2004**] Shrikanth Mohan, “*Iris Recognition for Personal Identification*”, Intelligent Systems, Electrical Engineering, Clemson University

[**J. Gonzalez et al, 2004**] J. Gonzalez, J. Fires, D. Ramos and J. Ortega, “*Bayesian Analysis of Fingerprint, Face and Signatures Evidence*”.

[**Koichi Ito et al, 2004**] Koichi Ito, H. Nakajima, H. Kabayashi, T. Auki and T. Higuchi, “*A Fingerprint Matching Algorithm Using Phase-Only Correlation*”.

[**J. Coetzer et al, 2004**] J. Coetzer, B. M. Herbst and J. A. du Preez, “*Offline Signature Verification Using the Discrete Radon Transform and a Hidden Markov Model*”.

[**F. M. Rodríguez et al, 2004**] F. M. Rodríguez and F. J. Suárez López, “*Identificación Dactilar Basada en Filtros Gabor*”. Universidad de Vigo.

[**K. Delac et al, 2004**] Kresimir Delac and Mislav Grgic, “*A Survey of Biometric Recognition Methods*”.

[**P. Rosenzweyz et al, 2004**] Paul Rosenzweyz, Alane Kochems and Ari Schwartz, “*Biometric Technologies: Security, Legal and Policy Implications*”. June, 2004.

[**Anil K. Jain et al. 2004a**] Anil K. Jain, Fellow, Arun Ross and Salil Prabhakar, “*An Introduction to Biometric Recognition*”.

[**Anil K. Jain et al. 2004b**] Anil K. Jain, Sharath Pankanti, Salil Prabhakar, Lin Hong, Arun Ross and James L. Wayman, “*Biometrics: A Grand Challenge*”.

[**Frederic Bimbot et al. 2004**] Frederic Bimbot, Jean-François Bonastre, Corinne Fredouille, Guillaume Gravier, Ivan Magrin-Chagnolleau, Sylvain Meignier, Teva Merlin, Javier Ortega-García, Dijana Petrovska-Delacrétaz and Douglas A. Reynolds, A “*Tutorial on Text-Independent Speaker Verification*”.

[**Joaquín Gonzalez-Rodriguez et al. 2004**] Joaquin Gonzalez-Rodriguez, Julian Fierrez-Aguilar, Daniel Ramos-Castro and Javier Ortega-García, “*Bayesian Analysis of Fingerprint, Face and signature evidences with Automatic Biometric Systems*”.

[**Unathi Mahola et al. 2004**] Unathi Mahola, Fulufhelo V. Nelwamondo and Tshilidzi Marwala, “*HMM Speaker Identification Using Linear and Non-Linear Merging Techniques*”.

[**Morné Breedts, 2004**] Morné Breedts and Martin S. Olivier, “*Using a Central Data Repository for Biometric Authentication in Passport Systems*”.

[**N. Yager et al, 2004**] N. Yager and A. Amin: “*Fingerprint verification based on minutiae features: a review*”. Pattern Analysis and Application, vol. 17, pp. 94-113. 2004.

[**RD1777, 2004**] Ministerio de Economía y Hacienda: “*Real Decreto 1777/2004, de 30 de Julio, por el que se aprueba el Reglamento del impuesto de Sociedades. Anexo: Tablas de Coeficientes de Amortización*”. 2004.

Disponible en: [http://www.boe.es/g/es/bases\\_datos/doc.php](http://www.boe.es/g/es/bases_datos/doc.php)

[**J. Fu et al, 2005**] J. Fu, H. J. Caulfield, S-M Yoo and V. Atluri, “*Use of Artificial Color Filtering to Improve Iris Recognition and Searching*”. Pattern Recognition Letters 26, 2244-2251, 2005.

[**L. Yu et al, 2005**] L. Yu, D. Zhang, K. Wang and W. Yang, “*Coarse Iris Classification Using Box-Counting to Stimulate Fractal Dimensions*”. Pattern Recognition 38, 1791-1798, 2005.

[**Anil K. Jain et al, 2005**] A.K. Jain, K. Nandakumar and A. Ross: “*Score normalization in multimodal biometric systems*”. Pattern Recognition, vol. 38, no. 12, pp. 2270-2285. 2005.

