



UNIVERSIDAD DE LAS PALMAS
DE GRAN CANARIA

Departamento de Informática y Sistemas

TESIS DOCTORAL

Análisis Biométrico de las Orejas

M^a Esther González Sánchez

Las Palmas de Gran Canaria

2008

Don Santiago Candela Solá, DIRECTOR DEL DEPARTAMENTO DE INFORMÁTICA Y SISTEMAS DE LA UNIVERSIDAD DE LAS PALMAS DE GRAN CANARIA,

CERTIFICA,

Que el Consejo del Departamento, en su sesión ordinaria de fecha 1 de octubre de 2008, tomó el acuerdo de dar el consentimiento para su tramitación a la tesis doctoral titulada "**Análisis Biométrico de las Orejas**", presentada por la doctoranda Doña M^a Esther González Sánchez y dirigida por los Doctores Don Luis Álvarez León y Don Luis Mazorra Manrique de Lara.

Y para que así conste, y a efectos de lo previsto en el Artº 73.2 del Reglamento de Estudios de Doctorado de esta Universidad, firmo la presente en Las Palmas de Gran Canaria, a uno de octubre de dos mil ocho.





UNIVERSIDAD DE LAS PALMAS
DE GRAN CANARIA

Departamento: **Informática y Sistemas**

Programa de Doctorado: **Percepción Artificial y Aplicaciones**

Análisis Biométrico de las Orejas

Tesis Doctoral presentada por **D^a M^a Esther González Sánchez**

Dirigida por el Dr. **D. Luis Álvarez León**

Codirigida por el Dr. **D. Luis Mazorra Manrique de Lara**

El director

Luis Álvarez León

El Codirector

**Luis Mazorra Manrique de
Lara**

La doctoranda

M^a Esther González Sánchez

Las Palmas de Gran Canaria, a 8 de Octubre de 2008

A mi madre, in memoriam.

*A Sergio y Elena, mis hijos,
que son mi razón de ser
y el motor de mi vida*

Agradecimientos

Hay varias personas que, de una u otra forma, han contribuido a que este trabajo sea una realidad y quisiera expresarles aquí mi agradecimiento.

Mi agradecimiento, en primer lugar, a Luis Álvarez, sin cuya inestimable ayuda no hubiera sido posible esta tesis, por la paciencia y comprensión que ha demostrado tener conmigo y mis circunstancias personales. Y, como no, a Luis Mazorra, por su colaboración como codirector de este trabajo. También a todos aquellos que, desinteresadamente, se prestaron a ser fotografiados para la realización de la base de datos de imágenes de las orejas.

Al rector de la U.L.P.G.C. y a Santiago Candela (Director del Departamento de Informática y Sistemas), por desarrollar políticas de apoyo a los TEUS para la realización de sus tesis doctorales, de las cuales me he podido beneficiar.

A Agustín Trujillo, Javier Sánchez y Miguel Alemán, por sus continuas palabras de aliento para sacar adelante este trabajo, por impartirme un cursillo acelerado de las herramientas de software utilizadas en el desarrollo y redacción de este trabajo, y por prestarse siempre a resolver, amablemente, todas mis dudas.

A Beatriz Correas, por su predisposición a colaborar conmigo en lo que fuera necesario y guiarme en los trámites a seguir para la presentación de esta tesis, cuando ella acababa de presentar la suya. También por haber confiado en mí para formar parte de su equipo directivo, facilitándome así un posible ascenso en la categoría laboral.

A Marcos, por estar siempre a mi lado, animándome y apoyándome, y por compartir conmigo esas dos personitas maravillosas que son la alegría de la casa. Y a mi padre, por su apoyo incondicional a todo lo que hago.

A todos, muchas gracias.

Índice general

1. Descripción general del trabajo realizado	5
1.1. Introducción	5
1.2. Aportaciones Principales	6
1.3. División del documento	10
2. Introducción	15
2.1. Biometría	15
2.2. Sistemas Biométricos. Aplicaciones	19
2.3. Biometría de la oreja	24
3. Estado del arte	29
3.1. Redes Neuronales	31
3.2. Diagramas de Voronoi	32
3.3. Transformación de campos de fuerza	34

3.4. Análisis de los componentes principales (PCA)	36
3.5. Algoritmos genéticos	38
3.6. Parches de Superficie	39
3.7. Extracción de características geométricas	40
3.8. Técnica de descomposición de ondículas Haar	43
4. Localización del contorno	45
4.1. Contornos Activos	45
4.1.1. Planteamiento clásico: Minimización de energía	46
4.1.2. Modelos Deformables Implícitos	48
4.2. Aproximación utilizada en este trabajo	53
4.2.1. Atracción del snake hacia los bordes	57
4.2.2. Regularización del contorno	61
4.2.3. Detalles de la Implementación	65
5. Modelo de Ovoide	69
5.0.4. Detalles de implementación	76
6. Comparación de dos orejas	79
7. Resultados Experimentales	87
7.1. Análisis Estadístico	87
7.1.1. Introducción	87

7.1.2.	Caracterización y comparación entre individuos	91
7.1.3.	Método de selección de parámetros	92
7.1.4.	Resultados experimentales del método de selección de parámetros propuesto	95
7.1.5.	Resultados experimentales de la clasificación de vistas	102
7.1.6.	Análisis de los datos normalizando la excentricidad de las orejas	107
7.2.	Análisis Discriminante Clásico	116
7.2.1.	Método de selección de variables	117
7.2.2.	Obtención de las funciones discriminantes	121
7.2.3.	Interpretación de las funciones discriminantes	127
7.2.4.	Evaluación del procedimiento de clasificación	131
7.2.5.	Resultados experimentales de la clasificación de vistas	134
7.2.6.	Análisis de los datos normalizando la excentricidad de las orejas	140
7.3.	Comparación de resultados	149
7.3.1.	Excentricidad Variable:	150
7.3.2.	Excentricidad Normalizada:	153
8.	Conclusiones	157
9.	Futuras Líneas de Desarrollo	161
9.0.3.	Automatización de la aproximación inicial:	161
9.0.4.	Optimización de la localización del centroide:	162

9.0.5. Implementación de Métodos de Reconocimiento y no sólo de
Identificación: 162

Descripción general del trabajo realizado

1.1. Introducción

Actualmente existe una creciente necesidad de establecer medidas de seguridad y control de acceso tanto en lugares públicos como privados (edificios oficiales, aeropuertos, estaciones, entidades bancarias, etc...) y, en general, de verificar la identidad de los individuos. Básicamente existen tres métodos diferentes de verificación de la identidad:

- algo que el usuario posee, como tarjetas de identificación, distintivos, llaves;
- algo que el usuario conoce, como claves de acceso, identificación de usuario, numeros de identificación personal (PIN);
- algo que el usuario es en sí mismo (biometría), como huellas dactilares, su cara, sus ojos, sus orejas.

Reconocimiento biométrico, o simplemente biometría, se refiere al reconocimiento automático de los individuos basándose en sus características fisiológicas o de comportamiento. La biometría ofrece una precisión mucho mayor que los métodos tradicionales de reconocimiento.

La utilización de la oreja como característica biométrica del individuo posee ciertas ventajas como son: (1) las orejas son una parte del cuerpo que está generalmente visible y (2) se puede tomar imágenes de ella fácilmente, incluso sin el conocimiento ni consentimiento de la persona examinada. La biometría de la oreja también se puede utilizar para acentuar la efectividad de otras biometrías ampliamente utilizadas (como voz, geometría de la mano o identificación de rostros), en la implementación de sistemas biométricos multimodales (híbridos), que están recibiendo mucha atención últimamente.

En cualquier caso, se hace necesario un método robusto de extracción de características, a partir de las imágenes tomadas de la oreja, que se pueda usar para determinar la identidad de algunos individuos. En este trabajo proponemos un método de identificación de individuos, basado en la extracción de ciertas características geométricas de las imágenes de sus orejas.

1.2. Aportaciones Principales

La aproximación que nosotros proponemos considera las características geométricas de los contornos de las orejas. El objetivo del trabajo no es diseñar un método completamente automático de identificación, sino analizar el poder discriminante de la geometría de los contornos externo e interno de las orejas, como elemento identificativo del individuo. Por ello, para alcanzar la máxima precisión en la obtención de los contornos de las orejas, se ha diseñado un método semiautomático donde, a partir de una primera aproximación grosera del contorno de la oreja, facilitada manualmente, se defina, posteriormente, dicho contorno de forma automática, aproximando finalmente el contorno de la oreja con gran precisión. En este sentido, hemos implementado una nueva forma de ajustar el contorno de una oreja en una imagen, en la que hemos aportado, básicamente, las siguientes contribuciones:

- La adaptación de un probado modelo de aproximación a curvas cerradas, como son los contornos activos (snakes), con numerosas e importantes aplicaciones a distintos tipos de imágenes, para que se ajuste a los requerimientos del contorno de la oreja, que es un contorno abierto. Esto lo hemos logrado insertando, en el polígono inicial, un último segmento, que une el primer y último puntos marcados, y que se mantendrá fijo a lo largo de las iteraciones del algoritmo.

- La aproximación del contorno externo de la oreja mediante un ovoide. Como primera aproximación a la forma del contorno externo de una oreja, utilizamos un modelo de ovoide. Explotamos aquí el hecho de que, en general, dicho contorno externo tiene, globalmente, una forma ovoidal. El modelo de ovoide que utilizamos en este trabajo es el de una elipse deformada, donde incluimos un parámetro de distorsión de la elipse (γ), que determinará el carácter ovoidal de la forma. Se obtienen, en una aproximación inicial, los parámetros de un ovoide que se ajuste al contorno de la oreja. Posteriormente, estos parámetros se refinan, utilizando un criterio de distancia euclídea. La ventaja de este modelo de ovoide, aparte de ser bastante simple (solo un parámetro más que la elipse) y ajustarse bastante bien al contorno externo de la oreja, es que, algunos de sus parámetros, se pueden usar para comparar diferentes orejas y alinear, de una forma muy precisa, sus contornos. Por ejemplo, los parámetros de distorsión y excentricidad ($\gamma, \frac{b}{a}$), son invariantes euclídeos que se pueden usar para comparar orejas independientemente de su tamaño y localización. Además de esto, también nos influyó en la elección, la existencia de numerosos trabajos donde se consigue clasificar correctamente, huevos pertenecientes a diferentes especies de pájaros, basándose en determinadas medidas geométricas extraídas de esta forma ovoide [35], [36], [37], [38].

- Normalización de los contornos obtenidos de diferentes imágenes de orejas, para ponerlas en correspondencia y poder así compararlas. Para ello nos basamos en los parámetros del ovoide que hemos asociado al contorno y aplicamos una trans-

formación para llevar el centro de la elipse deformada al $(0,0)$ y colocarla en vertical, normalizando, además, su área. El hecho de localizar el centro del nuevo sistema de coordenadas en el origen, hace que nuestro método sea invariante a rotaciones, traslaciones y escalados. Una vez hecho esto, existen dos opciones: comprimir/expandir las elipses deformadas obtenidas hasta alcanzar una excentricidad fija dada, para compararlas, o bien, dejar la excentricidad sin fijar y que puedan tomar, por tanto, valores variables. En el estudio hemos considerado las dos opciones.

- Creación de una base de datos de imágenes de prueba. Para validar los resultados de nuestro trabajo hemos creado una base de datos compuesta por las orejas de 100 individuos, tomadas desde 7 vistas diferentes: Cinco de las imágenes recogidas corresponden a la misma oreja (oreja derecha), con el individuo mirando al frente (FRONTAL, `_front`), mirando arriba y abajo (ARRIBA, `_up`, ABAJO, `_down`) y mirando a izquierda y derecha (IZQUIERDA, `_left`, DERECHA, `_right`). La sexta imagen corresponde a la misma oreja, con el individuo mirando al frente, pero modificando la distancia focal de la cámara con la que se captura la imagen (ZOOM, `_zoom`). La última imagen corresponde a la otra oreja del individuo (oreja izquierda), mirando al frente y tomada con la misma distancia focal de la cámara que las cinco anteriores (POSTERIOR, `_back`). Tener esta diversidad de vistas distintas, para los diferentes individuos, es muy importante en nuestro análisis, pues permite hacer validaciones cruzadas y verificar la robustez de los resultados. Una vez procesadas las 700 imágenes tomadas, se obtienen los vectores de características que componen la base de datos de prueba.
- Análisis discriminante y clasificación de las imágenes de orejas como pertenecientes a un individuo determinado. Con objeto de analizar el carácter discriminante, entre individuos, de los contornos de las orejas, y su poder de clasificación, hemos utilizado dos estrategias:

1. Asociando a cada individuo una clase formada por las 7 vistas, hemos realizado un Análisis Discriminante Clásico, basado principalmente en el estadístico Lambda de Wilks.

 2. Con objeto de introducir el poder de clasificación, como criterio en la selección de los parámetros discriminantes, hemos definido un método de selección de parámetros basado en la minimización de un funcional de energía, adaptado al caso de estudio, de manera que, el número de parámetros finalmente elegido sea el mínimo posible y, a la vez, el poder discriminante de esos parámetros sea máximo. Hemos analizado diferentes posibilidades para el cálculo de la distancia apropiada para comparar una imagen de oreja con la oreja "media" de un individuo, de cara a clasificarla como perteneciente o no a ese individuo. Concretamente hemos estudiado la distancia euclídea, la distancia normalizada a través de la varianza total y un refinamiento de esta última distancia mejorada, que es la que finalmente ha dado mejores resultados. Para aumentar la fiabilidad del análisis, se han utilizado algunas vistas para seleccionar los parámetros y otras diferentes para evaluar la capacidad discriminatoria de las mismas.
-
- Conclusiones principales obtenidas. Aunque a lo largo de la memoria se ilustrarán, con mucho más detalle, las conclusiones obtenidas, a título de resumen, podemos decir que la capacidad discriminatoria, entre individuos, de la forma que tienen los contornos interno y externo de las orejas, es excelente cuando manejamos vistas fronto-paralelas de la oreja. Sin embargo, cuando manejamos vistas de la otra oreja del mismo individuo, o vistas no fronto-paralelas, donde hay una distorsión de perspectiva significativa, la tasa de éxito del método propuesto se reduce notablemente. Otro aspecto interesante del trabajo desarrollado es que permite determinar qué zonas de la oreja son más significativas a la hora de clasificar individuos.

1.3. División del documento

A grandes rasgos, nuestra propuesta se divide en los siguientes pasos, que pasamos a analizar, más detalladamente, uno por uno, en los siguientes capítulos de esta tesis.

1. **Selección manual de la primera aproximación del contorno:** El usuario selecciona sobre la imagen, marcándolos con el ratón, unos puntos alrededor del contorno de la oreja, que se uniran formando un polígono cerrado, el cual se considerará la primera aproximación (bastante burda) al contorno de la oreja (imagen izquierda, figura 1-1).
2. **Localización del contorno de la oreja:** Usando una aproximación al método de los contornos activos, mejoramos la localización del contorno externo de la oreja en la imagen. Adaptamos el modelo de snakes para trabajar sobre el contorno de la oreja (que es una curva abierta) y obtenemos buenos resultados (imagen central, figura 1-1). En el capítulo 3 de esta tesis se describe en detalle este paso del proceso, que también se utiliza para la extracción del contorno interno de la oreja.
3. **Modelo de ajuste del lóbulo exterior de la oreja utilizando una elipse deformada:** Una vez que se ha estimado el contorno externo de la oreja, usaremos un modelo de ovoide para ajustar este contorno. Los parámetros que caracterizan este ovoide nos permitirán alinear los contornos extraídos de varias imágenes de orejas para compararlos, independientemente de su localización y tamaño (imagen derecha, figura 1-1). En el capítulo 4 de esta tesis se describe, en detalle, este proceso.
4. **Criterios de comparación entre la geometría de dos orejas:** Una vez obtenidas las coordenadas de los puntos de los contornos (tanto externo como interno) de las imágenes de las orejas correspondientes a cada individuo, necesitamos ponerlas todas en un sistema de referencia que nos permita comparar

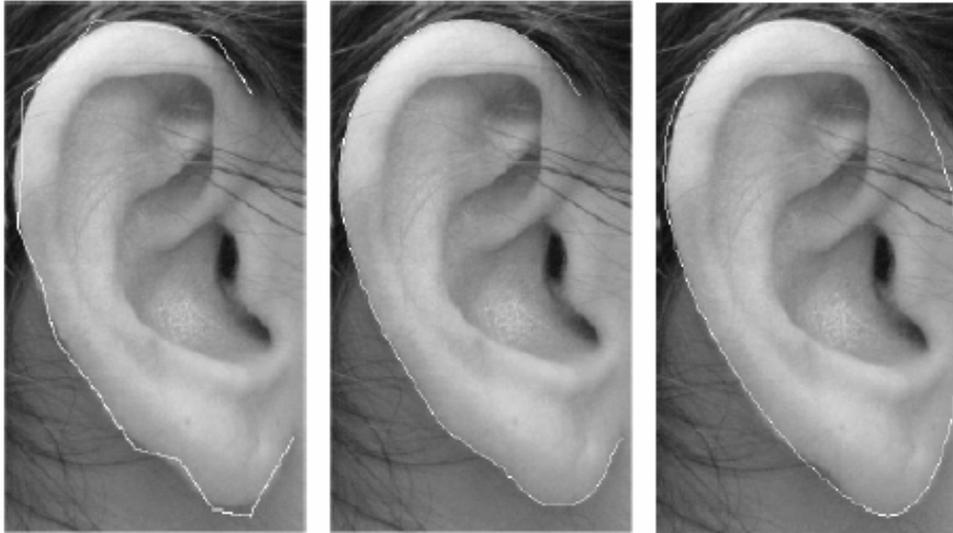


Figura 1-1: *Representación de las tres primeras fases del metodo propuesto. En la izquierda imagen original de la oreja con el polígono de aproximación al contorno inicial dibujado en blanco. En el centro, mejora de la aproximación del contorno usando los snakes (en blanco). A la derecha, ovoide estimado a partir del contorno de la oreja (en blanco).*

los contornos, para lo cual nos basamos en los parámetros del ovoide calculado anteriormente. Luego, para cada imagen, calculamos el centroide y obtenemos las distancias de este centroide a los puntos del contorno externo e interno de la oreja. De esta forma obtenemos un vector de características, que identifican la imagen y permiten clasificarla como perteneciente a un individuo determinado. Todo este proceso se explica más en detalle en el capítulo 5 de esta tesis (ver figura 1-2).

5. **Resultados experimentales y análisis estadístico:** Para la implementación del método hemos utilizado una base de datos con imágenes correspondientes a 100 individuos, 7 imágenes por individuo, tomadas todas en las mismas condiciones de iluminación. Para todas ellas hemos calculado el conjunto de distancias antes mencionado y, junto con los parámetros del mejor ovoide que ajusta cada contorno, se han sometido a un estudio estadístico (análisis discriminante) y

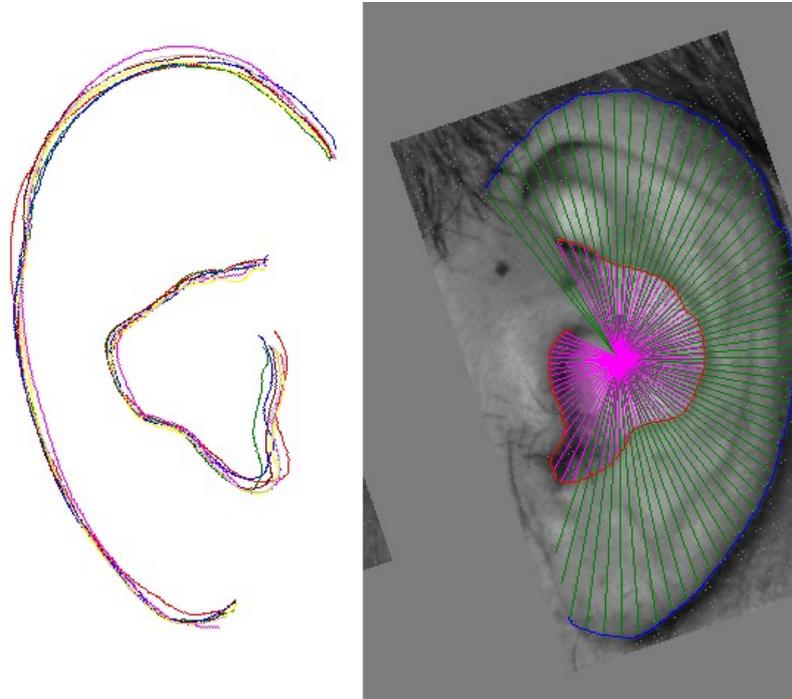


Figura 1-2: **(a)** Contornos internos y externos extraídos, correspondientes a 7 imágenes diferentes de un mismo individuo. **(b)** Distancias medidas desde el centroide a los contornos de la oreja

matemático para extraer, de entre todas estas medidas, el conjunto de características que mejor discrimina las imágenes. Los resultados obtenidos nos permiten señalar (para la población analizada) qué zonas de las orejas son más importantes, de cara a identificar al individuo, y otras zonas que presentan menos diferencias entre ellos. Este proceso, así como los resultados obtenidos, se explican en detalle en el capítulo 6 de esta tesis (ver figura 1-3).

- 6. Conclusiones y trabajos futuros:** El último capítulo de esta tesis se destina a exponer las conclusiones del trabajo realizado y a hacer propuestas de posibles ampliaciones futuras de este trabajo.

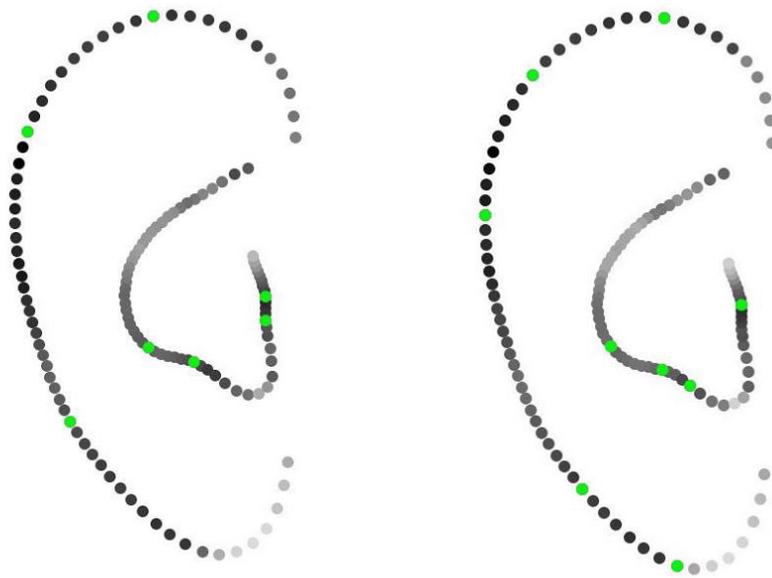


Figura 1-3: *Características que mejor discriminan las imágenes*

División del documento

Descripción general del trabajo realizado

Introducción

2.1. Introducción al problema del reconocimiento y la identificación personal. Biometría.

¿Es esta persona quien dice ser?. Hoy día esta pregunta aparece incesantemente. Una amplia variedad de sistemas requieren esquemas de reconocimiento personal, bien para confirmar o para determinar la identidad de un individuo que requiera sus servicios. El propósito de tales esquemas es asegurar que los servicios facilitados son accedidos solamente por los usuarios legitimados para ello y por nadie más. En muchas organizaciones como servicios financieros, e-comercio, telecomunicaciones, gobierno, tráfico, cuidado sanitario, etc... las medidas de seguridad se hacen cada vez más importantes. La ausencia de sistemas de reconocimiento personal robustos hace a estos sistemas vulnerables al ataque de impostores. Por esta razón, diferentes organizaciones están interesadas en sistemas automáticos de autenticación de identidad que mejoren la satisfacción del cliente y la eficiencia operativa, al mismo tiempo que ahorren costes y resulten más precisos que un humano [2].

Básicamente existen tres métodos diferentes de verificación de la identidad:

- algo que el usuario posee, como tarjetas de identificación, distintivos, llaves;

- algo que el usuario conoce, como claves de acceso, identificación de usuario, números de identificación personal (PIN);
- algo que el usuario es en sí mismo (biometría), como huellas dactilares, su cara, sus ojos, sus orejas.

Reconocimiento biométrico, o simplemente biometría, se refiere al reconocimiento automático de los individuos basándose en sus características fisiológicas o de comportamiento. La biometría ofrece una precisión mucho mayor que los métodos tradicionales de reconocimiento. Las posesiones pueden perderse, olvidarse o copiarse (replicarse) fácilmente. El conocimiento puede olvidarse. Ambas, las posesiones y el conocimiento, pueden ser robadas o compartidas con otras personas. En biometría estos inconvenientes no existen o existen en mucha menor medida [3],[4].

Los humanos han usado características tales como la cara, la voz, la forma de caminar, etc... durante miles de años para reconocerse unos a otros. Aunque la biometría tuvo su origen, y su mayor campo de utilidad, en servir de ayuda a las fuerzas legales en la identificación de criminales (autorizaciones de seguridad para empleados con trabajos comprometidos, determinación de la paternidad, ayuda a la identificación forense, identificación positiva de convictos o prisioneros, etc...) su uso para establecer el reconocimiento de personas en distintas aplicaciones civiles, se ha visto ampliamente incrementado ultimamente.

¿Qué cualidades debe reunir una característica biológica para servir como biometría? Una biometría ideal es *universal*, *única*, *permanente* y *medible*. Esto significa que cada persona debe poseer la característica (*universal*) y no debe haber dos personas que compartan la característica (*única*). La característica debe permanecer suficientemente invariante a lo largo del tiempo (*permanente*) y debe ser fácilmente recogida por algún sensor y cuantificable (*medible*) [2].

En la biometría se distinguen, básicamente, dos grupos de registros biométricos: los fisiológicos o morfológicos y los conductuales (ver figura 2-1).

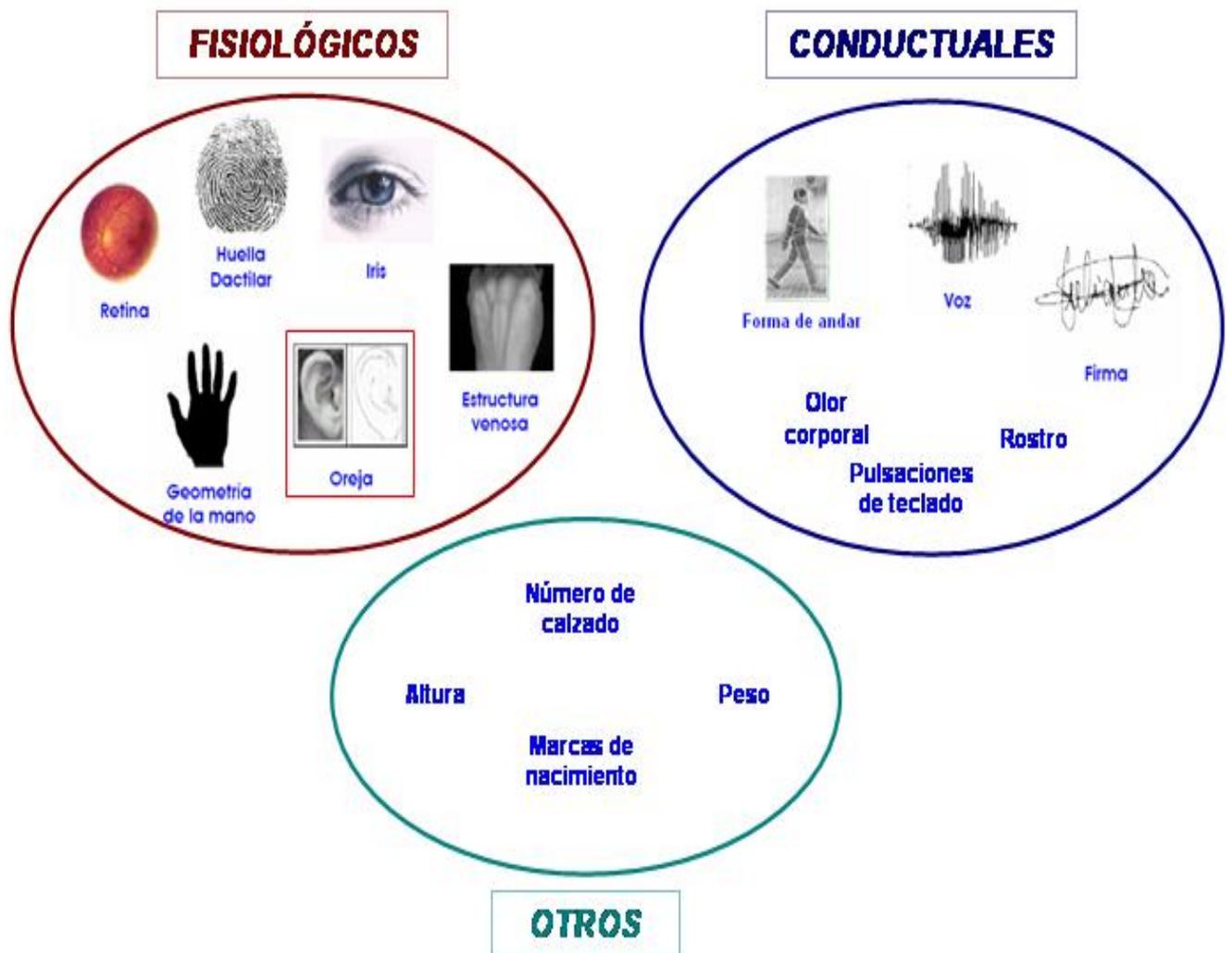


Figura 2-1: Métodos Biométricos

Los biométricos conductuales son aquellos que se soportan sobre características de la conducta del ser humano tales como: pulsaciones del teclado, discurso, dinámica de la firma, etc. Estos métodos requieren que el usuario tome parte activa en el proceso de identificación, repitiendo acciones específicas, cada vez que es examinado.

Los biométricos morfológicos o fisiológicos son aquellos que se soportan sobre características físicas inalterables y presentes en la mayoría de los seres humanos tales como: huella dactilar, geometría de la mano, características del iris, patrones vasculares de la retina, mano, etc. La mayor ventaja de este tipo de métodos es que son pasivos, y los sistemas que los implementan trabajan sólo con imágenes adquiridas de la parte del cuerpo en cuestión. Todo lo que el usuario tiene que hacer es colocarse frente a una cámara o tocar un sensor y esperar que se realice la identificación. Algunos sistemas, que probablemente serán el futuro de la biometría, son capaces de verificar la identidad de los humanos sin su cooperación e incluso sin su conocimiento. Monitorizaciones de lugares públicos como aeropuertos o recintos deportivos son las aplicaciones más importantes que requieren de este tipo de soluciones. Una de las novedades más interesantes en el campo de la identificación pasiva de los humanos es el uso de la oreja como fuente de información [5].

Cada característica biométrica tiene sus ventajas e inconvenientes dependiendo de la aplicación y no hay ninguna "óptima" que cubra los requerimientos de todas las aplicaciones. Muchas de estas características se usan actualmente en aplicaciones comerciales para identificación, siendo la huella digital, el reconocimiento facial, la geometría de la mano, el iris, la voz y la firma, las que más apoyo reciben por las industrias comerciales. La oreja se ha propuesto como biometría [6], aunque todavía no se ha implementado ningún sistema automatizado de reconocimiento de la oreja, ya que presenta la dificultad de que existen pocos detalles específicos para describirlas. Las personas poseemos una capacidad extraordinariamente desarrollada para el reconocimiento e identificación de las caras, ya que ésta posee muchos elementos diferenciadores (color de los ojos, tamaño de la nariz, grueso de los labios, etc...). Esta capacidad se utiliza por ejemplo

para la construcción de retratos-robot de determinados delincuentes. Sin embargo, no es habitual que una persona se fije en las orejas para identificar a otra persona (a menos que sean realmente inusuales).

2.2. Sistemas Biométricos. Aplicaciones

Un sistema biométrico es esencialmente un sistema de reconocimiento de patrones que opera adquiriendo datos biométricos de un individuo, extrayendo un conjunto de características de los datos adquiridos y comparando este conjunto de características con un conjunto de plantillas o patrones contenido en una base de datos. En la práctica, no todas las características biológicas, incluso cumpliendo las condiciones de universalidad, unicidad, permanencia y medibilidad indicadas anteriormente, sirven para un sistema biométrico. En un sistema biométrico hay otros aspectos que también deben tenerse en cuenta, a saber:

Según [2] *rendimiento*, *aceptación* y *vulnerabilidad*. *Rendimiento* se refiere a la precisión del sistema y la velocidad: si el sistema es muy lento o comete muchos fallos, no se usará. La *aceptación* por el usuario es importantísima: si la gente no acepta el sistema como parte de su rutina diaria, el sistema no se usará. La *vulnerabilidad* refleja como de fácil resulta engañar al sistema usando métodos fraudulentos: la tasa de engaños debe ser lo más baja posible.

Según [7] y [8], cuando se usa la biometría para propósitos de autenticación el sistema debe ser además *cómodo* (lo que significa que la duración de la verificación debe ser tan corta como sea posible y el sistema debe ser fácil de usar), *preciso* (la tasa de errores debe ser la menor posible), el sistema debe estar *disponible* cuando y donde se necesite y su *coste* no debe ser excesivo.

Los dispositivos biométricos tienen tres partes principales; por un lado, disponen de un mecanismo automático que lee y captura una imagen digital o analógica de

la característica a analizar. Además disponen de una entidad para manejar aspectos como la compresión, almacenamiento o comparación de los datos capturados con los guardados en una base de datos (que son considerados válidos), y también ofrecen una interfaz para las aplicaciones que los utilizan. En todos ellos se requiere un paso previo de registro o inscripción de usuarios, donde se introducen los datos de los usuarios habilitados (ver figura 2-2).

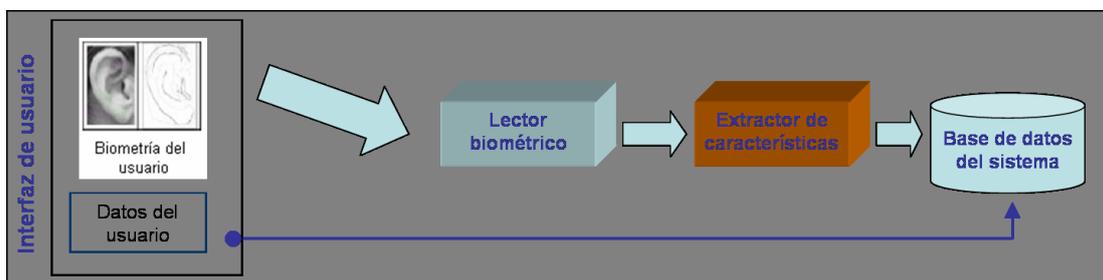


Figura 2-2: Módulo de Registro o Inscripción de usuarios

Dependiendo del contexto de aplicación un sistema biométrico puede operar en modo verificación (o autenticación) o en modo identificación (reconocimiento):

- En el proceso de autenticación o verificación (ver figura 2-3) los rasgos biométricos se comparan solamente con los de un patrón ya guardado, este proceso se conoce también como uno-para-uno (1:1). Este proceso implica conocer presuntamente la identidad del individuo a autenticar, por lo tanto, dicho individuo ha presentado algún tipo de credencial, que después del proceso de autenticación biométrica será validada o no. Es en esta decisión donde principalmente entran en juego las dos características básicas de la fiabilidad de todo sistema biométrico (en general, de todo sistema de autenticación): las tasas de falso rechazo y de falsa aceptación. Por tasa de falso rechazo (False Rejection Rate, FRR) se entiende la probabilidad de que el sistema de autenticación rechace a un usuario legítimo porque no es capaz de identificarlo correctamente, y por tasa de falsa aceptación (False Acceptance Rate, FAR) la probabilidad de que el sistema

autentique correctamente a un usuario ilegítimo; evidentemente, una FRR alta provoca descontento entre los usuarios del sistema, pero una FAR elevada genera un grave problema de seguridad: estamos proporcionando acceso a un recurso a personal no autorizado a acceder a él.

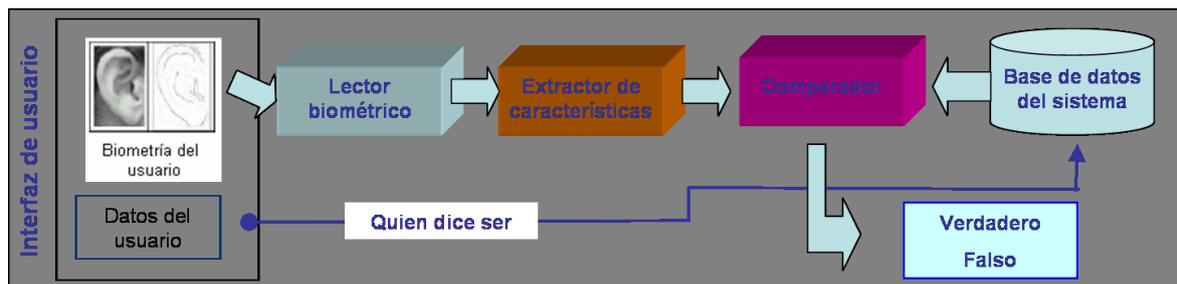


Figura 2-3: *Módulo de Verificación*

- En el proceso de identificación o reconocimiento (ver figura 2-4) los rasgos biométricos se comparan con los de un conjunto de patrones ya guardados, este proceso se conoce también como uno-para-muchos (1:N). Este proceso implica no conocer la identidad presunta del individuo, la nueva muestra de datos biométricos es tomada del usuario y comparada una a una con los patrones ya existentes en el banco de datos registrados. El resultado de este proceso es la identidad del individuo, mientras que en el proceso de autenticación es un valor verdadero o falso.

El proceso de autenticación o verificación biométrica es más rápido que el de identificación biométrica, sobre todo cuando el número de usuarios (N) es elevado. Esto es debido a que la necesidad de procesamiento y comparaciones es más reducido en el proceso de autenticación. Por esta razón, es habitual usar autenticación cuando se quiere validar la identidad de un individuo desde un sistema con capacidad de procesamiento limitada o se quiere un proceso muy rápido.

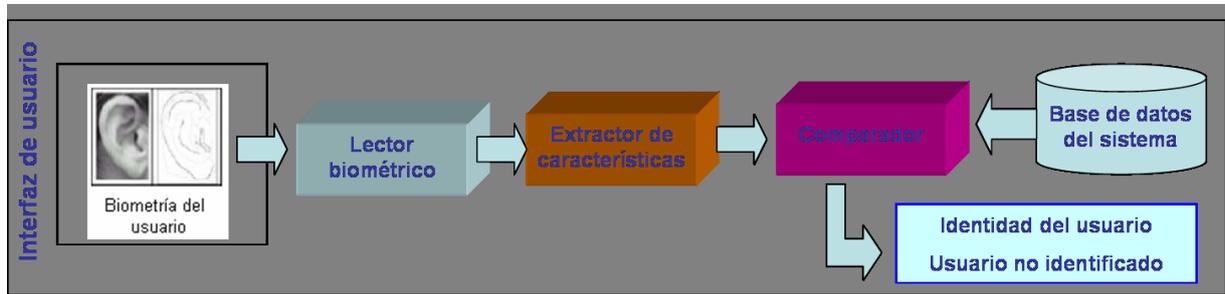


Figura 2-4: Módulo de Identificación

Las aplicaciones actuales de los sistemas biométricos se pueden dividir en tres grandes grupos:

1. Aplicaciones comerciales: Login en redes de ordenadores, seguridad de datos electrónicos, e-comercio, acceso a internet, tarjetas de crédito, control de acceso físico (a edificios, instalaciones, etc...), teléfonos móviles, manejo de registros médicos, enseñanza a distancia, etc...
2. Aplicaciones gubernamentales: Documento nacional de identidad electrónico, permisos de conducir, seguridad social, control de fronteras, control de pasaportes, etc...
3. Aplicaciones forenses: identificación de cadáveres, investigación criminal, identificación de terroristas, determinación de paternidad, niños perdidos, etc...