[**Alen Peacock et al, 2005**] Alen Peacock, Xian Ke and Mathew Wilkerson, “*Typing Patterns: A Key to user Identification*”. Massachusetts Institute of Technology.

[**Rui Xu et al, 2005**] Rui Xu and Donald Wunsch, “*Survey of Clustering Algorithms*”. IEEE Transactions on Neural Networks, Vol. 16, NO. 3, May 2005.

[**José García-Hernández et al, 2005**] José García-Hernandez y Roberto Paredes, “*Biometric Identification Using Palmprint Local Features*”.

[**M. Wölfel et al, 2005**] Matthias Wölfel and Hazim Kemal Ekenel, “*Feature Weighted Mahalanobis Distance: Improved Robustness for Gaussian Classifiers*”. Institut fuer Thoretische Informatik, Universität Karlsruhe, May 2005.

[**D. Dessimoz et al, 2005**] D. Dessimoz, J. Richiardi, C. Champod and A. Drygajlo: “*Multimodal biometrics for identity documents*”. University of Lausanne and EPFL (European Biometrics Portal). Technical Report PFS 341-08.05. 2005.

[**Koichi Ito et al, 2005**] Koichi Ito, Ayumi Morita, T. Aoiki, T. Higuchi, H. Nakajima and K. Kobayashi, “*A Fingerprint Algorithm Using Phase-Based Image Matching for Low-Quality Fingerprints*”.

[**T. Marwala et al, 2005**] T. Marwala, S. Chakraverty, U. Mahola, “*Neural Networks and Support Vector Machines For Fault Identification In Cylinders*”, Proceedings of International Symposium on Neural Networks and Soft Computing in Structural Engineering, Cracow, Poland, 2005.

[**Arslan Brömme, 2005**] Arslan Brömme, “*A Risk Analysis Approach for Biometric Authentication Technology*”. International Journal of Network Security, Vol.2, No. 1, PP.52-63, January 2006.

[**O. Coltell et al, 2006**] O. Costell, J. M. Badia and G. Torres, “*Biometric Identification System based in Keyboard System*”.

[**A. Zilinskas et al, 2006**] A. Zilinskas and J. Zilinskas, “*On Multidimensional Scaling with Euclidean and City Block Metrics*”. Institute of Mathematics and Informatics, Akademijos, Vilnius, Lithuania. ISSN 1392-8619 print/ISSN 1822-3613 online. 2006, Vol XII, No 1, 69-75. <http://www.tede.vgtu.it>

[**J. Fierrez-Aguilar et al, 2006**] J. Fierrez-Aguilar, J. Ortega-Garcia, D. Torre-Toledano, J. Gonzalez-Rodriguez, “*BioSec Baseline Corpus: A Multimodal Biometric Database*”. 12 January 2006.

[**C. E. Vivaracho et al, 2006**] C. E. Vivaracho, V. Cardeñoso y C. Gonzalez, “*Mejoras en un Sistema de Reconociendo de Locutor Basado en RNA Mediante Entrenamiento con Normalización de Canal*”.