Tradicionalmente las aplicaciones comerciales usaban sistemas basados en conocimiento (PIN's y passwords), las aplicaciones gubernamentales utilizaban sistemas basados en posesiones (tarjetas de identificación y distintivos) y aplicaciones forenses utilizaban expertos humanos para comparar características biométricas. Actualmente los sistemas biométricos han visto incrementado su uso en un amplio espectro de aplicaciones civiles, como muestra la figura 2-5:



Figura 2-5: Ejemplos de aplicaciones biométricas: (a) Sistema de verificación de huella dactilar fabricado por Digital Persona Inc., se utiliza para la identificación de usuarios en ordenadores y redes. (b) Terminal de punto de venta basado en huella dactilar, fabricada por Indivos, Inc. que verifica la identificación de los clientes para el cambio de tarjetas de créditos y para agilizar los pagos en tiendas al detalle, cafeterías, restaurantes, etc... (c) Bloqueador de puertas basado en huella dactilar, fabricado por Biothentica Corporation, usado en lugares de acceso restringido. (d) Sistema acelerado de servicio de pasajeros del Ministerio de Inmigración y Naturalización de EE.UU. (INSPASS), instalado en los mayores aeropuertos americanos, está basado en una tecnología de verificación de la geometría de la cabeza, desarrollado por Recognition Systems, Inc. que reduce significativamente el tiempo de proceso de los trámites de inmigración. (e) Sistema de reconocimiento de pasajeros que cruzan la frontera, usando reconocimiento de iris, instalado en el aeropuerto de Heathrow en Londres. (f) El aeropuerto de Ben Gurion, en Tel Aviv (Israel) usa un sistema basado en el reconocimiento de la geometría de la mano para seguridad e inmigración. (g) El sistema de identificación de rostros FacePass de Viisage se utiliza en aplicaciones de verificación POS como ATM's, evitando así la necesidad de los PIN. (h) El sistema basado en reconocimiento de huella Identix TouchClock se utiliza en aplicaciones de control horario y de asistencia de personal.

2.3. Empleo de imágenes de la oreja para identificación o reconocimiento.

El uso de la oreja en la identificación de personas podemos remontarlo, al menos, 100 años atrás. Alfonso Bertillón (1852 - 1914), jefe de la división de identificación criminal del departamento de policía de París, utilizó la oreja (entre otras medidas corporales) como medio de identificación de criminales.

La biometría de la oreja es, a menudo, comparada con la biometría facial [9], [6], [10], aunque la oreja presenta varias ventajas frente a la cara:

1. La oreja no cambia de manera significativa a partir del estado adulto del individuo; la cara, por el contrario, sigue cambiando (lentamente) durante toda la vida del individuo.
2. La cara cambia notablemente de aspecto con la expresión (tristeza, alegría, miedo, sorpresa, etc...), lo cual es una seria dificultad para su procesamiento por el ordenador. Por el contrario la forma y aspecto de la oreja son fijos (salvo mutilaciones u otros casos extremos).
3. La distribución del color es más uniforme en la oreja que en la cara (piénsese en los ojos, por ejemplo), por lo que se pierde menos información al trabajar con imágenes de tonos de gris o blanco y negro.
4. La superficie ocupada por la oreja es más pequeña (aproximadamente $1/20$ ó $1/25$ de la de la cara), lo que permite utilizar imágenes de menor resolución y hacer más eficiente y económico el proceso.

También presenta algunas ventajas respecto a otras biometrías como la voz o las huellas dactilares:

1. Mientras que la biometría de la oreja solo necesita el uso de una cámara CCD ordinaria, el reconocimiento de las huellas dactilares requieren el uso de sensores especialmente diseñados, que pueden resultar muy caros para su empleo a gran escala.
2. El reconocimiento de las orejas es un procedimiento menos invasivo que el de las huellas dactilares, lo cual puede ser esencial en escenarios no cooperativos (i.e. identificación de terroristas). Esta naturaleza pasiva de la captura de imágenes reduce la oportunidad de tapar/dañar el dispositivo de captura.
3. La calidad de la voz de un sujeto puede variar enormemente con el estado de salud y verse afectado por ruido de fondo, mientras que la apariencia de la oreja es invariante respecto a estas variables.

Sin embargo hay varias posibilidades de error en la identificación de la oreja:

1. Básicamente, la forma de la oreja es la misma durante toda la vida y el crecimiento proporcional. A pesar de ello, la gravedad puede causar algún estiramiento o alargamiento (más apreciable en la parte del lóbulo). Este estiramiento es mayor entre la edad de 4 meses hasta los 8 años y, nuevamente, a partir de los 70 años [12].
2. Las orejas pueden aparecer total o parcialmente tapadas por el pelo o por pendientes; aunque lo mismo le ocurre a la cara (si bien en menor medida) con el maquillaje, barba, peinado, etc..., o al iris con el uso de gafas o lentes de contacto.
3. Como en cualquier representación plana de un objeto espacial, hay información que se pierde. En el caso de la cara, la longitud de la nariz, etc... (aunque con ayuda de la imagen de perfil pueden recuperarse en gran medida); en el caso de la oreja, la separación del cráneo, la profundidad de los diversos pliegues, etc...
4. En ambos casos hay que tener cuidado con las sombras creadas por la iluminación lateral.

La estructura de la oreja es bastante compleja (ver figura 2-6). El pabellón auricular está constituido por un esqueleto cartilaginoso revestido por tegumentos. La lámina de cartílago se pliega sobre sí misma formando relieves y depresiones que confieren al pabellón su forma característica que ha sido sobradamente descrita en los textos anatómicos. Hay evidencias claras que sugieren que la forma de la oreja es única. El trabajo más famoso en este sentido fue realizado por Alfred Iannarelli en 1989, quien, habiendo reunido imágenes de unas 10.000 orejas, encontró que todas ellas eran diferentes. Un segundo estudio posterior probó que incluso las orejas de gemelos idénticos tenían similares, pero no idénticas, características fisiológicas [6], [7], [9] y [13]. Ya en 1906, Imhofer había encontrado que, en un conjunto de 500 orejas, sólo 4 características eran necesarias para indicar la unicidad de las mismas. La principal limitación de estas técnicas es que requieren de la intervención de un especialista y, algunas de las medidas que se toman, son subjetivas y difíciles de calcular de manera automática y robusta.

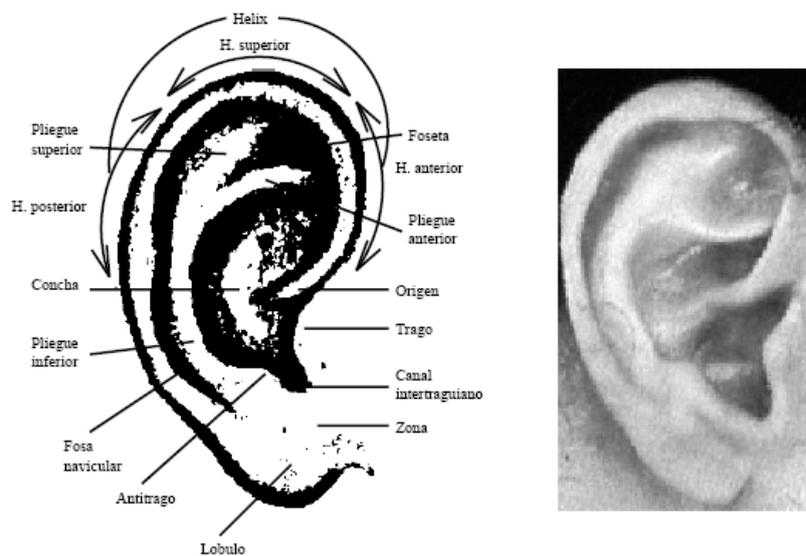


Figura 2-6: *Estructura de la oreja*

Un sistema de reconocimiento de orejas consta básicamente de cinco fases:

1. **Adquisición de la imagen:** Existen varios métodos posibles de captura de la imagen para identificación de la oreja. Los más usados son estos tres:

- a) *Comparación de fotos:* Es el método utilizado en el trabajo de esta tesis y el que mayor número de investigadores comparten, tal como se explicará en detalle en el próximo capítulo. A partir de una imagen de la oreja, que puede ser una fotografía o una imagen extraída de una secuencia de video, se obtienen ciertas características de la oreja, que se comparan entre las distintas imágenes.
- b) *Huellas de la oreja:* Este método es de uso más reciente y principalmente utilizado por la medicina legal y forense para la identificación de delincuentes. La huella de oreja es una reproducción en dos dimensiones de las partes del pabellón auricular que se han puesto en contacto con una determinada superficie y que, habitualmente, son las regiones más prominentes del mismo. Este tipo de huellas, al igual que las dactilares, se producen por los restos de descamación, sudor y grasa que deja la piel al contacto con una superficie. No son manifiestas a simple vista, pero se recuperan de forma muy simple con reveladores físicos o químicos. La razón habitual de producirse las huellas es sencilla: los delincuentes que van a robar un domicilio apoyan la oreja sobre la puerta para comprobar que no hay nadie en su interior y ahí dejan el indicio. La primera identificación de un delincuente por la huella de oreja se realizó en Suiza en 1965. En Inglaterra 4 delincuentes han sido juzgados entre 1996 y 1989 usando sólo las huellas de la oreja [15]. En Holanda existen ya más de 200 casos judiciales de identificaciones por huella de oreja [7]. En España la primera sentencia condenatoria basada parcialmente en este tipo de evidencia se produjo el 19 de noviembre de 2001. A esta sentencia le han seguido al menos otras dos, y se han realizado más de 20 identificaciones. Existen en la actualidad bases de datos de huellas de oreja en Palencia, Valladolid, Santander y Lleida. En febrero de 2002 se puso en marcha el proyecto

FEARID (“Forensic EAR IDentification”), aprobado por la Unión Europea, como respuesta a la necesidad de una investigación científica estricta y un estudio sistemático de las huellas de orejas y el desarrollo de herramientas de apoyo automatizado para su cotejo. El objetivo es establecer un proceso estándar de detección, recuperación, almacenamiento e identificación de las huellas de oreja y el establecimiento mediante programas informáticos de una base de datos paneuropea de huellas de oreja que permita un cálculo estadístico y aumente la potencia de la prueba dando una validez científica y judicial a la identificación por huellas de oreja [17].

c) Termogramas: En este caso también se utilizan imágenes de las orejas. Cuando la oreja está parcialmente cubierta por el pelo este puede ser diferenciado en la imagen usando termogramas. En el termograma se usan diferentes colores y texturas para encontrar las diferentes partes de la oreja. La oreja es fácilmente detectable y localizable usando termogramas, simplemente buscando las áreas de mayor temperatura en la imagen [13].

2. **Preprocesado:** Se utilizan técnicas estándares tales como equalización y normalización de histogramas
3. **Interpretación de los datos y extracción de características:** Muchos investigadores han realizado varias propuestas que difieren drásticamente unas de otras en términos de adquisición e interpretación de datos y en la extracción de características. Algunas de estas propuestas ya se había demostrado que eran prácticas en el campo del reconocimiento humano (por ejemplo PCA [6], redes neuronales [16], etc...), mientras que otras presentan una perspectiva totalmente novedosa (por ejemplo la transformación de campos de fuerza [10]). Las comentaremos en más detalle en el próximo capítulo.
4. **Comparación de patrones:** Se utilizan habitualmente métodos estándares de análisis estadístico de errores.

Estado del arte

En este capítulo daremos un breve repaso a las principales contribuciones realizadas hasta el momento.

El primer método de reconocimiento de orejas, manual, fue llevado a cabo por Alfred Iannarelli [18], quien realizó dos estudios para identificación de orejas, a gran escala, en 1989. El segundo de estos estudios, destinado a demostrar la unicidad de las orejas, se llevó a cabo entre gemelos idénticos y no idénticos y obtuvo como conclusión que, incluso los gemelos idénticos, tenían orejas parecidas pero no iguales.

Iannarelli trabajó durante 30 años como ayudante del sheriff en el condado de Alameda, en California, como policía del campus en la Universidad estatal de California, en Hayward, y en varios otros puestos relacionados con las fuerzas de seguridad. Empezó a interesarse por las orejas en 1948 y, en los 14 años siguientes, clasificó cerca de 7000 orejas a partir de fotografías. La primera versión de su libro describiendo su método de clasificación fue publicado en 1964. La segunda edición se publicó en 1989. Iannarelli no usaba base académica alguna para apoyar sus estudios [19].

Alfred Iannarelli creó, para clasificar las orejas, un sistema de 12 medidas, el "Sistema Iannarelli"(ver figura 3-1).

Usaba la oreja derecha de cada individuo en fotografías especialmente alineadas y

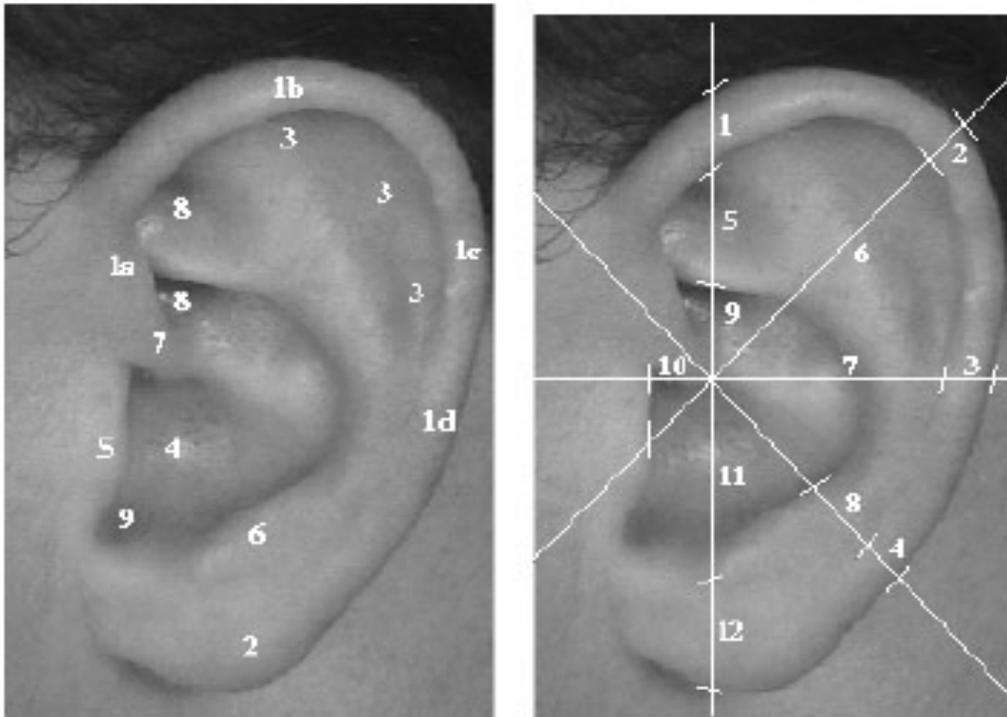


Figura 3-1: *(a) Anatomía, (b) Medidas. (a) 1.Llanta helicoidal, 2.Lóbulo, 3.Antihélice, 4.Concha, 5.Trago, 6.Antitrago, 7.Cruce de hélice, 8.Fosa triangular, 9.Incisura intertrágica. (b) Localización de las medidas antropométricas usadas en el "Sistema Iannarelli" [12]*

normalizadas. Para normalizar las imágenes, éstas se ampliaban hasta que encajaran en un caballete predefinido. Después se tomaban las medidas directamente sobre las fotografías. Se medían las distancias entre las áreas numeradas y se les asignaba un valor entero. Para la clasificación se utilizaban las 12 medidas obtenidas junto con información referida al sexo y la raza del individuo. Burge [12] comentó que este método no era apropiado para la visión artificial debido a la dificultad para la localización, de forma automática, de los puntos anatómicos clave descritos en el sistema. Además de esto, si el primer punto no estaba perfectamente definido, ninguna de las medidas restantes resultaba útil. El mismo Iannarelli había reconocido esta limitación de sus sistema.

Después de la clasificación de Iannarelli han aparecido muchos otros métodos, más científicos, para identificación de las orejas:

3.1. Redes neuronales usando combinación de varias técnicas

Moreno y colaboradores [4] investigaron la actuación de varios clasificadores neuronales y propusieron una técnica combinada (de clasificación múltiple) para la identificación de las orejas. La imagen de la oreja de una persona era analizada por clasificadores neuronales usando puntos característicos del contorno externo de la oreja, información obtenida de la forma de la oreja y sus pliegues, y macro-características extraídas por redes de compresión (ver figura 3-2). Los resultados de esta clasificación se combinaban entonces con otras técnicas diferentes de clasificación. Ellos proponían concretamente técnicas de 3 tipos: clasificadores deterministas como Voto Mayoritario (Majority Vote), clasificadores de rango como Contador Borda (Borda Count) y clasificadores probabilísticos como Pesos Bayesianos (Weighting Bayesian).

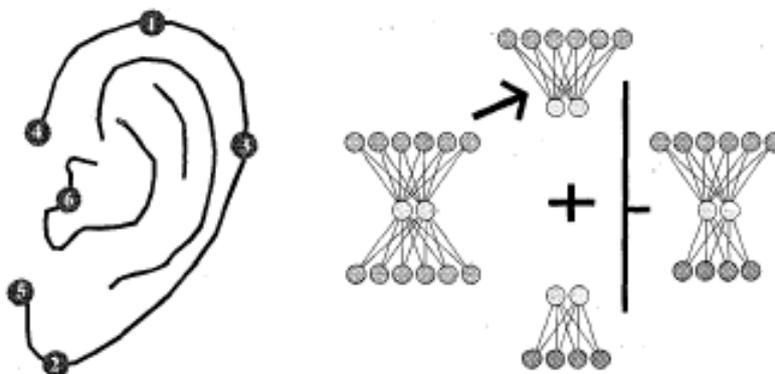


Figura 3-2: (a) Puntos característicos considerados en el contorno externo de la oreja. (b) Construcción de una red neuronal de compresión.

La base de datos usada en su experimento estaba separada en dos grupos. El primer grupo, compuesto por 168 imágenes, incluía 28 individuos distintos, con 6 fotos por persona. El segundo grupo estaba formado por 20 imágenes, de 20 individuos diferentes, 1 foto por persona, siendo todos los individuos diferentes de los del primer grupo. El primer grupo de imágenes se dividía, a su vez, en 3 subgrupos: un grupo de entrenamiento (3 imágenes por individuo), un grupo de validación (1 imagen por individuo) y el resto de imágenes se agrupaban formando el grupo de prueba. El segundo grupo se usó enteramente para prueba.

Se llevaron a cabo dos categorías de experimentos, unos con un umbral de rechazo fijado de antemano y otros sin él. Los resultados mostraron que las redes de compresión alcanzaban los mejores resultados de identificación (93 %) sin considerar umbrales de rechazo. Además, la combinación de clasificadores no incrementaba, en general, la tasa de identificación, ya que los clasificadores no eran del todo independientes. Sin embargo, si se tenían en cuenta los umbrales de rechazo, estas técnicas combinadas incrementaban las identificaciones correctas (84 % por el método pesos bayesianos frente a 79 % con redes de compresión).

3.2. Encaje de gráficos de vecindad usando diagramas de Voronoi

Burge y colaboradores propusieron el uso de la coincidencia gráfica para la identificación de las orejas [12] [13] y [14]. En esta propuesta, la oreja de cada individuo se modelaba como un gráfico de adyacencia construido a partir de los segmentos de curva en un diagrama de Voronoi, el cual se extraía por detección de bordes, como se muestra en la figura (ver figura 3-3).

Para la autenticación, se usaba un algoritmo de corrección de los errores producidos en el encaje gráfico, que tenía en cuenta los segmentos de curva erróneos que pudieran

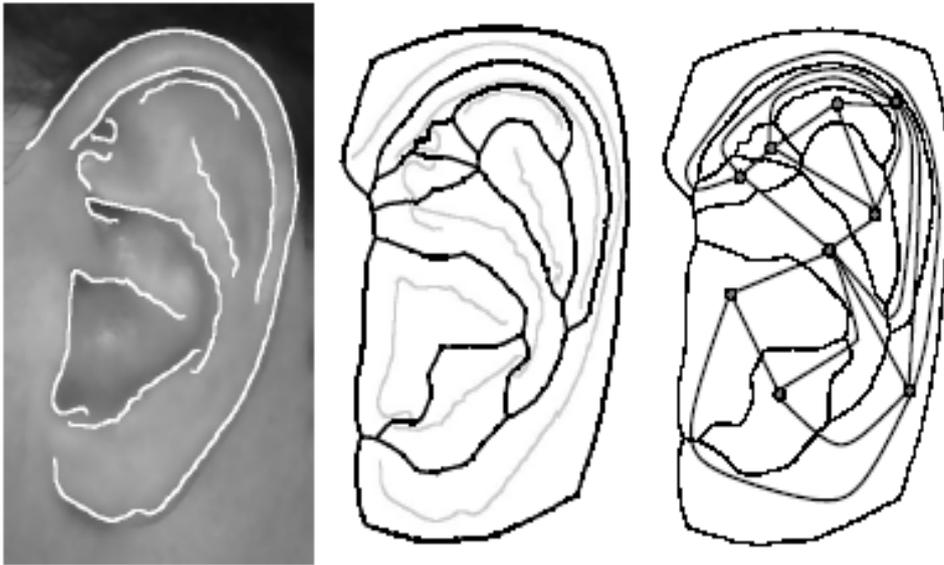


Figura 3-3: (a) Impresión de la oreja. (b) Diagrama de Voronoi. (c) Gráfico de vecindad

darse debido a cambios en la imagen de la oreja.

En su estudio no se facilitaban resultados experimentales, sin embargo, en un intento de implementar esta propuesta [9], se encontró que la descripción básica de la oreja utilizada en este modelo era bastante inestable. Un problema era que, los bordes detectados a partir de la imagen de la oreja, podían ser diferentes debido a la presencia de cambios de orientación cámara-oreja o de la iluminación, incluso aunque estos cambios fueran relativamente pequeños. Los bordes detectados en la imagen de la oreja aparecían, la mayor parte de las veces, debido a contornos ocultos más que a discontinuidades de la superficie; motivo por el cual, estos bordes, naturalmente cambiarían de forma sustancial ante cualquier cambio en la orientación o iluminación.

Burge y colaboradores también sugirieron el uso de imágenes termográficas (ver figura 3-4) como ayuda para reconocer partes de la oreja parcialmente tapadas por el pelo [12]. Segmentando las áreas de baja temperatura localizadas en la anatomía externa de la oreja se podría eliminar la oclusión provocada por el pelo. Otra observación

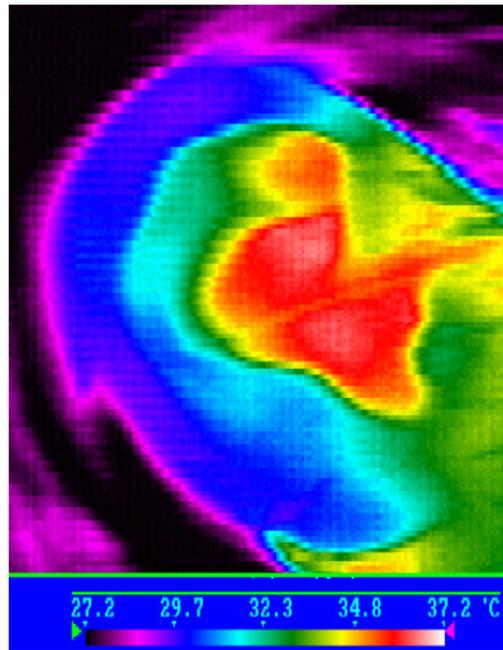


Figura 3-4: Termograma de una oreja. Imagen facilitada por Brent Griffith, Laboratorio Termográfico de Infrarrojos, Laboratorio Nacional Lawrence Berkeley [12]

interesante era que, el pasaje que conduce al oído interno (meatus) era fácilmente localizable, por ser la parte más caliente de la imagen, con una diferencia de temperatura respecto del pelo que puede taparlo, de hasta 8 grados centígrados.

3.3. Transformación de campos de fuerza

Hurley y colaboradores [11] atacaron el problema de descripción de la oreja simulando el proceso natural de los campos de fuerza electromagnéticas. Tratando cada pixel de la imagen como un atractor gaussiano, la imagen de la oreja se transformaba en un campo de fuerzas (un vector de partículas mutuamente atrayentes, cuya fuerza de atracción era proporcional a la intensidad del pixel). Cuando pixeles exploratorios (de valor unitario) se colocan en el campo de fuerza, estos tendían a moverse naturalmente hacia los extremos locales de la superficie de energía potencial ("pozos de potencial")

bajo la influencia de la fuerza, formando canales.^{en} su camino hacia la estabilidad. La posición de los pozos de potencial servían de base para generar el vector de características (ver figura 3-5), que resultaba compacto ya que se extraían pocos pozos de potencial.

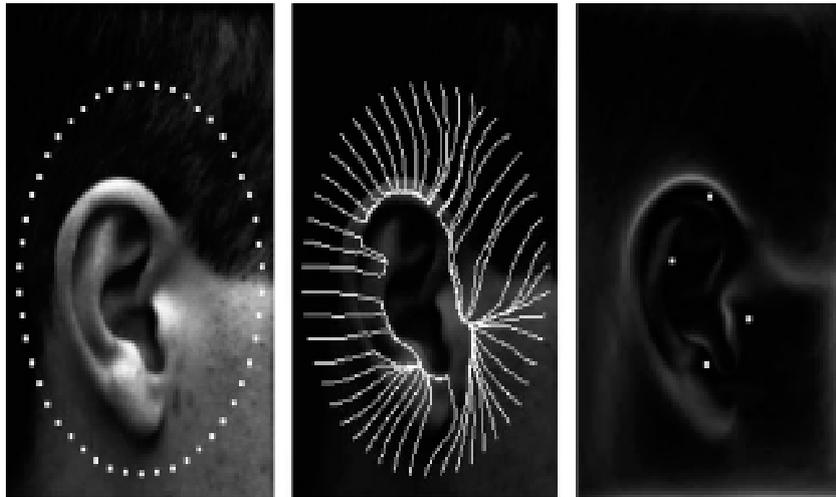


Figura 3-5: *Extracción de los pozos de potencial y los canales. (a) Inicialización, (b) Formación de los canales, (c) Posición de los pozos de potencial*

Los resultados mostraban que el vector de características era altamente inmune a inicialización, rotación, escalado y ruido. Se demostraba también que la transformación del campo de fuerzas era una transformación lineal invertible y, por lo tanto, toda la información contenida en la imagen se conservaba en la transformación [10]. Utilizando esta transformación se hicieron pruebas de reconocimiento automático de orejas en una pequeña base de datos [11]. Los resultados iniciales fueron prometedores, aunque no se facilitaron algunos detalles del experimento y de la implementación. La topología de los canales de potencial, que no se usó en el experimento, se podría aprovechar mejor, de cara a aumentar la discriminación, con poblaciones mayores de orejas, que necesitarían una mayor extracción de información [10].

3.4. Análisis de los componentes principales (PCA)

El Análisis de los Componentes Principales (PCA) es, con diferencia, el método más ampliamente adoptado, de todos los métodos usados en la investigación de la biometría de la oreja [6], [9], [20]. El hecho de que el PCA haya sido ampliamente investigado en el campo de la visión por ordenador, especialmente en el reconocimiento facial, hace obvia la elección de esta propuesta. PCA es una técnica para reducir la dimension del vector de características preservando las variaciones del conjunto de datos. Para la clasificación, se usa un espacio de menor dimensión que viene definido por un conjunto de autovectores (Componentes Principales) del conjunto de datos. A veces se utiliza el término 'eigen ears' para referirse al conjunto de autovectores que describen la oreja humana [6], en el mismo sentido en que se inventó el término 'eigen-face'.

Víctor y colaboradores compararon la aplicación del PCA al reconocimiento de la cara y de la oreja [6]. En sus experimentos se usó un total de 294 sujetos. El número de imágenes usadas en el entrenamiento fue de 207 para ambas cosas (orejas y caras). Las imágenes de entrenamiento para las caras eran tomadas del sujeto con una expresión facial normal mientras que, para las orejas, eran imágenes de la oreja derecha del individuo. La colocación relativa del sujeto, las luces y la cámara, para la captura de imágenes, seguían el protocolo FERET de adquisición de imágenes del rostro, que se ha convertido en un estándar de hecho para la evaluación de algoritmos de reconocimiento facial [9]. Para evaluar la implementación se utilizó una galería diferente de imágenes (los mismos individuos tomadas en diferentes fechas) y conjuntos de prueba. Se tuvieron en cuenta factores tales como la iluminación, la expresión y el lapso de tiempo transcurrido entre la adquisición de imágenes. Los resultados mostraron que, en todos los experimentos, el reconocimiento de rostros funcionaba mejor que el reconocimiento de orejas.

Sin embargo, en experimentos similares, llevados a cabo por Chang y sus colaboradores [9], se alcanzaba una conclusión diferente. Se encontró que no había diferencia

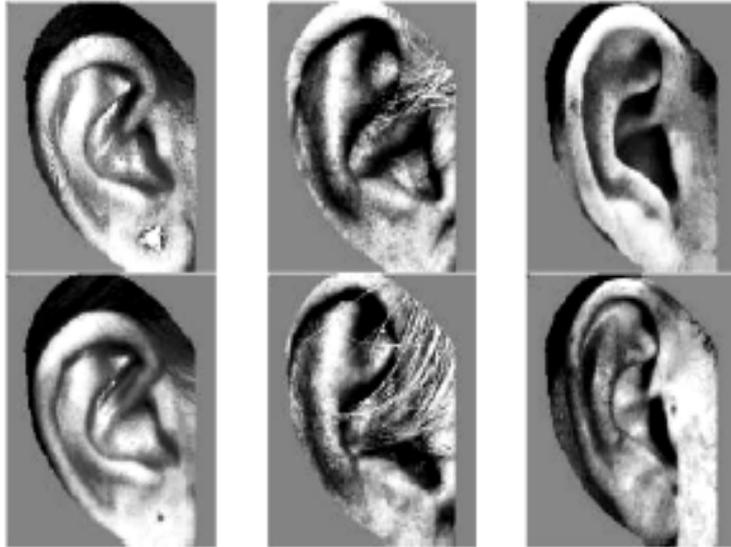


Figura 3-6: *Problemas que pueden aparecer con las imágenes de las orejas. (a) Pendientes, (b) Ocultación de la oreja, (c) Iluminación inconsistente.*

significativa entre el reconocimiento de rostros y orejas. La calidad de las imágenes de caras y orejas en la base de datos de estos últimos experimentos estaba más rigurosamente controlada y se habían eliminado del estudio aquellas imágenes en las que la cara o la oreja quedaban sustancialmente oscurecidas por el pelo o por el uso de pendientes (ver figura 3-6). Los resultados así obtenidos sugerían que la cara y la oreja podrían tener un valor similar para el reconocimiento biométrico (en un experimento se obtuvieron tasas de reconocimiento de 70.5% y 71.6% para orejas y caras respectivamente). Otro hallazgo importante de este estudio fue que el reconocimiento bimodal (usando a la vez la cara y la oreja) ofrece una mejora, estadísticamente significativa, de los resultados obtenidos usando cada uno de los biométricos de forma separada (por ejemplo un 90.9% en un experimento).

La diferencia de conclusión alcanzada por los dos estudios podría deberse a la calidad de los datos, ya que la base de datos de imágenes utilizada en el primer estudio tenía menor control sobre variaciones tales como el uso de pendientes, el pelo cubriendo las orejas, etc... Sin embargo, los resultados de este estudio no deberían considerarse

sesgados, en tanto en cuanto el objetivo perseguido fue reflejar la calidad media de las imágenes que se pueden adquirir en aplicaciones reales.

3.5. Algoritmos genéticos

El algoritmo genético (GA) es una aplicación de la regla de selección natural - sobrevive el que mejor se adapta-, encontrada en la evolución de los organismos, a la resolución de problemas complejos. Generalmente se aplica a la resolución de problemas que conllevan una complejidad computacional que aumenta de forma exponencial.

Yuizono y sus colaboradores [20] trataron el problema de reconocimiento de imágenes de orejas como un problema de optimización y le aplicaron GA. Además de las operaciones genéticas básicas tales como selección, cruces y mutaciones, al método GA se le añade una búsqueda local, lo cual se sugiere que incrementará la eficiencia en la búsqueda de valores óptimos locales. Este método combinado recibe el nombre de búsqueda genética local.

En el experimento realizado en [20], los autores emplearon 660 imágenes de 110 individuos (6 por participante). 3 imágenes de las primeras 80 personas se usaron para determinar los parámetros de reconocimiento usando GA. Las primeras 50 personas se etiquetaron como registradas, otras 30 como no registradas y las 30 restantes se trataron como datos completamente desconocidos para la evaluación del experimento de reconocimiento. Se probaron y evaluaron varios métodos de selección GA diferentes como: búsqueda local y mantenimiento de la élite, suma de la élite, búsqueda local y suma de la élite, y mantenimiento de la élite. Los resultados mostraron que, la tasa más alta de reconocimiento de los registrados, alcanzaba prácticamente el 100% y la tasa de rechazo de los no registrados también se aproximaba al 100%. Con una selección óptima de los parámetros de búsqueda, el tiempo medio de búsqueda para un registrado era de 1.5 segundos en un PC, a 633Mhz y ejecutando JDK 1.4.

3.6. Comparación de trozos de superficie local usando sensores de alcance

Todas las investigaciones previas en la biometría de la oreja usaban imágenes de intensidad en 2D obtenidas usando cámaras CCD, viéndose afectadas por tanto, de forma inevitable, por problemas en la imagen tales como sombras o variaciones de la iluminación o de la posición. Una propuesta alternativa, realizada por Bhanu y Cheng [21], utilizaba sensores de alcance, que son insensibles a los problemas de la imagen antes mencionados y permiten obtener directamente información geométrica de la oreja en 3D. En esta propuesta, se usaba un nuevo descriptor local para la representación de la superficie.

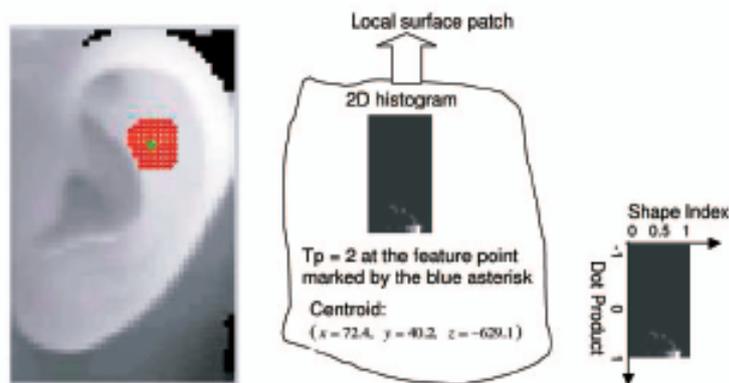


Figura 3-7: Ilustración de un trozo de superficie local (Local Surface Patch). (a) El punto característico P es marcado con un asterisco y sus N vecinos se marcan con puntos interconectados. (b) Representación de la LSP incluyendo un histograma 2D, un tipo de superficie y las coordenadas del centroide. (c) Se muestra el histograma 2D como una imagen gris en la que las áreas brillantes corresponden a zonas donde se deposita la frecuencia más alta de ocurrencia.

Este descriptor local de superficie (ver figura 3-7) viene definido por un centroide, su tipo de superficie y un histograma 2D. El tipo de superficie se determina por los signos de la curvatura media y la curvatura gaussiana.

El histograma 2D consta de índices de formas, calculados a partir de las curvaturas principales, y ángulos entre la normal del punto de referencia y las de sus vecinos. Como el coste de cálculo del descriptor local de superficie, para cada pixel en la imagen de alcance, es muy alto computacionalmente hablando, se calculan sólo los de los puntos de referencia. Los puntos de referencia se definen como los máximos y mínimos locales de los índices de forma, calculados a partir de las curvaturas principales.

Se obtiene una lista de posibles correspondencias entre trozos de la superficie local, comparando los de la imagen de prueba con los de una imagen modelo y se aplican luego algunas restricciones geométricas para filtrar las parejas correspondientes. La verificación se lleva a cabo mediante la estimación inicial de algunas transformaciones y la posterior alineación de puntos de la imagen de prueba con la imagen modelo. Se localizan luego los pares de correspondencia más cercanos usando como medida la distancia euclídea y, finalmente, se calcula la calidad de ajuste midiendo la ratio entre los puntos de correspondencia y el número de puntos totales del modelo.

Se usaron las imágenes de alcance de 10 orejas como base de datos de prueba del experimento. Los resultados mostraron que la mayoría de los números más altos de pares correspondientes pertenecían a los modelos de oreja correctos. El modelo con la máxima calidad de ajuste se elegía como la oreja reconocida. Se alcanzó una tasa del 100% de reconocimiento para el conjunto de datos utilizado.

3.7. Extracción de características geométricas

Michael Chorás [5] considera que, las características geométricas que representan la forma de los contornos de las orejas, son más apropiados para el reconocimiento de este tipo de imágenes que la textura, el color o características globales. Por este motivo, propone otro método para extraer las características necesarias para la clasificación de las imágenes de las orejas. El método propuesto es una aproximación geométrica

automática (no necesita ninguna operación manual) e invariante, que se realiza en varios pasos (ver figura 3-8).

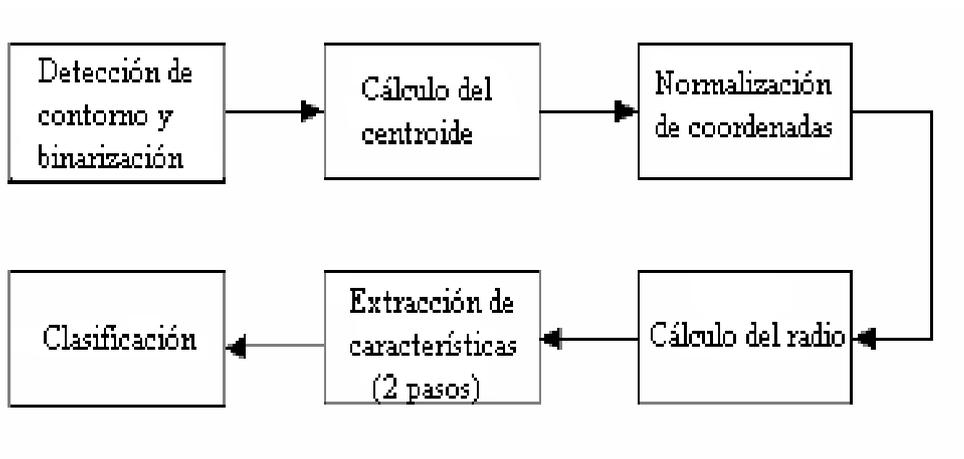


Figura 3-8: Diagrama de flujo del modelo

Primero se ejecuta un algoritmo de detección de contornos, del cual resulta una imagen binaria (ver figura 3-9); se calcula el centroide de esa imagen, y se normalizan las coordenadas, de manera que el centroide se convierta en el punto central de la imagen. Localizando el centro del nuevo sistemas de coordenadas en el centroide, el método resulta invariante a rotaciones, traslaciones y escalado.