- [**R. L. van Renesse, 2006**] Rudolf L. van Renesse, “*Implication of Applying biometrics to Travel-Document*”. TNO Institute of Applied Physics.
- [**R. Capelli et al, 2006**] R. Cappelli, D. Maio, D. Maltoni, J.L. Wayman and A.K. Jain.: “*Performance evaluation of fingerprint verification systems*”. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 28, no.1, pp. 3-18. 2006.
- [**Julián Fierrez, 2006**] Autor: Julián Fierrez Aguilar, Director: Javier Ortega-García, “*Esquemas Adaptados de Fusión para Autenticación Biométrica Multimodal*”.
- [**D. Dessimoz et al, 2006**] D. Dessimoz, J. Richiardi, C. Champod and Drygajlo, “*Multimodal Biometrics for Identity Documents*”. State-of-the-Art, Research Report, Version 2.0, June 2006.
- [**P. Zezula et al, 2006**] P. Zezula, G. Amato, V. Dohnal and M. Batko, “*Similarity Search: The Metric Space Approach*”. Berlin, Germany: Springer, 2006.
- [**Xianchao Qin et al, 2007**] Xianchao Qin, Zhenan Sun and Tieniu Tan, “*Coarse Iris Classification by Learned Visual Dictionary*”. ICB 2007, LNCS 4642, pp. 770-779, 2007.
- [**R. Mukherjee, 2007**] R. Mukherjee, “*Indexing Techniques for Fingerprint and Iris Database*”. M. S. dissertation, Lane Dept. Comput. Sco. Elect. Eng., West Virginia University, Morgantown, WV, 2007.
- [**R. S. Reillo et al, 2007**] R. S. Reillo, C. S. Ávila y M<sup>a</sup>B. F. Saavedra, “*La Evaluacion de la Seguridad en Sistemas Biométricos*”.
- [**M. Filippone et al, 2007**] M. Filippone, F. Camastra, F. Masulli and S. Rovetta, “*A Survey of Kernel and Spectral Methods for Clustering*”. Department of Computer and Information Science, University of Genova and CNISM. Department of Applied Science, University of Naples Parthenope, Italy 2007.
- [**J. Fierrez et al, 2007**] Julián Fierrez, Javier Ortega-Garcia, Daniel Ramos, Joaquín González-Rodríguez, “*HMM-based on-line signature verification: Feature extraction and signature modeling*”.
- [**Amjad Zaim, 2007**] Amjad Zaim, “*Automatic Segmentation of Iris Images for the Purpose of Identification*”. InfoMD Consultancy Group Toledo, Ohio USA [www.infomd.org](http://www.infomd.org)
- [**L. Puente et al, 2008**] L. Puente Rodríguez, A. García Crespo, M. J. Poza Lara and B. Ruiz Mezcuca: “*Study of different fusion techniques for multimodal biometric authentication*”. Proceedings of the 2008 IEEE International Conference on Wireless & Mobile Computing, Networking & Communication. pp. 666-671. 2008.

[**Felipe Ochoa et al. 2008**] Felipe Ochoa, César San Martín, Roberto Carrillo, “*Identificación Biométrica de Locutores para el Ámbito Forense: Estado del Arte*”, VI Congreso Iberoamericano de Acústica – FIA 2008.

[**S. M. Nataranjan et al, 2008**] S. M. Nataranjan and Thangarajan R., “*Structural Parking of Natural Language Text in Tamil Using Phrase Structure Irbid Language Model*”. International Journal of Computer, Information and Systems Science, and Engineering, 2008.

[**C. Conde et al, 2008**] Cristina Conde and Enrique Cabello, “*Biometric Facial Verification using 2D and 3D Information*”.

[**D. T. Toledano et al, 2008**] D. T. Toledano, D. Hernández-López, C. Esteve-Elizalde, J. Fierrez, J. Ortega-Garcia, D. Ramos and J. Gonzalez-Rodriguez. “*BioSec Multimodal Biometric Database in Text-Dependent Speaker Recognition*”. ATVS Biometric Recognition Group, Universidad Autónoma de Madrid.

[**Feng Hao et al, 2008**] Feng Hao, J. Daugman and Piotr Zielinski, “*A Fast Search Algorithm for a Large Fuzzy Database*”. IEE Transaction Forensics and Security, VOL. 3, NO. 2, June 2008.

[**Karthik Nandakumar et al. 2009**] Karthik Nandakumar, Anil K. Jain and Arun Ross, “*Fusion in Multibiometric Identification Systems: What about the missing data?*”

[**M. Martinez-Diaz et al, 2009**] M. Martinez-Diaz, Julian Fierrez and Javier Ortega-Garcia, “*Dynamic Signature Verification Using Portable Devices*”.

[**J. Galbally et al, 2009**] J. Galbally, J. Fierrez, J. Ortega-Garcia, C. McCool and S. Marcel, “*Hill-Climbing Attack to an Eigenface-Based Face Verification System*”.

[**K. Niinuma et al, 2009**] K. Niinuma and Anil K. Jain, “*Continuous User Authentication Using Temporal Information*”. Department of Computer Science & Engineering, Michigan State University, East Lansing, MI.

[**Unsang Park et al, 2009**] Unsang Park, Arun Ross and Anil K. Jain, “*Periocular Biometrics in the Visible Spectrum: A Feasibility Study*”. To appear in Biometrics: Theory, Applications and Systems (BTAS 09), Washington DC, September 2009.

[**Hasan Demirel et al, 2009**] Hasan Demirel and Gholamreza Anbarjafari, “*Iris Recognition System Using Combined Histogram Statistics*”. Department of Electrical and Electronic Engineering, Eastern Mediterranean University, Gazimağusa, KKTC, Mersin 10, TURKEY.

[**S. Young et al, 2009**] S. Young, G. Evermann, M. Gales, T. Hain, D. Kershaw, X. Liu, G. Moore, J. Odell, D. Ollason, D. Povey, V. Valtchev and P. Woodland, “*The HTK*

*Book*". 1995-1999 Microsoft Corporation, 2001-2009 Cambridge University Engineering Department. March, 2009.

**[D. Carrero et al., 2009]** D. Carrero and B. Ruiz, "*Front-End for a Biometric Extended Experiment Platform (BEEP)*". Proyecto Fin de Carrera, Julio 2009.

**[1]** Diehl, Frank (Temic), Fischer, Volker (IBM), Kiessling, Andreas (Ericsson) Marasek, Krzysztof (Sony) "*Specification of Databases - Specification of recording scenarios*" Deliverables of the SpeeCon Project D211, 212, 213, 214, 215. <http://www.speecon.com>



## **ANEXO I:**

# **GESTION DEL PROYECTO**



En las siguientes páginas se detalla la planificación y el presupuesto para el proyecto. El capítulo presenta la estimación de los recursos económicos y humanos necesarios y a continuación muestra el plan de utilización de los mismos a lo largo de las diferentes etapas del proyecto.

## 11.1 INTRODUCCIÓN

La planificación surge de la necesidad de estimar los costes económicos y temporales que requiere el desarrollo de un proyecto así como del control de los factores que pueden alterar su evolución [W.S. Humphrey, 1995]. En base al conjunto de funcionalidades propuestas, se hace imprescindible definir un calendario de trabajo con el fin de ordenar la ejecución de cada una de las tareas necesarias para su implementación.

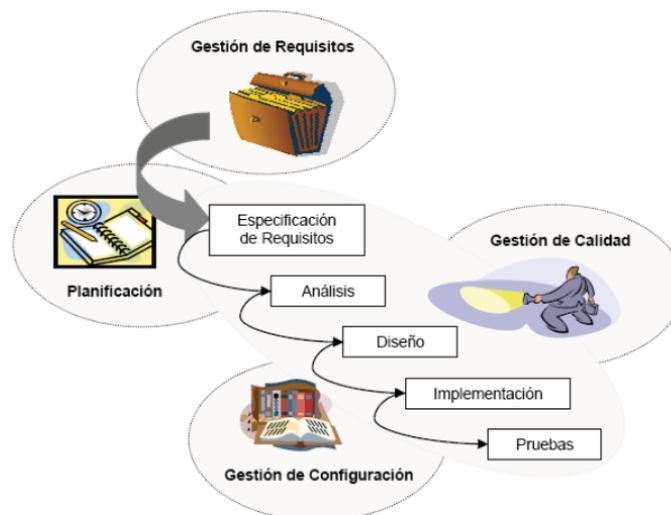


Figura 53: Proceso de Desarrollo Software

El proceso software es la descripción de las etapas que se siguen durante la ejecución de un proyecto. Su correcta definición permite asegurar una adecuada asignación de recursos y conocer el estado del proyecto en cada momento. Establece el marco de trabajo, tanto técnico como de gestión, en la aplicación de los métodos, las herramientas y las personas a las tareas de desarrollo de software [A. Cuevas, 2003].

Una metodología de trabajo definida permite que cada nuevo proyecto sea construido en base a la propia experiencia y a la de los predecesores. Su utilización permite identificar las causas de los problemas para corregirlos. En la Figura 53 se muestra el proceso donde se unen las actividades de desarrollo y gestión. Estas últimas se orientan a controlar el desarrollo del proceso software y a corregir las desviaciones

con respecto a los parámetros de calidad establecidos. Los procesos de gestión se ejecutan de manera paralela a las operaciones de desarrollo.

El ciclo de vida de un proyecto es el periodo que comienza cuando se concibe un producto software y termina cuando dicho producto deja de estar disponible. Se divide normalmente en fases que estructuran y organizan las etapas de concepción, desarrollo y mantenimiento del sistema software.