Figura 3-9: Algoritmo de extracción de contornos

El centroide es también el punto de referencia clave para el algoritmo de extracción

de características, que se divide en dos pasos (ver figura 3-10):

- En el primer paso se crean círculos, centrados en el centroide, y se cuenta en número de puntos de intersección de cada radio y la suma de todas las distancias entre esos puntos. Todos estos puntos se almacenan como el primer vector de características, correspondientes a los radios.
- En el segundo paso se utilizan los círculos creados pero, en esta ocasión, se cuentan los puntos de intersección con cada línea de contorno. Además, mientras se trazan las líneas de contorno, se detectan puntos característicos como terminaciones y bifurcaciones. Se almacena todo esto, junto con los puntos de intersección para cada contorno, en el segundo vector de características, correspondiente a la topología de los contornos.

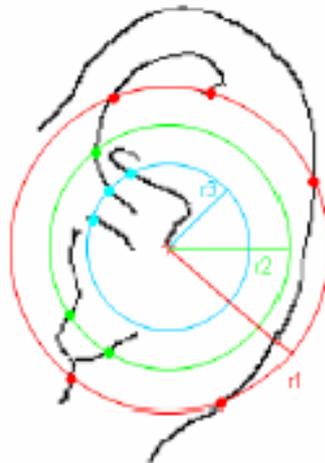


Figura 3-10: Representación simbólica del algoritmo para $Nr=3$

La clasificación se realiza basándose, simplemente, en la comparación entre los vectores de características de la imagen de entrada y todos los vectores de características de la base de datos.

Los resultados obtenidos eran buenos aunque, los mismos autores comentan que se podría mejorar el método incrementando los parámetros considerados en los vectores de características.

3.8. Técnica de descomposición de ondículas Haar

En la India, concretamente en el Departamento de Informática de IIT, en Kanpur, se ha adoptado una nueva aproximación para la extracción de características a partir de una imagen de oreja, usando una técnica de descomposición de ondículas Haar, propuesta por Sana y colaboradores en el año 2007 [32].

La descomposición de ondículas Haar se aplica a las imágenes de las orejas y, los coeficientes de las ondículas extraídos, representan el patrón de la oreja. Para la identificación de una imagen de prueba, se trata de encajar esta con n imágenes de control, usando un algoritmo de distancias de Hamming.

Este estudio incorpora, por primera vez, información de la biometría de la oreja que no sólo es geométrica, sino física, para mejorar el proceso de identificación. Esta otra información adicional consiste en datos antropométricos y antroposcópicos (por ejemplo estatura, peso y color de la piel del individuo), que se recogen en el momento de la inscripción o registro. Esta información se utiliza para filtrar, reduciendo la zona de búsqueda en las grandes bases de datos, mejorando, de esta forma, la velocidad y la eficiencia de la búsqueda en un sistema biométrico. Por ejemplo, si la forma de la oreja es redondeada en la imagen de prueba, la búsqueda se puede restringir sólo a los individuos con ese perfil, registrados en la base de datos. El método se implementó y probó con dos bases de datos de imágenes, con imágenes pertenecientes a 600 individuos de IIT Kanpur y 350 individuos de la Universidad de Saugor. Los resultados mostraron una precisión del 98.4% y 97.3%, para las bases de datos de IIT Kanpur y la Universidad de Saugor respectivamente.

Localización del contorno de la oreja

4.1. Introducción a los Contornos Activos

Desde que fueron introducidos por Kass y colaboradores [23] en 1987, los contornos activos o snakes han ganado popularidad y se han convertido en una herramienta eficaz en numerosas tareas relacionadas con el procesamiento y el análisis de imágenes como pueden ser, entre otras, la segmentación y el seguimiento de objetos móviles o su deformación.

Un contorno activo (comúnmente llamado *snake*) consiste en una curva elástica aproximada que, colocada sobre una imagen, empieza a deformarse a partir de una forma inicial con el fin de delimitar las regiones de interés en la escena. Esta deformación se produce mediante la aplicación de *fuerzas internas*, intrínsecas al snake y que controlan la suavidad de la curva, así como mediante *fuerzas externas*, procedentes de la imagen bajo tratamiento, que empujan al snake hacia las características salientes de la imagen. Esta habilidad para explotar el control mixto ascendente (datos de imágenes) y descendente (datos basados en conocimiento) convierte a los modelos deformables en una herramienta eficaz en múltiples tareas, como en el análisis de imágenes médicas,

donde la baja relación señal/ruido hace insuficientes los resultados obtenidos mediante técnicas clásicas.

4.1.1. Planteamiento clásico: Minimización de energía

Un *snake* colocado sobre una imagen empieza a deformarse a partir de una forma inicial con el fin de ajustarse a las características notables de la escena. Esta deformación se produce por atracción, hacia características destacables de la imagen, provocada por las fuerzas externas, que se ven contrarrestadas por las fuerzas internas, que intentan mantener la condición de suavidad en la forma de la curva. Dentro de este esquema de comportamiento los contornos activos evolucionan hacia una posición y forma determinados por un mínimo en la energía total definida como:

$$\varepsilon(u) = E_{int}(u) + E_{ext}(u)$$

donde E_{int} y E_{ext} son, respectivamente, la energía interna y la energía externa del *snake*. $u(s)$ representa el contorno como una proyección del dominio paramétrico unitario $s \in [0, 1]$ en el plano de la imagen \mathbb{R}^2 .

La energía interna de deformación en su expresión clásica [27], [28], viene dada por:

$$E_{int}(u) = \int_0^1 \left[\alpha(s) \left| \frac{du(s)}{ds} \right|^2 + \beta(s) \left| \frac{d^2u(s)}{ds^2} \right|^2 \right] ds$$

donde $u(s) = (x(s), y(s))$ representa la curva del *snake* y s la longitud de arco de la curva. Con esta expresión la energía interna se corresponde con la suma de la energía de membrana, función del estiramiento del *snake*, y la energía de deformación plana, que expresa la flexión. Las funciones paramétricas de elasticidad $\alpha(s)$ y $\beta(s)$ controlan la suavidad de la curva. Estas funciones son útiles para manipular el comportamiento físico y la continuidad local del modelo. Así, por ejemplo, un aumento en la magnitud de

$\alpha(s)$ da lugar a incrementos en la tensión de la curva con lo que se tiende a eliminar lazos o rizos mediante la reducción de la longitud del contorno. Por otra parte, el aumento de $\beta(s)$ incrementa la rigidez del *snake* haciéndolo más suave en su forma y menos flexible. En particular, haciendo $\alpha(s_0) = \beta(s_0) = 0$ se permite una discontinuidad en la posición, y haciendo únicamente $\beta(s_0) = 0$ se permite una discontinuidad tangente en s_0 .

La energía externa empuja al *snake* hacia extremos de intensidad, bordes u otras características de interés más complejas, dependiendo de la aplicación. Es conveniente pensar en las fuerzas externas que actúan sobre el contorno como derivadas a partir de un potencial P . De este modo la energía externa se evalúa como la integral del potencial a lo largo de la curva del *snake*

$$E_{ext}(u) = \int_0^1 P(u(s)) ds$$

La definición del potencial es, en general, dependiente de la aplicación. Expresiones típicas son las siguientes:

- $P_I = \pm \gamma I(u(s))$. Este potencial genera una fuerza que atrae a la curva hacia puntos de alta/baja intensidad [27].
- $P_G = \pm \delta |G_\sigma(u(s)) * I(u(s))|$, definido para provocar la atracción de la curva hacia puntos de gradiente bajo/alto. La operación $G_\sigma * I$ denota la imagen convolucionada con un filtro de suavización Gaussiano cuyo ancho característico controla la extensión espacial de la depresión atractiva de P [29].
- $P_E = \pm \eta e^{-d(u(s))^2}$, donde $d(u(s))$ representa la distancia entre el punto $u(s)$ y el punto más próximo dentro del mapa de propiedades de interés (bordes, valles, crestas, etc). Este potencial da lugar a estados energéticos mayores (menores) a medida que avanza el *snake* hacia la localización, previamente determinada, de características de interés en la imagen [28].

Las constantes γ, δ, η ponderan la influencia de cada uno de los términos anteriores en la expresión total de la energía externa.

Los modelos de *snakes* tradicionales no están exentos de limitaciones. Por una parte, debido a su naturaleza secuencial, estas técnicas requieren, en mayor o menor medida, un tiempo de computación alto, lo que puede representar un inconveniente en aplicaciones como la segmentación de objetos en movimiento o cualquier otra que requiera una alta velocidad de respuesta. Otra limitación importante está relacionada con las transformaciones topológicas. La mayoría de los algoritmos basados en modelos de contornos activos, debido a su naturaleza paramétrica, son incapaces de separar un único contorno en varios o unir varios en un único snake. Esto limita su campo de aplicación a tareas donde el número de objetos y localización aproximada son conocidos a priori. Con el objetivo de superar tales restricciones, surgen nuevas estrategias como los denominados Modelos Deformables Implícitos.

4.1.2. Modelos Deformables Implícitos

Una alternativa diferente al modelo paramétrico clásico la constituyen los Modelos Implícitos, introducidos independientemente por [24] y [25], en base al trabajo de Osher y Sethian [26]. Se trata de esquemas independientes de la topología, basados en el modelo de propagación de frentes de onda con velocidad dependiente de la curvatura. La propagación del frente de onda es visto como la evolución de un conjunto de nivel de alguna función definida implícitamente.

Sea $[C : [0, 1] \longrightarrow \mathbb{R}^2, s \rightarrow C(s)]$ un contorno inicial cerrado definido en el espacio euclídeo \mathbb{R}^2 y $C(s, t)$ la familia de curvas resultantes de su evolución en la dirección normal N con velocidad dependiente de la curvatura $F(k)$:

$$\left\{ \begin{array}{l} \frac{\partial C}{\partial t} = F(k)N \\ C(s, 0) = C_0(s) \end{array} \right\} \quad (4.1)$$

La implementación de la evolución de la curva puede abordarse desde la óptica de las técnicas clásicas de modelos deformables, aunque con sus correspondientes limitaciones a la hora de manejar posibles cambios de topología en el frente de onda. Estas limitaciones son superadas en los modelos implícitos mediante la definición del contorno activo como un conjunto de nivel cero de una función de orden superior Ψ definida usualmente como:

$$\Psi(x, t = 0) = \pm d$$

donde d representa la distancia desde $x \in \mathbb{R}^2$ a $C_0(s)$ con el signo dependiente de si se trata de un punto interior o exterior al contorno. En consecuencia se cumple que

$$C_0(s) = (x \mid \Psi(x, t = 0) = 0)$$

El objetivo es producir una ecuación para la función envolvente $\Psi(x, t)$ que contenga la curva $C(s, t)$ resultante del desplazamiento del contorno inicial, como el conjunto de nivel $\Psi = 0$. Por lo tanto, dado un punto $x(t)$ perteneciente al frente de onda $C(t)$ debe cumplirse:

$$\Psi(x(t), t) = 0$$

Aplicando la regla de la cadena

$$\frac{\partial \Psi}{\partial t} + \sum_{i=1}^2 \frac{\partial \Psi}{\partial x_i} \frac{\partial x_i}{\partial t} = 0 \tag{4.2}$$

donde x_i es la i -ésima componente de x . Dado que para cualquier conjunto de nivel el vector normal viene dado por $\frac{\nabla \Psi}{|\nabla \Psi|}$, podemos reescribir el segundo sumando de la ecuación 4.2 relacionándola con la ecuación 4.1

$$\sum_{i=1}^2 \frac{\partial \Psi}{\partial x_i} \frac{\partial x_i}{\partial t} = (\Psi_{x_1}, \Psi_{x_2}) \cdot \left(\frac{\partial x_1}{\partial t}, \frac{\partial x_2}{\partial t} \right) = F(k) |\nabla \Psi|$$

y obtener por tanto la ecuación de la evolución de Ψ ,

$$\frac{\partial \Psi}{\partial t} + F(k) |\nabla \Psi| = 0$$

Esta formulación para la propagación del frente de onda presenta importantes ventajas entre las que cabría destacar las siguientes:

- La función envolvente Ψ se mantiene bien definida siempre y cuando la velocidad de evolución $F(k)$ sea suave, aunque el frente de onda $C(s, t)$ asociado al conjunto de nivel cero puede cambiar de topología tal y como se ilustra en la figura 4-1.

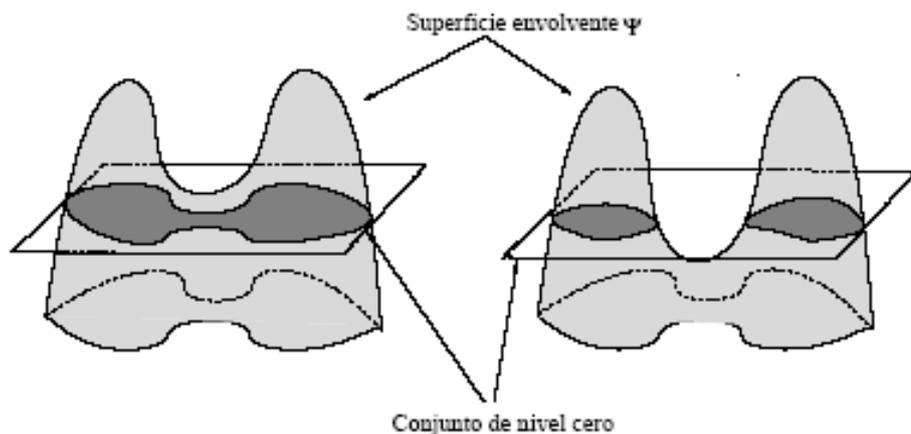


Figura 4-1: Ejemplo de transformación topológica en la formulación de conjuntos de nivel. Obsérvese como aún cuando se ha producido un cambio de topología en el conjunto de nivel cero, la superficie envolvente permanece bien definida.

- Precisamente el hecho de que permanezca bien definida durante su evolución permite realizar simulaciones numéricas de un modo sencillo en base a esquemas de diferencias finitas.
- Pueden estimarse diferentes propiedades geométricas de la curva directamente a partir de la función de conjuntos de nivel. El vector normal viene dado por $N = \frac{\nabla \Psi}{|\nabla \Psi|}$ y la curvatura de cada conjunto de nivel por $k = \nabla \cdot \frac{\nabla \Psi}{|\nabla \Psi|}$

- La formulación se puede extender directamente a frentes de dimensiones superiores.

Con vistas a la utilización de esta técnica en la segmentación de imágenes, es necesario definir una función de velocidad que dirija el frente de onda hacia los límites de las regiones de interés. Inicialmente se propuso una velocidad de propagación (deformación) dependiente de dos términos:

1. Un término basado en la regularidad y dependiente de la geometría del frente (curvatura local), F_I . Este término de difusión suaviza las regiones de elevada curvatura dando lugar a un efecto regularizador del frente similar al realizado por la energía de deformación interna (término de flexión) en los modelos deformables clásicos.
2. Un término de advección independiente de la geometría, F_A . Su efecto es similar a las fuerzas de inflación/deflación (dependiente de si el contorno inicial es exterior o interior a la estructura a delimitar) utilizada en los modelos paramétricos.

Ambos términos son multiplicados por una cantidad g dependiente de la imagen y que detiene al frente en los límites de la región de interés. Así, por ejemplo

$$g(x, y) = \frac{1}{1 + |\nabla G_\sigma * I(x, y)|} \quad (4.3)$$

toma valores próximos a cero en regiones de alto gradiente y valores próximos a la unidad en regiones homogéneas. En base a las propiedades de interés se pueden sintetizar criterios de parada más complejos que sitúen al contorno en los límites de los objetos a segmentar. La ecuación de movimiento del frente de onda sería entonces

$$\frac{\partial C}{\partial t} = g [F_A + F_I(k)] N$$

El principal inconveniente de la formulación clásica de modelos implícitos es que el contorno únicamente se detendrá cuando g sea nulo, lo cual sólo ocurre en bordes ideales. En los casos donde el gradiente toma diferentes valores a lo largo de los bordes de las regiones de interés (frecuente en las imágenes reales), g no se anulará y, por tanto, el frente continuará evolucionando a través de dichas discontinuidades. Esto obliga a iniciar el proceso con un alto grado de simetría en relación con la forma del objeto a delimitar, así como a diseñar criterios de parada más sofisticados que garanticen la detención del frente de onda en los límites de las estructuras de interés, lo cual generalmente no es una operación sencilla de abordar.

Con el propósito de superar esta limitación, Caselles y colaboradores propusieron una variante de los modelos implícitos conocida con el nombre de *contornos activos geodésicos* [30]. Esta incorpora una nueva componente de velocidad relacionada con la búsqueda de curvas geodésicas, o curvas de distancia mínima en un espacio de Riemann, cuya métrica se define a partir del contenido de la imagen, de tal modo que el estado de mínima energía se corresponda con la segmentación deseada. Como resultado, el contorno evoluciona hacia un mínimo del siguiente funcional de energía:

$$E [C(s)] = \int_0^1 g (|\nabla I (C(s))|) \left| \dot{C}(s) \right| ds$$

donde

- $\dot{C}(s)$ es la derivada parcial de la curva respecto del parámetro s
- $g(\cdot)$ se define con características similares a la función en 4.3.

Es decir, el mínimo de energía se corresponde con la curva de mínima longitud (i. e. geodésica) en relación a una nueva longitud de arco ponderada por la función $g(\cdot)$. Esta función objetivo se minimiza mediante la resolución de la ecuación de Euler-Lagrange asociada:

$$\frac{\partial C(t)}{\partial t} = g(I)kN - (\nabla g \cdot N)N$$

Por tanto, en el guiado del contorno intervienen dos términos:

$F_I = g(I)k$. Equivalente al término regularizador de los modelos deformables implícitos clásicos.

$F_G = \nabla g \cdot N$ que atrae la curva hacia los bordes de interés.

Al igual que ocurre con los modelos paramétricos y los modelos implícitos clásicos, se puede incluir una fuerza de advección F_A que empuje al *snake* y evite que quede anclado en bordes espurios, haciendo el modelo más insensible frente a la inicialización.

Finalmente, considerando el contorno activo como un conjunto de nivel de una función de una dimensión superior, obtenemos la siguiente ecuación de movimiento para la superficie envolvente:

$$\frac{\partial \Psi(t)}{\partial t} = g(I)k \nabla \Psi - (\nabla g \cdot \nabla \Psi) \nabla \Psi$$

4.2. Aproximación utilizada en este trabajo

La implementación que hemos utilizado de la técnica de contornos activos en 2D, propuesta en [30], está basada en la siguiente ecuación en derivadas parciales:

$$\frac{\partial u}{\partial t} = g_\sigma(I) \operatorname{div} \left(\frac{\nabla u}{\|\nabla u\|} \right) \|\nabla u\| + \lambda \nabla u \nabla g_\sigma(I) \quad (4.4)$$

$$u(0, x, y) = u_0(x, y)$$

donde $\lambda \geq 0$, la función $I(x, y)$ representa la intensidad de los niveles de gris de la imagen real original sobre la cual se va a utilizar el snake, $g_\sigma(I)$ se define como:

$$g_\sigma(I) = \frac{1}{\sqrt[2]{1 + \alpha \|\nabla G_\sigma * I\|^2}} \quad (4.5)$$

con $\alpha > 0$ y $G_\sigma * I$ es la convolución de I con el núcleo gaussiano

$$G_\sigma(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (4.6)$$

La imagen $u(t, x, y)$ representa la evolución del snake y $u(0; x, y)$ representa la función característica asociada al contorno inicial P_0 suministrado por el usuario. Hay diferentes modos de hacer eso. El más habitual consiste en que el usuario haga unos clicks sobre la imagen para dibujar un polígono inicial (ver figura 4-2), que denotamos por $P_0(s) = (x_0(s), y_0(s))$, cerca del borde buscado (en nuestro caso el contorno exterior e interior de la oreja).

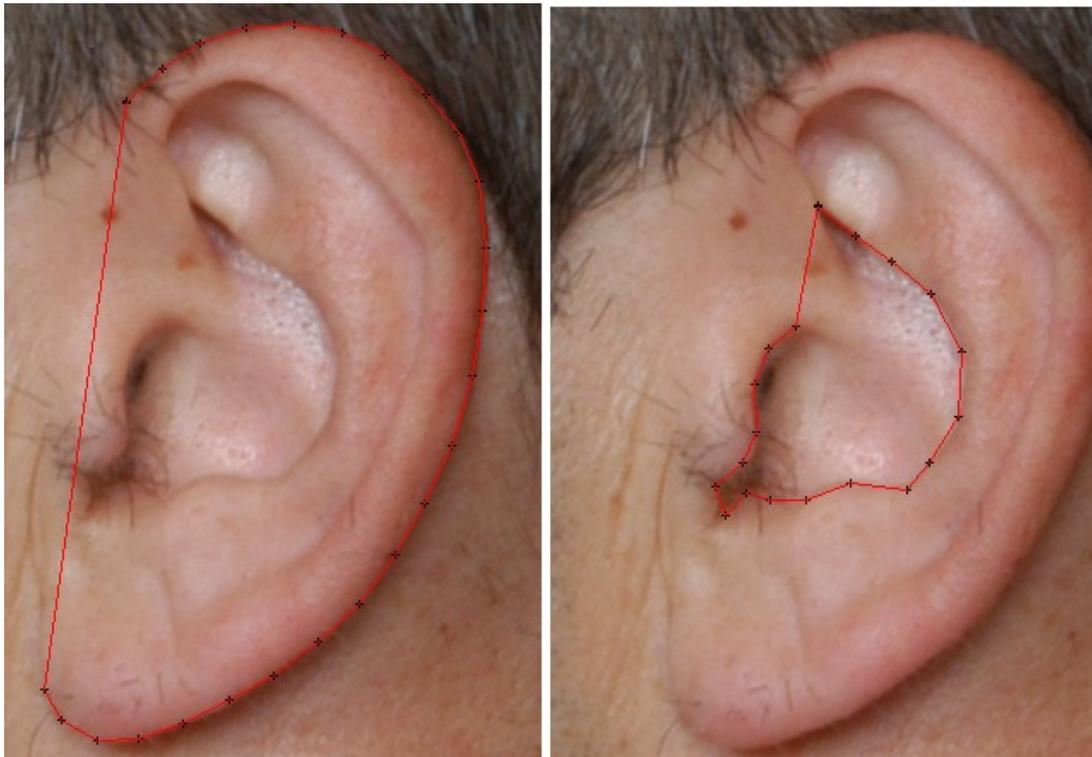


Figura 4-2: Polígonos de aproximación inicial a los contornos externos e interno de una oreja. El usuario señala los puntos (marcados en negro en la imagen) y el algoritmo trabaja con el polígono cerrado (marcado en rojo), que resulta de unir todos los puntos marcados, en el orden en que se marcaron, y unir el primer y último punto para cerrarlo.

A partir de $P_0(s)$ podemos definir $u(0, x, y)$ como una superficie tal que $P_0(s)$ corresponda a una curva de nivel de $u(0, x, y)$. Una curva de nivel se define como la frontera de un conjunto de nivel es decir

$$\text{curva de nivel} = P_0(s) \equiv \partial \{(x, y) : u(0, x, y) > \lambda\}$$

Por ejemplo, podemos definir $u(0, x, y)$ como una superficie tal que

$$u(0, x, y) = \begin{cases} \text{distancia}((x, y), P_0(s)) & \text{si } (x, y) \text{ está en el interior de } P_0(s) \\ -\text{distancia}((x, y), P_0(s)) & \text{si } (x, y) \text{ está en el exterior de } P_0(s) \end{cases} \quad (4.7)$$

en ese caso $P_0(s)$ viene representado por la curva de nivel asociada a $\lambda = 0$. Otra forma más sencilla definir $u(0, x, y)$ es utilizando una función característica, es decir:

$$u(0, x, y) = \begin{cases} l_1 & \text{si } (x, y) \text{ está en el interior de } P_0(s) \\ l_2 & \text{si } (x, y) \text{ está en el exterior de } P_0(s) \end{cases} \quad (4.8)$$

siendo l_1 y l_2 los niveles de intensidad donde el snake está incluido. En este caso, $P_0(s)$ viene dado por la curva de nivel asociado a cualquier λ que verifique $0 < \lambda < 1$. Para cada $t > 0$, el snake viene definido por la curva de nivel de la función $u(t, x, y)$ asociada al nivel λ correspondiente.

$$\text{curva de nivel} = P_t(s) \equiv \partial \left\{ (x, y) : u(t, x, y) > \frac{l_1 + l_2}{2} \right\}$$

En particular $P_\infty(s) \equiv \text{Lim}_{t \rightarrow \infty} P_t(s)$ representa el contorno de salida devuelto por la ecuación.

En el algoritmo de implementación utilizamos una imagen (ver figura 4-3), que llamamos *image_level_set*, que contiene el valor de la intensidad del nivel de gris de la imagen de conjuntos de nivel, en precisión flotante, dado por $\lim_{t \rightarrow \infty} u^{\sigma_1}(t, x, y)$. Es un parámetro de salida.

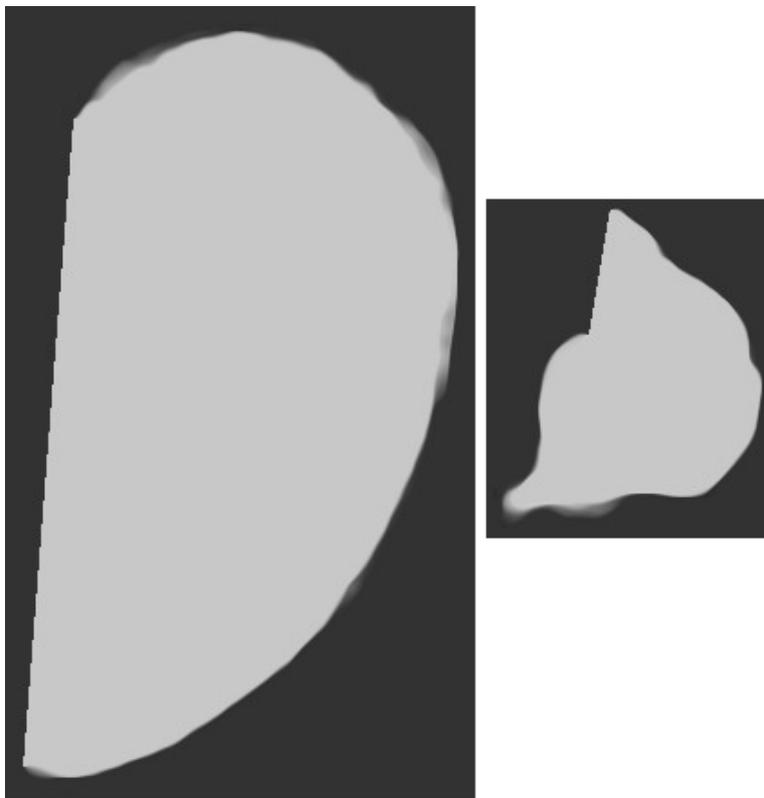


Figura 4-3: *Imagen del valor de intensidad del nivel de gris, de la imagen de conjuntos de nivel.*

A partir de esta imagen podemos calcular fácilmente el contorno final de salida, C_∞ , con precisión de pixel, usando, por ejemplo, como criterio, el hecho de que un punto (x, y) en la imagen, pertenece al contorno final C_∞ , si $image_level_set[y * width + x] < (l_1 + l_2)/2$ y en los 4 puntos vecinos usuales de (x, y) hay uno (x_{n_1}, y_{n_1}) donde $image_level_set[y_{n_1} * width + x_{n_1}] < (l_1 + l_2)/2$ y otro (x_{n_2}, y_{n_2}) donde $image_level_set[y_{n_2} * width + x_{n_2}] > (l_1 + l_2)/2$. Como los valores de l_1 y l_2 son irrelevantes, en el algoritmo los hemos fijado a $l_1 = 200$ y $l_2 = 50$, siendo, por tanto, $(l_1 + l_2)/2 = 125$.

En la ecuación diferencial 4.4, podemos observar dos partes que actúan simultáneamente y compiten entre sí. En primer lugar analizaremos la segunda parte, cuya función es atraer el snake actual hacia los contornos de la imagen. Posteriormente veremos cómo actúa la primera parte regularizando el contorno.

4.2.1. Atracción del snake hacia los bordes

La segunda parte de la ecuación viene definida por:

$$\frac{\partial u}{\partial t} = \lambda \nabla u \nabla g_\sigma(I)$$

Si denotamos $\nabla g_\sigma(I) = (a(x, y), b(x, y))$. La ecuación anterior se escribe como:

$$\frac{\partial u}{\partial t} = a(x, y) \frac{\partial u}{\partial x} + b(x, y) \frac{\partial u}{\partial y}$$

Esta ecuación se denomina de transporte porque cada punto tiende a desplazarse en la dirección $-(a(x, y), b(x, y))$. Efectivamente, si $u_0(x, y)$ es el dato inicial para la ecuación, y consideramos que es localmente constante, es decir, $(a(x, y), b(x, y)) = (a, b)$ es inmediato demostrar que $u(t, x, y) = u_0(x + ta, y + tb)$ es la solución de la ecuación diferencial, lo cual representa un desplazamiento en la dirección $-(a, b)$.

Para ilustrar este efecto, supongamos que tenemos una región de nivel alto sobre un fondo de nivel bajo. Por tanto, el gradiente señalará hacia el interior de la región (ver

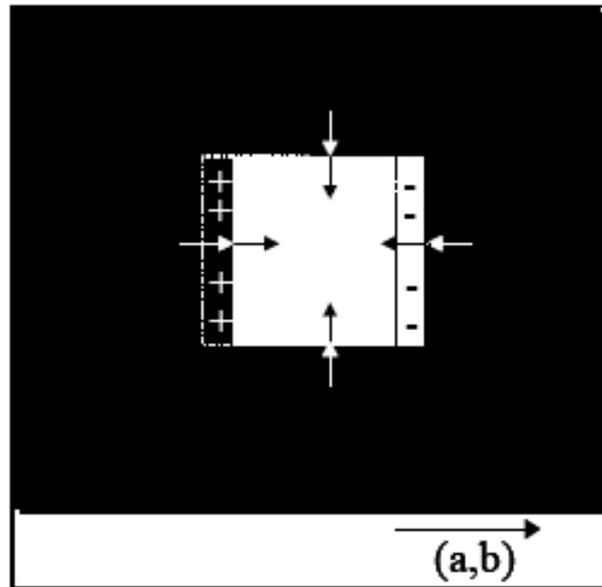


Figura 4-4: Efecto de la ecuación de transporte: el cuadrado se desplaza en sentido contrario al indicado por (a, b)

4-4). En los puntos donde el gradiente tiene la dirección de (a, b) , que son los situados en el lado opuesto al que indica (a, b) a partir del centro de la región (en el ejemplo de la figura, a la izquierda del cuadrado), el producto será positivo y por tanto tenderán a aumentar su valor, con lo que la región crecerá por esa zona. En los puntos donde el gradiente tiene la dirección opuesta a (a, b) (en el ejemplo de la figura, a la derecha del cuadrado), el producto será negativo y tenderán a disminuir su valor. Por lo tanto, la región se reducirá por esa zona. Cuando el gradiente es perpendicular a (a, b) , el resultado será 0 y no se producirán cambios. Como resultado de estos efectos, la región se habrá desplazado en la dirección opuesta a (a, b) .

En nuestro caso, el vector (a, b) ha sido sustituido por $\nabla g_\sigma(I)$, donde

$$g_\sigma(I) = \frac{1}{\sqrt[2]{1 + \alpha \|\nabla I_\sigma\|^2}}$$

siendo I_σ la imagen original convolucionada con una Gaussiana de desviación estándar

σ . En las figuras 4-5 vemos las imágenes I_σ correspondientes a los contornos externo e interno de la oreja y en la figura 4-6 las imágenes $g_\sigma(I)$ correspondientes.

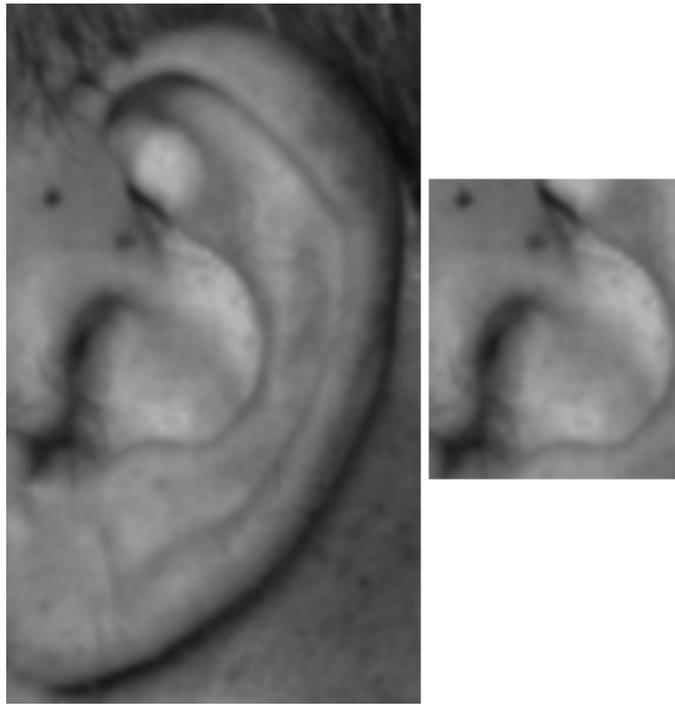


Figura 4-5: Imágenes I_σ . Resultado de convolucionar la imagen original de los contornos externo e interno de la oreja, con una gaussiana de desviación estándar σ .

Cuanto mayor sea ∇I_σ , menor será $g_\sigma(I)$. Por lo tanto, el menor valor de $g_\sigma(I)$, se alcanzará en los puntos de borde de I y, a medida que nos alejamos del borde, va aumentando, a un lado y a otro, como si fuese un valle cuyo cauce está determinado por los bordes de I_σ . De esta forma, ∇I_σ es un vector orientado en sentido contrario a donde se encuentra el borde, tanto si estamos a un lado del mismo, como si estamos en el opuesto (ver 4-7).

Dado que la ecuación de transporte produce un desplazamiento en sentido opuesto

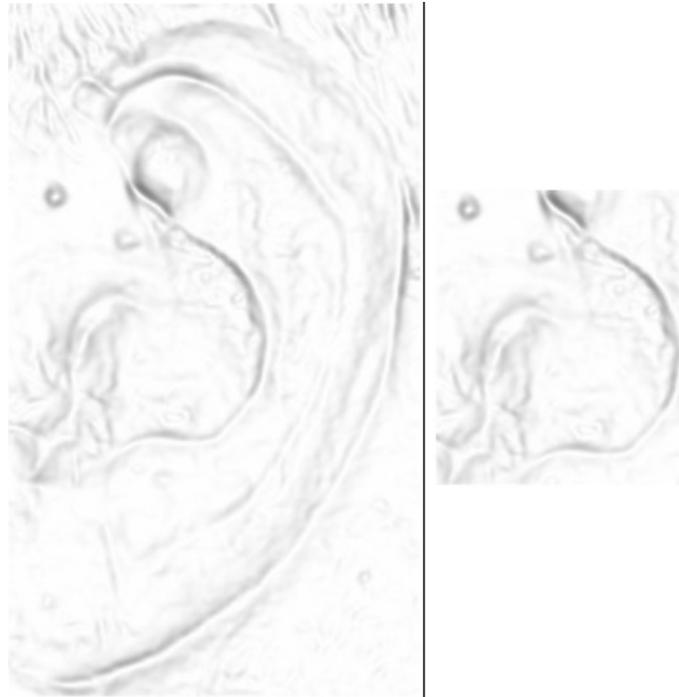


Figura 4-6: *Imágenes $g_{\sigma}(I)$ correspondientes a las imágenes I_{σ} anteriores.*

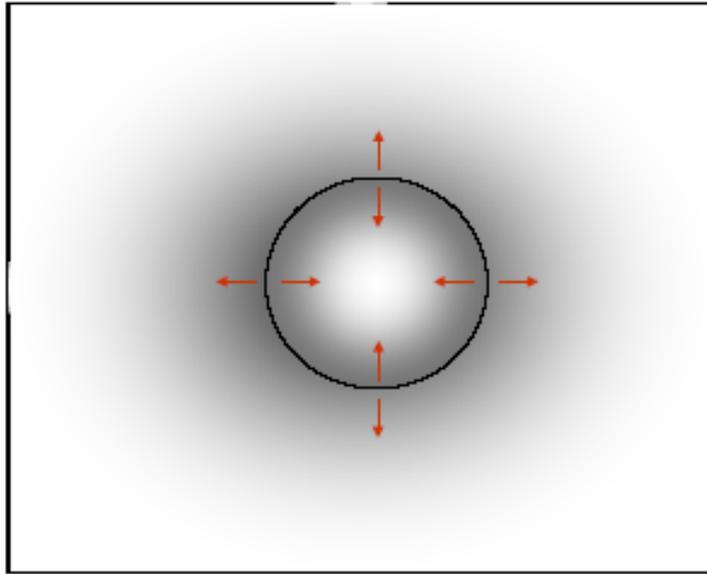


Figura 4-7: Valores de $g_\sigma(I)$ y dirección de ∇I_σ . Los valores mínimos de $g_\sigma(I)$ se producen en los bordes y tanto hacia el interior como hacia el exterior aumenta su valor; por tanto, el gradiente apunta en sentido contrario al borde.

al que indica el vector utilizado, en este caso, la ecuación

$$\frac{\partial u}{\partial t} = \nabla u \nabla g_\sigma(I)$$

producirá un desplazamiento de u en sentido opuesto al indicado por $\nabla g_\sigma(I)$, es decir, producirá un desplazamiento hacia los bordes de I (ver 4-8). El parámetro λ_s regula con qué fuerza será atraído hacia el borde el snake.