## 11.2 PLANIFICACIÓN DEL PROYECTO

La consecución de todo proyecto debe seguir una estricta planificación con el fin de poder ser terminado en los plazos establecidos. En este apartado se describe la planificación que se ha tenido en cuenta en la consecución del proyecto.

La planificación aquí descrita está formada por dos fases, cada fase contiene una serie de etapas y cada una de estas etapas contendrá sus respectivas tareas:

- FASE A: Trabajo realizado
  - ETAPA 1: Documentación y estudio del estado del arte.
- FASE B: Trabajo de campo
  - ETAPA 1: Establecimiento de los criterios de agrupación.
  - ETAPA 2: Creación y adaptación del software a los criterios.
  - ETAPA 3: Evaluación de los criterios de agrupación.
  - ETAPA 4: Optimización de los resultados y tiempos de cómputo.
  - ETAPA 5: Redacción de la memoria del proyecto.

## 11.3 ESTIMACIÓN DE LOS RECURSOS TEMPORALES

La estimación de los recursos temporales hace referencia al tiempo necesario para la consecución del proyecto. El proyecto tiene como fecha de comienzo el 3 de Agosto del 2009 y la fecha de finalización con la actual planificación se emplaza hasta el 20 de Junio de 2010. Entre estas dos fechas se sitúa el desarrollo actual del proyecto con las sucesivas tareas. El reparto de horas para cada una de las fases y etapas se muestra en la Tabla 15.

PLANIFICACION		HORAS
FASE A	Trabajo realizado	
	ETAPA 1	Documentación y estado del arte
FASE B	Trabajo de campo	
	ETAPA 1	Establecimiento de criterios de agrupación
	ETAPA 2	Creación y adaptación del software a los criterios
	ETAPA 3	Evaluación de los criterios de agrupación
	ETAPA 4	Optimización de los resultados y tiempos de computo
	ETAPA 5	Redacción de la memoria del proyecto
TOTALES		1.074

Tabla 15: Recursos Temporales por Fases del Proyecto

El diagrama de Gantt de la Figura 54 muestra la sucesión temporal de cada una de las tareas que componen el proyecto

## 11.4 ESTIMACIÓN DE LOS RECURSOS ECONÓMICOS

En este apartado se detalla el coste económico que conlleva la realización del Proyecto. Estos costes son divididos en dos bloques: costes materiales, el constituido por los componentes empleados en la realización del proyecto y costes de personal, donde se detallan los honorarios de las personas que participaron, así como el tiempo invertido por cada una de ellas. El último bloque incluye el coste final, la suma de todos los costes.

### 11.4.1 Recursos Materiales

La estimación del coste y de los periodos de amortización de los recursos materiales se han obtenido en base a lo descrito en [RD1777, 2004]. En este documento se establece un plazo de amortización de 8 años para los elementos de tratamiento de información y de 6 años para los programas informáticos. En base a esta información, el coste asociado a cada elemento viene determinado por la siguiente expresión.

$$Coste = Precio \cdot Coef = Precio \cdot \left( Unidades \cdot \frac{d}{a} \right)$$

En la expresión anterior,  $a$  y  $d$  corresponden al periodo de amortización del elemento y a la duración del proyecto en años, respectivamente. La Tabla 16 muestra el desglose de los costes asociados a los recursos materiales.

Concepto	Precio (€)	Coefficiente (%)	Precio (€)
PC Sobremesa Phoenix	1.290,31	9,4	121,28
Window XP Profesional OEM	124,43	12,5	15,55
Impresora Canon Pixma multifunction MX300	68,22	9,4	6,41
Microsoft Office 2007 Home&Student 32 bits retail	98,99	12,5	12,37
Sistema usuario Visual Studio 2007	60	12,5	7,5
Monitor 18.5'' Hacer X193HQB TFT Panorámico	89	9,4	8,36
Conexión ADSL a 6 Mbps durante 10 meses a 39,90 €/mes	399	100	399
Fungibles	100	100	100
TOTAL			670,47

Tabla 16: Recursos Materiales para el Desarrollo del Proyecto

### 11.4.2 Recursos Humanos

Los recursos temporales presentados en el apartado 10.3 son realizados por personal dando lugar a los costes de recursos humanos. El PFC será dirigido por el tutor que hará las veces de Director de Proyecto, y será realizado por el proyectando que hará las veces de Jefe de Proyecto, Analista-Programador. En la Tabla 17 se muestran los perfiles de los profesionales requeridos para la ejecución de las diferentes tareas y los costes de cada uno de ellos.