4.2.2. Regularización del contorno

La primera parte de la ecuación está definida por:

$$\frac{\partial u}{\partial t} = g_\sigma(I) \operatorname{div} \left(\frac{\nabla u}{\|\nabla u\|} \right) \|\nabla u\|$$

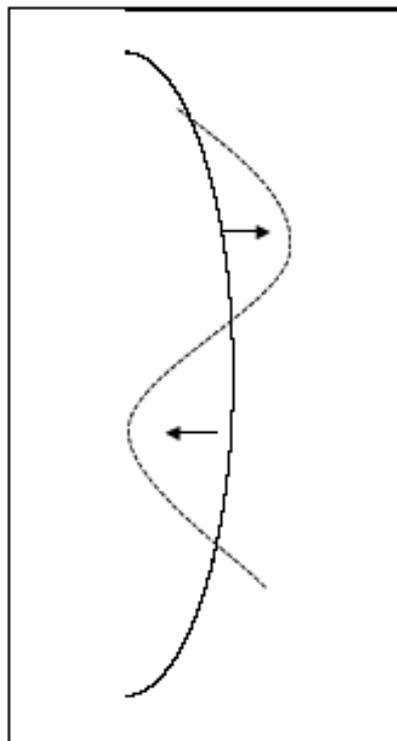


Figura 4-8: *Atracción del contorno hacia los bordes. La línea continua representa el snake y la línea discontinua el borde en la imagen original*

Esta basada en la ecuación de curvatura media:

$$\frac{\partial u}{\partial t} = \operatorname{div} \left(\frac{\nabla u}{\|\nabla u\|} \right) \|\nabla u\|$$

Si desarrollamos $\operatorname{div} \left(\frac{\nabla u}{\|\nabla u\|} \right)$, obtenemos:

$$\operatorname{div} \left(\frac{u_x}{\|\nabla u\|}, \frac{u_y}{\|\nabla u\|} \right) = \frac{\partial \left(\frac{u_x}{\|\nabla u\|} \right)}{\partial x} + \frac{\partial \left(\frac{u_y}{\|\nabla u\|} \right)}{\partial y}$$

$$\operatorname{div} \left(\frac{\nabla u}{\|\nabla u\|} \right) = \frac{(u_x)^2 u_{yy} - 2 u_x u_y u_{xy} + (u_y)^2 u_{xx}}{((u_x)^2 + (u_y)^2)^{\frac{3}{2}}}$$

que se corresponde con la curvatura de la función u , teniendo en cuenta que la curvatura k de una curva $C = (x(s), y(s))$ viene dada por:

$$k = \frac{y''x' - x''y'}{((x')^2 + (y')^2)^{\frac{3}{2}}}$$

Nuestra curva u depende de s (parámetro usado para definirla paramétricamente) y de t , que define la evolución en el tiempo. En cada instante de tiempo, los puntos de la curva se caracterizan por tener el mismo valor:

$$u(t, x(t, s), y(t, s)) = L \quad \forall s, t$$

Por lo tanto, si derivamos con respecto al tiempo:

$$\frac{du(t, x(t, s), y(t, s))}{dt} = u_t + u_x x_t + u_y y_t = 0 \tag{4.9}$$

Si queremos regularizar la función u , nos interesa que en las zonas convexas evolucione hacia el interior y en las zonas cóncavas evolucione hacia el exterior. Es decir, en la dirección de la normal y en sentido contrario al que indica la curvatura. Si la región tiene un nivel más alto que el fondo, el gradiente señalará hacia el interior de la región;

en las regiones cóncavas, la curvatura será positiva y evolucionará hacia el exterior (sentido contrario al indicado por la normal); en las zonas convexas, la curvatura será negativa y evolucionará hacia el interior (el mismo sentido que la normal):

Por lo tanto, nuestro objetivo es que

$$Ct = -k \vec{N}$$

o lo que es lo mismo:

$$\begin{pmatrix} x_t \\ y_t \end{pmatrix} = -k \frac{\nabla u}{\|\nabla u\|} = \begin{pmatrix} -k \frac{u_x}{\|\nabla u\|} \\ -k \frac{u_y}{\|\nabla u\|} \end{pmatrix}$$

Sustituyendo en la ecuación 4.9:

$$u_t - k \frac{u_x}{\|\nabla u\|} u_x - k \frac{u_y}{\|\nabla u\|} u_y = 0$$

$$u_t = k \frac{(u_x)^2 + (u_y)^2}{\|\nabla u\|} = k \|\nabla u\|$$

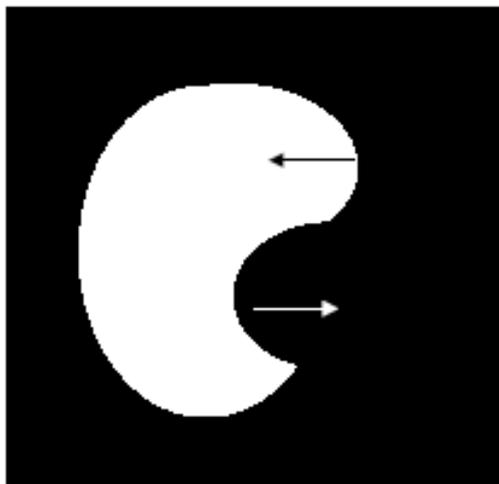


Figura 4-9: Efecto de la ecuación de curvatura media: en la zonas convexas atrae el contorno hacia el interior y en las cóncavas hacia el exterior.

Esta parte de la ecuación hará que la región tienda a redondearse y reducirse. Si sustituimos k por $\operatorname{div} \left(\frac{\nabla u}{\|\nabla u\|} \right)$, obtenemos la ecuación

$$\frac{\partial u}{\partial t} = \operatorname{div} \left(\frac{\nabla u}{\|\nabla u\|} \right) \|\nabla u\|$$

Al multiplicar por $g_\sigma(I)$ se consigue que el efecto sea tanto mayor cuanto más lejos estemos del borde y se detenga cuando alcancemos uno.

4.2.3. Detalles de la Implementación

El principal inconveniente que nos encontramos en nuestro caso particular de aplicación es que, el contorno de la oreja, que intentamos aproximar, es un contorno abierto, mientras que toda la técnica de contornos activos que hemos explicado trabaja con un contorno inicial, que denotamos por $P_0(s) = (x_0(s), y_0(s))$, cerca del borde buscado que, en todos los casos, es una curva cerrada. Por tanto, hemos tenido que adaptar el modelo a esta situación particular, para lo cual, en nuestro algoritmo, se añade un último segmento que cierra el polígono que se toma como aproximación inicial del contorno de la oreja. Este polígono se obtiene uniendo los puntos que el usuario marca en la imagen original, cerca del borde de la oreja y, el último segmento añadido, une el primer y el último punto marcados por el usuario, cerrando de esta manera el polígono inicial. Este último segmento que se añade al polígono inicial, es forzado a permanecer invariante a lo largo de las iteraciones del algoritmo consiguiendo, de esta manera, que la evolución del snake se adapte al contorno de la oreja sin modificarse la parte de la misma unida a la cabeza.

La implementación usual del tipo de ecuaciones diferenciales 4.4 es la siguiente: En primer lugar se discretizan todos los operadores diferenciales que aparecen en la ecuación utilizando máscaras 3x3, de tal manera que se implementa el esquema explí-

cito:

$$u_{i,j}^{n+1} = u_{i,j}^n + \delta t \left(g_\sigma(I_{i,j}) \operatorname{div} \left(\frac{\nabla u_{i,j}^n}{\|\nabla u_{i,j}^n\|} \right) \|\nabla u_{i,j}^n\| + \lambda \nabla u_{i,j}^n \nabla g_\sigma(I_{i,j}) \right) \quad (4.10)$$

El esquema usual con más precisión para seguir el snake se basa en el siguiente bucle iterativo en n :

1. A partir del polinomio $P_n(s)$ se calcula la superficie $u_{i,j}^n$ dada por la función distancia (vease 4.7). Esto no es necesario hacerlo sobre toda la imagen, basta con hacerlo en una franja alrededor de la posición de $P_n(s)$
2. Se da una iteración del esquema numérico 4.10 para calcular $u_{i,j}^{n+1}$
3. Se calcula $P_{n+1}(s)$ como la curva de nivel de $u_{i,j}^{n+1}$ asociada al λ , correspondiente.

Por supuesto no pretendemos que esta sea considerada la mejor discretización del modelo 4.4 ni de los algoritmos de los snakes en general. En nuestra opinión la principal desventaja del algoritmo radica en que, si en alguna parte del contorno final, el gradiente no es suficientemente grande, el algoritmo puede no trabajar apropiadamente. Otra limitación del algoritmo es que, debido a la naturaleza geométrica de la ecuación 4.4, el algoritmo alcanza más fácilmente la parte convexa del contorno final que la parte cóncava. Para ayudar al algoritmo a ajustarse al contorno final correcto, una buena idea es definir el polígono inicial tomando puntos fuera del contorno final en la parte convexa y puntos dentro del contorno final en la parte cóncava.

La implementación se ha realizado con dos funciones en lenguaje C. La primera de ellas, llamada *ami_snake_segmento_fijo*, implementa la evolución del snake comentada, manteniendo fijo, a lo largo de las iteraciones del esquema 4.10, el segmento final del polígono de partida. Para agilizar el algoritmo, éste se ejecuta sólo en la subventana de la imagen que incluye el polígono. Un parámetro de entrada indica la distancia extra que añadimos a los límites del polígono para definir los bordes de la subventana

de trabajo. Generalmente un valor de 10 de este parámetro es suficiente. Como trabajamos con una subventana de la imagen original, debemos mantener información de la posición relativa de la subventana respecto de la imagen total. Dos parámetros de la función (xd, yd) tienen esta utilidad, de manera que, las coordenadas en pixels del punto (x,y) en la subventana, corresponde a $(xd+x; yd+y)$ en la imagen original.

La segunda de las funciones, llamada *ami_extraer_contorno_segmento_final*, es la que finalmente extrae el conjunto de puntos correspondientes al contorno (externo o interno), a partir de la imagen de conjuntos de nivel final devuelta por la función anterior. En la figura 4-10 vemos los contornos obtenidos partiendo de los polígonos marcados en la figura 4-2.

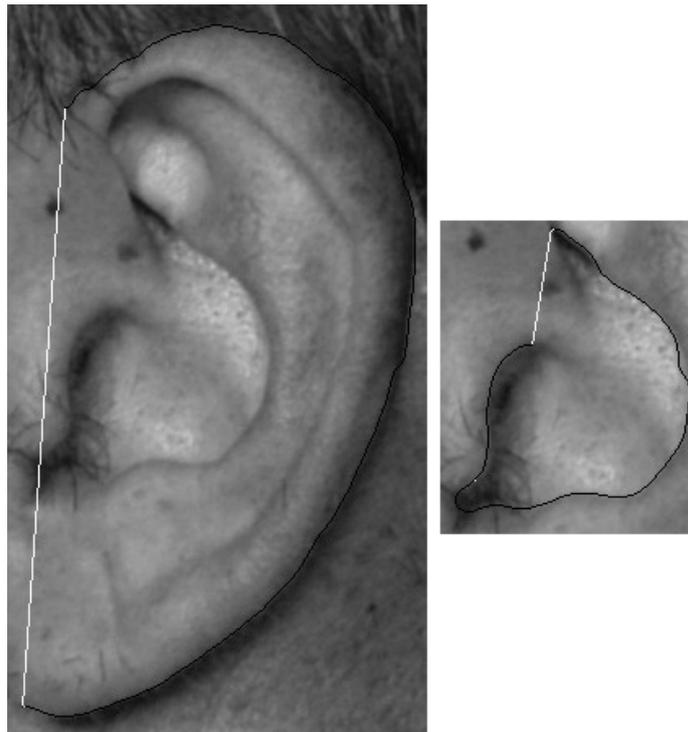


Figura 4-10: *Imagen de los conjuntos de puntos (dibujados en negro) obtenidos como contornos externo e interno de la oreja de la imagen, después de aplicar el algoritmo. El segmento dibujado en blanco es el segmento del polígono inicial que se mantiene fijo a lo largo de las iteraciones del algoritmo.*

Modelo de ajuste del lóbulo exterior de la oreja utilizando una elipse deformada

Los ovoides son curvas cerradas, planas y convexas, de la misma naturaleza que los óvalos o elipses y, por lo tanto, tienen sus mismas propiedades. Sin embargo, hay una diferencia importante: así como los óvalos son simétricos respecto a sus dos ejes, los ovoides son más estrechos en uno de sus extremos, y sólo son simétricos respecto a su eje mayor, lo que les confiere su aspecto característico, parecido a un huevo de gallina.

La forma de un ovoide se puede obtener cambiando ligeramente la forma del óvalo. Multiplicando el término en y por un término $t(x)$ adecuado, de manera que la y sea mayor en la parte derecha del eje x y más pequeña en la izquierda (o al revés), manteniéndose el mismo valor de y en $x = 0$.

El ovoide que utilizamos para ajustar el contorno de la oreja, se obtiene, de esta forma, por deformación de una elipse. El modelo paramétrico usual para una elipse es:

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} x_0 + a \cos t \\ y_0 + b \sin t \end{pmatrix}$$

Para ajustarla mejor al contorno de una oreja, le añadimos a la elipse un factor de distorsión dado por un nuevo parámetro γ , obteniendo entonces el modelo de elipse deformada:

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} x_0 + a \cos t \\ y_0 + b(1 + \gamma \cos t) \sin t \end{pmatrix}$$

El factor γ , como se puede apreciar en la figuras 5-1, 5-2, 5-3, provoca una distorsión en el eje vertical de la elipse y permite adaptar mejor su forma a geometrías más complejas como las que se presentan en la forma de una oreja humana.

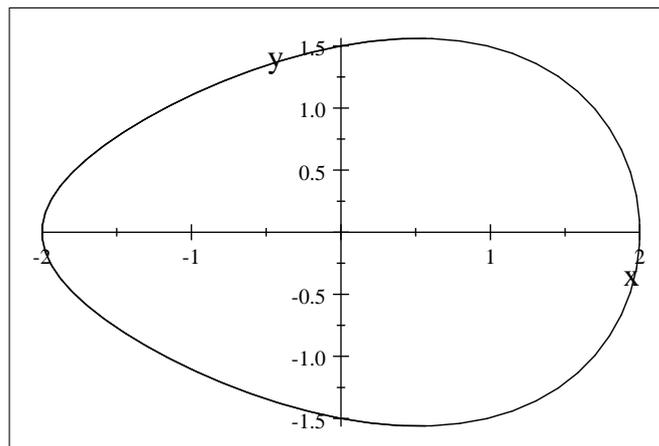


Figura 5-1: *Imagen de ovoide sin rotación, con valores de parámetros $(x_0, y_0) = (0, 0)$, $a = 2$, $b = 1,5$, $\alpha = 0$ y $\gamma = 0,3$. Su ecuación es $(2 \cos t, 1,5(1 + 0,3 \cos t) \sin t)$*

El modelo general también incluye una rotación respecto al centro de la elipse que vendría dada por un ángulo α y su correspondiente matriz de rotación

$$R(\alpha) = \begin{pmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{pmatrix}$$

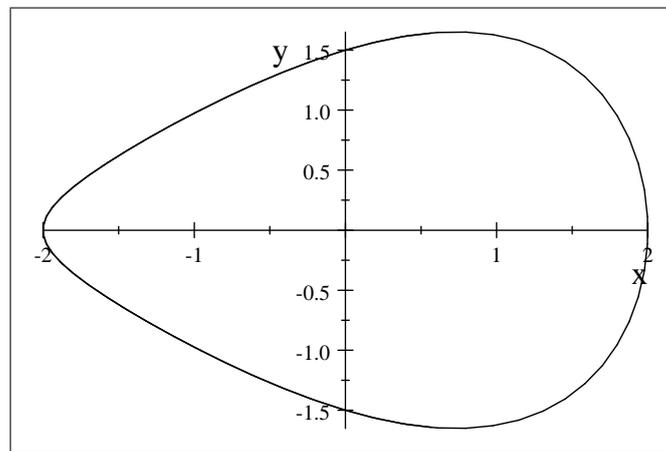


Figura 5-2: Imagen de ovoide. Todos los parámetros son los mismos que los de la figura anterior excepto $\gamma = 0,5$. Su ecuación es $(2 \cos t, 1,5(1 + 0,5 \cos t) \sin t)$

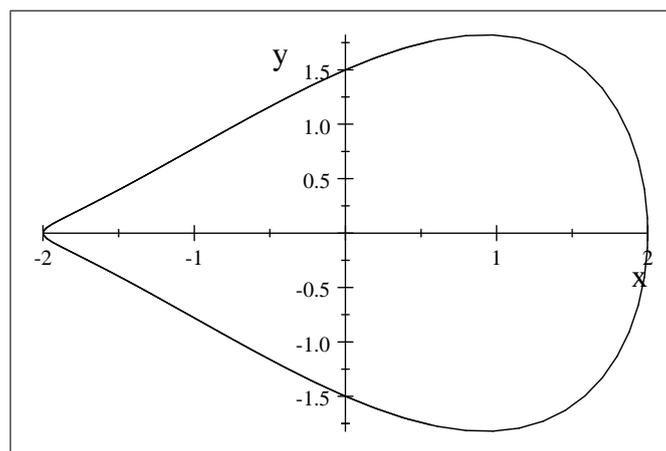


Figura 5-3: Imagen de ovoide. Todos los parámetros son los mismos que los de la figura anterior excepto $\gamma = 0,8$. Su ecuación es $(2 \cos t, 1,5(1 + 0,8 \cos t) \sin t)$

dicha matriz de rotación actúa de la siguiente manera

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} + R(\alpha) \begin{pmatrix} x(t) - x_0 \\ y(t) - y_0 \end{pmatrix}$$

La ecuación paramétrica del ovoide es:

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} x_0 + \cos \alpha a \cos t + \sin \alpha b(1 + \gamma \cos t) \sin t \\ y_0 - \sin \alpha a \cos t + \cos \alpha b(1 + \gamma \cos t) \sin t \end{pmatrix}$$

que tiene 6 grados de libertad: (x_0, y_0) que representan las coordenadas del centro del ovoide; a, b , que representan los parámetros usuales de los tamaños de los ejes de la elipse; el ángulo de rotación u orientación α , y la distorsión γ , que representa la deformación del ovoide respecto a la elipse. $\gamma = 0$ corresponde a una elipse, cuanto mayor es γ , más fuerte es la deformación de la elipse que tenemos.

En la figura 5-4, a la derecha, podemos observar la forma del ovoide, y podemos apreciar que se ajusta bastante bien a la forma de la oreja.

La ventaja de este modelo de ovoide que proponemos es que es bastante simple (tiene un sólo parámetro más que una elipse), puede ajustarse muy bien a la forma de una oreja humana y algunos de sus parámetros pueden usarse para comparar dos orejas. Por ejemplo, los parámetros $(\frac{b}{a}, \gamma)$ son invariantes euclídeos de la oreja y se pueden usar para comparar orejas independientemente de su localización y tamaño. Por otro lado, los parámetros del ovoide se pueden usar para alinear los contornos de dos orejas para compararlos de una forma más precisa.

En la práctica, este modelo de ovoide no nos sería muy útil si no dispusiéramos de un algoritmo que nos permitiera estimar los parámetros del ovoide, a partir del contorno de la oreja. Por ello, hemos desarrollado algunos algoritmos que estiman, de forma automática, los parámetros del ovoide que mejor se ajusta a un contorno de oreja dado, usando un criterio de distancia euclídea, y que pasamos a detallar.