Ocupación	Horas Invertidas	Precio (€/hora)	Importe (€)
Director de Proyecto	44,5	90	4.005
Jefe de Proyecto	74,5	65	4.842,5
Analista-Programador	956	35	33.460
TOTAL			42.307,5

Tabla 17: Recursos Humanos

Los perfiles profesionales requeridos para desarrollar cada una de las tareas son los siguientes:

- Director de Proyecto: Profesional titulado en Ingeniería Informática con experiencia. Requiere habilidades como liderazgo, capacidad de comunicación e intermediación en conflictos. Los honorarios de este profesional se sitúan en 90 €/hora.
- Jefe de Proyecto: Profesional titulado en Ingeniería Informática con amplia formación en gestión de proyectos y recursos humanos. Los honorarios de este profesional se encuentran en 65 €/hora.

- Analista de Servicios y Programador: Profesional titulado en Ingeniería de Informática con conocimientos de arquitectura de redes, programación en diferentes lenguajes, protocolos de comunicaciones y seguridad en entornos web. Así mismo, ha de poseer conocimientos en biometría. Los honorarios de este profesional se encuentra en 35 €/hora.

### 11.4.3 Costes Totales

El coste total al que asciende el proyecto es de 59.825,33€ (cincuenta y nueve mil ochocientas veinticinco con treinta y tres céntimos). La Tabla 18 muestra el desglose del coste total obtenido.

Concepto	Valor (€)
Recursos Materiales	670,47
Recursos Humanos	42.307,5
<b>Subtotal</b>	42.977,97
Gastos Generales (20% Subtotal)	8.595,59
<b>Base Imponible</b>	51.573,56
I.V.A. (16% Base Imponible)	8.251,77
<b>TOTAL</b>	59.825,33

Tabla 18: Costes Totales del Proyecto

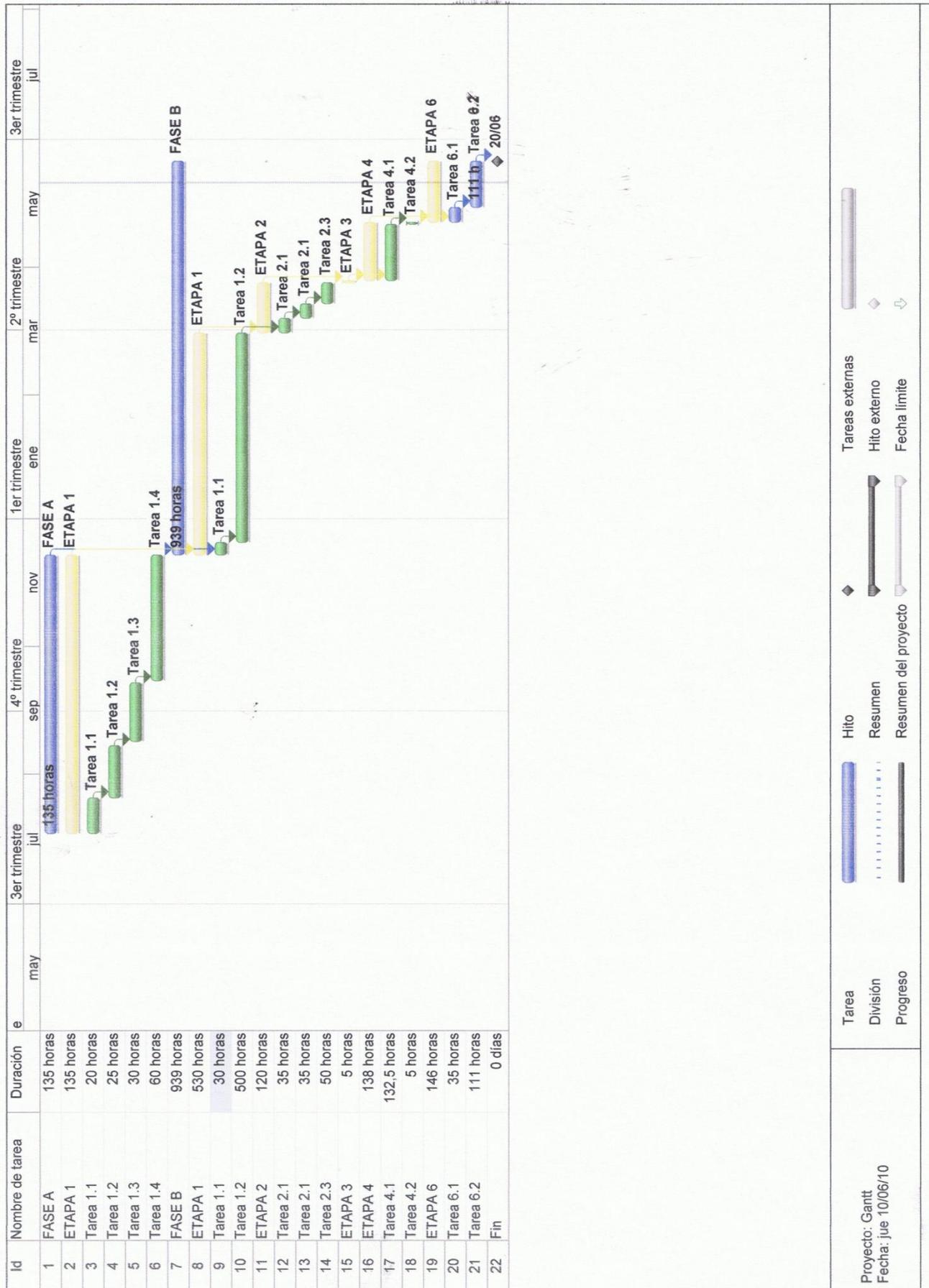


Figura 54: Diagrama Gantt

## **ANEXO II:**

# **HERRAMIENTAS UTILIZADAS**



A continuación se presentan las herramientas utilizadas tanto para la implementación de la aplicación como para la gestión del proyecto y la redacción de la memoria.

### **Microsoft Visual Studio 2005**



Microsoft Visual Studio<sup>12</sup> es un entorno de desarrollo integrado para sistemas basados en Windows. Soporta lenguajes de programación como Visual C++, Visual C#, Visual J#, ASP.NET y Visual Basic .NET y actualmente se están desarrollando extensiones para muchos otros lenguajes. Visual Studio permite crear aplicaciones, sitios y aplicaciones web, así como servicios web en cualquier entorno que soporte la plataforma. Visual Studio se ha utilizado para codificar el código para la generación del GMT.

### **Microsoft Office 2007**



Microsoft Office 2007<sup>13</sup> es la versión más reciente del paquete ofimático de Microsoft. Cuenta con herramientas para editar textos, realizar hojas de cálculo, presentaciones de diapositivas y otras muchas aplicaciones. La elaboración de la presente memoria se ha realizado mediante MS Word, para los diagramas UML se ha utilizado MS Visio, la planificación del proyecto se ha realizado con MS Project, la gestión de la base de datos se ha realizado mediante MS Access y los costes del proyecto se han calculado mediante MS Excel.

### **Mozilla Firefox 3**



Mozilla Firefox<sup>14</sup> es un navegador de Internet libre y de código abierto desarrollado por la Corporación Mozilla, la Fundación Mozilla y un gran número de voluntarios externos. Firefox es un navegador multiplataforma y está disponible en versiones para Microsoft Windows, Mac OS X, GNU/Linux y algunos sistemas basados en Unix. Firefox se ha empleado para la búsqueda de información.

### **Microsoft Windows Server System**



Es una marca que abarca una línea de productos servidor de Microsoft Corporation<sup>15</sup>. Windows Server System es un sistema operativo para la gestión de servidores que ha sido utilizado para la ejecución de todos los experimentos.

<sup>12</sup> <http://www.microsoft.com/spanish/msdn/vs2005/default.msp>

<sup>13</sup> <http://office.microsoft.com/es-es/default.aspx>

<sup>14</sup> <http://www.mozilla-europe.org/es/firefox/>

<sup>15</sup> <http://www.microsoft.com/servers/home.msp>