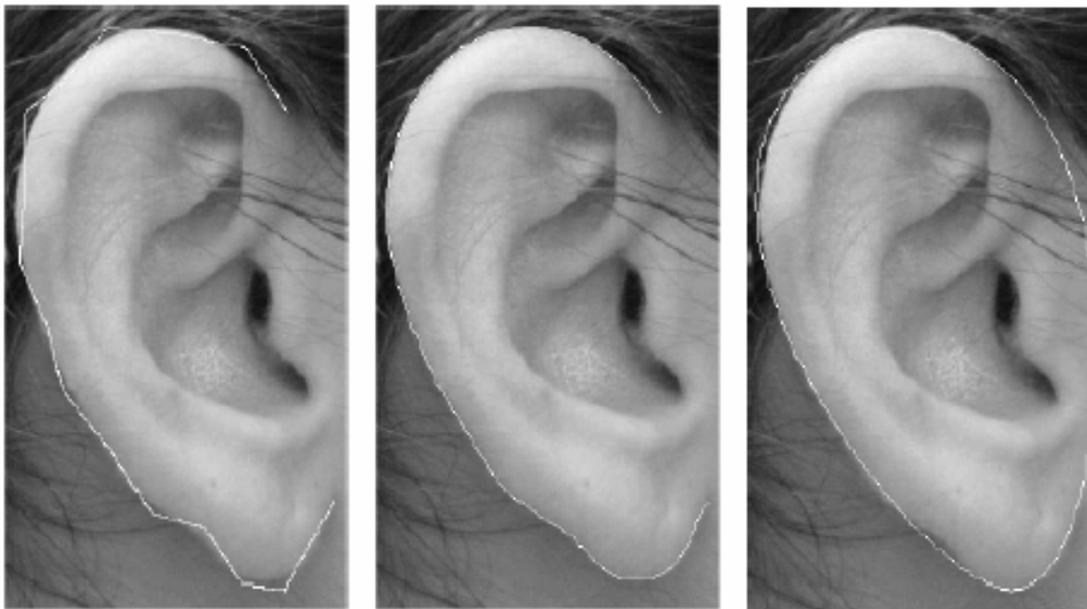


Figura 5-4: *Izquierda: imagen original de la oreja con el polígono de aproximación al contorno inicial dibujado en blanco. Centro: mejora de la aproximación del contorno (en blanco) usando los snakes. Derecha: ovoide (en blanco) estimado a partir del contorno de la oreja.*

En forma implícita y sin rotación, la ecuación de la elipse deformada es

$$\frac{x^2}{a^2} \left(1 + \gamma \frac{x}{a}\right)^2 + \frac{y^2}{b^2} = \left(1 + \gamma \frac{x}{a}\right)^2$$

despejando y nos queda

$$y = \pm b \left(1 + \gamma \frac{x}{a}\right) \sqrt{1 - \frac{x^2}{a^2}}$$

Observamos que los valores admisibles para el parámetro γ son $-1 \leq \gamma \leq 1$, pues de otro modo la curva generada dejaría de ser una curva de Jordan y se interceptaría a sí misma. La función $y = f(x)$ anterior admite una primitiva que viene dada por

$$\int b \left(1 + \gamma \frac{x}{a}\right) \sqrt{1 - \frac{x^2}{a^2}} dx = b \left(\frac{a}{2} \arcsin \frac{x}{a} + \frac{1}{2} x \sqrt{1 - \left(\frac{x}{a}\right)^2} \right) + \gamma b a \left(\frac{1}{3} \frac{x^2}{a^2} \sqrt{1 - \frac{x^2}{a^2}} - \frac{1}{3} \sqrt{1 - \frac{x^2}{a^2}} \right)$$

De esta integral podemos sacar algunas conclusiones interesantes:

- La primera de ellas es que el área de la elipse y la elipse deformada es la misma para cualquier valor del parámetro $-1 \leq \gamma \leq 1$. Es decir

$$AREA = 2 \int_{-a}^a b \left(1 + \gamma \frac{x}{a}\right) \sqrt{1 - \frac{x^2}{a^2}} dx = \pi ab$$

- La segunda conclusión es que área encerrada por la curva en el semiplano $x > 0$ viene dada por

$$AREA_{\{x>0\}} = 2 \int_0^a b \left(1 + \gamma \frac{x}{a}\right) \sqrt{1 - \frac{x^2}{a^2}} dx = \frac{\pi ab}{2} + \frac{2}{3} ab \gamma$$

Estas relaciones nos permiten diseñar fácilmente un algoritmo para extraer los diferentes parámetros de la elipse deformada que son 6 (el centro (x_0, y_0) , los semiejes a, b , el ángulo de rotación α , y la distorsión γ), a partir de una superficie en el plano

S , cuya geometría se ajusta a una elipse deformada. Concretamente, el algoritmo se puede desglosar en las siguientes etapas:

1. Calculamos el área de la forma (m_{00}) y su centroide (m_{10}, m_{01}) utilizando los momentos

$$m_{00} = \int_S dx dy \quad m_{10} = \frac{1}{m_{00}} \int_S x dx dy \quad m_{01} = \frac{1}{m_{00}} \int_S y dx dy$$

donde S representa la superficie encerrada por el contorno de la oreja extraído del modelo de snakes antes comentado. Nótese que el centroide de la forma no coincide con el centro de la elipse deformada, pues dicho centroide estará desplazado en la dirección del eje de simetría de la elipse.

2. Construimos la matriz de momentos centrados de orden dos $M_2 = \begin{pmatrix} m_{20} & m_{11} \\ m_{11} & m_{02} \end{pmatrix}$ dados por:

$$m_{20} = \int_S (x - m_{10})^2 dx dy \quad m_{02} = \int_S (y - m_{01})^2 dx dy$$

$$m_{11} = \int_S (x - m_{10})(y - m_{01}) dx dy$$

El autovector $v_{\text{máx}} = (x_{\text{máx}}, y_{\text{máx}})$ de autovalor mayor de la matriz M_2 determina la dirección principal de simetría de la elipse deformada. Por tanto, la orientación α de la elipse deformada se extrae simplemente a través de la relación:

$$\tan(\alpha) = \frac{y_{\text{máx}}}{x_{\text{máx}}}$$

3. A continuación proyectamos todos los puntos de la forma S sobre la recta que pasa por el centroide y tiene como vector director $v_{\text{máx}}$. Si denominamos $p_{\text{máx}}$ el punto proyectado que dista más del centroide en la dirección de $v_{\text{máx}}$, y por $p_{\text{mín}}$ el punto proyectado que dista más del centroide en la dirección opuesta $-v_{\text{máx}}$,

entonces podemos extraer el centro (x_0, y_0) y semieje a de la elipse deformada haciendo:

$$(x_0, y_0) = \frac{p_{\text{máx}} + p_{\text{mín}}}{2} \qquad a = \frac{\|p_{\text{máx}} - p_{\text{mín}}\|}{2}$$

4. Para obtener b (el semieje menor de la elipse deformada) basta con tener en cuenta la relación $m_{00} = \pi ab$, que viene dada por el hecho de que el área de la elipse deformada y el área del ovoide son iguales. Finalmente, para obtener la distorsión γ , calculamos la superficie A_+ de la elipse deformada que queda en uno de los semiplanos generado por la recta que pasa por el centro de la elipse deformada y tiene como vector perpendicular la dirección principal $v_{\text{máx}}$. Obtendremos γ a partir de la relación:

$$A_+ = \frac{\pi ab}{2} + \frac{2}{3} ab\gamma$$

5. Finalmente, para mejorar los parámetros del ovoide minimizamos la distancia euclídea:

$$E(x_0, y_0, \alpha, a, b, \gamma) = \sum_{i=0}^N \text{dist}(\text{Ovoid}, (x^i, y^i))$$

donde (x^i, y^i) representan los puntos del contorno de la oreja. Minimizamos tal distancia euclídea usando un algoritmo de gradiente descendente tomando como aproximación inicial los parámetros del ovoide calculados en los pasos del 1 al 4.

5.0.4. Detalles de implementación

La implementación de esta parte se ha realizado con otras dos funciones en lenguaje C. La primera de ellas, llamada *ami_inicializar_elipse_deformada*, calcula (siguiendo los pasos 1 al 4 del esquema anterior) la elipse deformada inicial, que se ajusta a una región plana dada por una superficie de nivel, que no es otra que el conjunto de nivel devuelto por las funciones de la etapa anterior (cálculo de los contornos). La parte más sensible a errores del algoritmo es el cálculo de la dirección principal (eje de simetría

de la elipse deformada) $v_{\text{máx}}$. En el algoritmo se permiten otras opciones, además de la señalada en los pasos del esquema, para calcular dicha dirección. Concretamente las opciones implementadas son :

- Utilizar los momentos de orden 2 de toda la forma (la descrita en el algoritmo)
- Utilizar los momentos de orden 2 de solo el contorno
- Utilizar la dirección del punto que más dista del centroide.

La aproximación inicial del ovoide obtenida con esta función se dibuja sobre la imagen original con un determinado nivel de gris.

La segunda de las funciones, llamada *ami_calcular_elipse_deformada_nube_puntos*, mejora los parámetros de la elipse deformada inicial, ajustándola al contorno exterior extraído de la oreja en la etapa previa. Utiliza para ello al método del gradiente paso adaptativo.

El ovoide mejorado también se dibuja sobre la imagen original, en un nivel de gris diferente al anterior, dando lugar a una imagen como la de la figura 5-5

Una propiedad interesante del modelo de distorsión de la elipse propuesto es que se comporta bien ante transformaciones por homotecias (cambios de tamaño). De hecho el parámetro de distorsión γ es una propiedad geométrica de la forma que no depende del tamaño de la elipse deformada. Efectivamente, si hacemos el cambio $(x, y) \rightarrow (\lambda x, \lambda y)$ en la forma, utilizando la ecuación implícita del modelo obtenemos

$$\frac{\lambda^2 x^2}{a^2} \left(1 + \gamma \frac{\lambda x}{a}\right)^2 + \frac{\lambda^2 y^2}{b^2} = \left(1 + \gamma \frac{\lambda x}{a}\right)^2$$

que se ajusta de nuevo a un modelo de elipse deformada de parámetros $\tilde{a} = \frac{a}{\lambda}$, $\tilde{b} = \frac{b}{\lambda}$ y $\tilde{\gamma} = \gamma$



Figura 5-5: En la imagen se pueden apreciar los puntos que se han obtenido como contorno externo de la imagen (dibujados en negro y unidos por un segmento blanco que cierra todo el contorno), el primer ovoide obtenido como aproximación a este contorno (todo en negro), y el ovoide final mejorado (en blanco).

Criterios de comparación entre la geometría de dos orejas.

El objetivo final de nuestro trabajo es, partiendo de la imagen de una oreja de un individuo dado, intentar obtener un conjunto de características medibles (vector de características), de manera que, basándonos en él, seamos capaces de clasificar esa imagen de oreja como perteneciente o no a un individuo determinado. La clasificación se realizaría basándose en la comparación entre los vectores de características de la imagen de entrada y todos los vectores de características de la base de datos con la que trabajamos.

Hasta ahora, hemos extraído algunas características de una imagen de oreja, como los parámetros que caracterizan al mejor ovoide que aproxima su contorno externo y una serie de puntos, que podemos considerar como la aproximación más exacta a los puntos de su contorno externo e interno. Si ahora estamos interesados en comparar más de una imagen de orejas entre sí, necesitaremos transformar las coordenadas de los puntos de los contornos obtenidos, para ponerlos todos en un mismo sistema de referencia que nos permita realizar la comparación.

Para esta transformación nos apoyamos en los parámetros del ovoide que hemos asociado al contorno externo. Vamos a transformar esta elipse deformada en una elipse

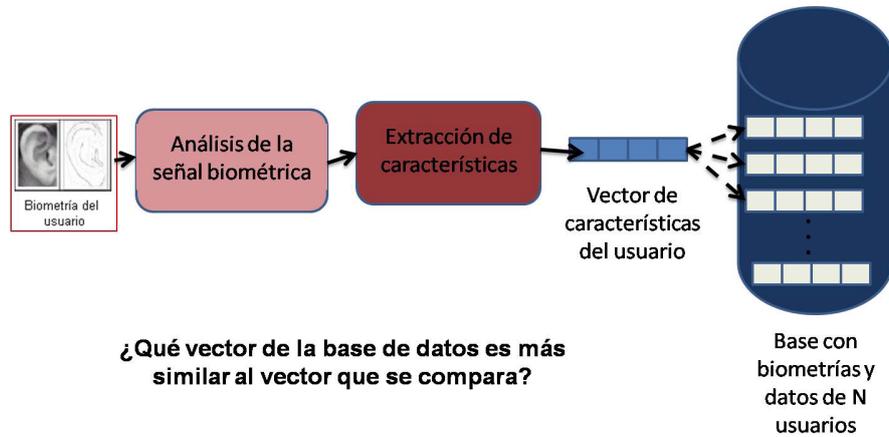


Figura 6-1: Clasificación de imágenes

centrada en el origen, sin rotación y de área fija.

La ecuación general de la elipse, en forma paramétrica, es:

$$x(t) = x_c + a \cos(t) \cos(\alpha) + b \sin(t) \sin(\alpha)$$

$$y(t) = y_c - a \cos(t) \sin(\alpha) + b \sin(t) \cos(\alpha)$$

La transformación euclídea que lleva el centro de la elipse al $(0, 0)$, la deja horizontal y con un área de 256^2 es:

$$H = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) & -\cos(\alpha)x_c + \sin(\alpha)y_c \\ \sin(\alpha) & \cos(\alpha) & -\sin(\alpha)x_c - \cos(\alpha)y_c \\ 0 & 0 & \frac{\sqrt[2]{\pi ab}}{256} \end{pmatrix}$$

En este punto, tenemos dos opciones:

- Podemos, simplemente, aplicar una rotación de $(-\frac{\pi}{2})$ para poner las elipses transformadas en vertical (las orejas, por su disposición, se aproximan mejor a un ovoide de eje mayor vertical) y dejar la excentricidad sin fijar (con lo cual con-

Comparación de dos orejas

servamos más información, aunque no se corrige la deformación de perspectiva), por lo que cada imagen tendrá una excentricidad variable. En este caso, la transformación final a aplicar a cada punto de la imagen, así como a los puntos que hemos señalado como contorno externo e interno, sería:

$$\begin{aligned}
 H'' &= \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) & -\cos(\alpha)x_c + \sin(\alpha)y_c \\ \sin(\alpha) & \cos(\alpha) & -\sin(\alpha)x_c - \cos(\alpha)y_c \\ 0 & 0 & \frac{\sqrt[2]{\pi ab}}{256} \end{pmatrix} = \\
 &= \begin{pmatrix} \sin(\alpha) & \cos(\alpha) & -\sin(\alpha)x_c - \cos(\alpha)y_c \\ -\cos(\alpha) & \sin(\alpha) & \cos(\alpha)x_c - \sin(\alpha)y_c \\ 0 & 0 & \frac{\sqrt[2]{\pi ab}}{256} \end{pmatrix}
 \end{aligned}$$

- O bien, además de la rotación anterior, podemos aplicar otra transformación que comprime/expande la elipse para fijar su excentricidad a un valor fijo dado E (en nuestro algoritmo de implementación, E es un parámetro de entrada que hemos fijado al valor 1,9). Con esta transformación perdemos algo de información pero, a cambio, se corrige parcialmente la deformación de la perspectiva de algunas imágenes. La matriz de transformación que realiza esto sería:

$$H' = \begin{pmatrix} \sqrt[2]{\frac{Eb}{a}} & 0 & 0 \\ 0 & \sqrt[2]{\frac{a}{Eb}} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

En este caso, la transformación final a aplicar sería:

$$H'' = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \sqrt[2]{\frac{Eb}{a}} & 0 & 0 \\ 0 & \sqrt[2]{\frac{a}{Eb}} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) & -\cos(\alpha)x_c + \sin(\alpha)y_c \\ \sin(\alpha) & \cos(\alpha) & -\sin(\alpha)x_c - \cos(\alpha)y_c \\ 0 & 0 & \frac{\sqrt[2]{\pi ab}}{256} \end{pmatrix} =$$

$$= \begin{pmatrix} \sqrt[2]{\frac{Eb}{a}} \sin(\alpha) & \sqrt[2]{\frac{Eb}{a}} \cos(\alpha) & \sqrt[2]{\frac{Eb}{a}} (-\sin(\alpha)x_c - \cos(\alpha)y_c) \\ -\sqrt[2]{\frac{Eb}{a}} \cos(\alpha) & \sqrt[2]{\frac{Eb}{a}} \sin(\alpha) & \sqrt[2]{\frac{Eb}{a}} (\cos(\alpha)x_c - \sin(\alpha)y_c) \\ 0 & 0 & \frac{\sqrt[2]{\pi ab}}{256} \end{pmatrix}$$

En la imagen 6-2 podemos apreciar, para un mismo individuo, los contornos internos y externos transformados, dejando variable la excentricidad, de siete imágenes de sus orejas. Las curvas correspondientes a los contornos interno y externo de la misma imagen están dibujadas en el mismo color. Cinco de esas imágenes corresponden a la misma oreja (oreja derecha), tomadas a la misma distancia focal de la cámara, en distintas posiciones: con el individuo mirando al frente (FRONTAL, `_front`), mirando arriba y abajo (ARRIBA, `_up`, ABAJO, `_down`) y mirando a izquierda y derecha (IZQUIERDA, `_left`, DERECHA, `_right`), en unas marcas de referencia previamente establecidas. La sexta imagen corresponde a la misma oreja, con el individuo mirando al frente, pero modificando la distancia focal de la cámara con la que se captura la imagen (ZOOM, `_zoom`). La última imagen corresponde a la otra oreja del individuo (oreja izquierda), mirando al frente y tomada con la misma distancia focal de la cámara que las cinco anteriores (POSTERIOR, `_back`).

De esta manera, estaríamos ya en disposición de comparar, visualmente, varias imágenes de orejas; pero necesitamos extraer algunas características adicionales, para sumar al vector de características obtenido hasta ahora, de forma que esta comparación pueda realizarse automáticamente.

En cada imagen calculamos el centroide del conjunto de puntos transformados que aproxima al contorno interno de la oreja y, vamos a obtener distintas medidas de distancias desde el centroide hasta los puntos de los contornos externo e interno transformados.

El contorno interno de las orejas tiene un "punto característico", fácilmente identificable a simple vista, que es donde se une el pliegue anterior y el pliegue superior, en

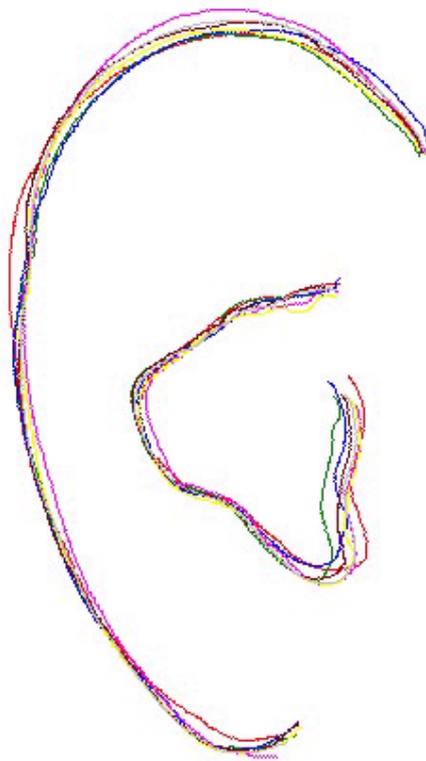


Figura 6-2: Contornos externos e internos transformados, usando excentricidad variable, de 7 imágenes de las orejas de un mismo individuo. Los contornos interno y externo correspondientes a la misma imagen están dibujados en el mismo color.

la base de la Foseta que, además, es el que hemos elegido como primer punto a marcar en el polígono inicial para la obtención del contorno. Este punto es el que va a servir de referencia para la obtención del resto de las características con las que tratamos de identificar las imágenes de cada oreja.

Primeramente calculamos la tangente de la recta que uniría el centroide obtenido con este punto, lo que nos proporciona un ángulo de partida para el resto del proceso. Añadiéndole a este ángulo inicial un determinado incremento, vamos a ir barriendo los contornos interno y externo, obteniendo un determinado número de distancias desde el centroide a los puntos de los contornos de la oreja. El número de distancias a calcular es, en nuestra implementación, un parámetro de entrada definido por el usuario y que nosotros hemos fijado a 64.

El barrido para los puntos del contorno interno comienza, como ya hemos señalado, en ese "punto característico" (con el ángulo de referencia original) y, para intentar abarcar todo el contorno, se van añadiendo 63 incrementos de ángulo, iguales, hasta completar un ángulo total de barrido de $\frac{11\pi}{6}$, con lo que obtenemos la medida de 64 distancias desde el centroide a los puntos del mencionado contorno.

Para el contorno externo, partimos del mismo ángulo inicial de referencia pero, para incluir el primer punto de dicho contorno, debemos barrer hacia atrás un determinado ángulo (para las imágenes de la base de datos con la que hemos trabajado, hemos considerado suficiente un ángulo de $\frac{\pi}{12}$) y luego, en el otro sentido, hasta alcanzar el lóbulo. El ángulo total barrido para este contorno es de $\frac{4\pi}{3}$. A lo largo de este recorrido obtenemos 64 distancias desde el centroide a los puntos del contorno externo. Obsérvese que las 64 distancias desde el centroide a los puntos del contorno interno y externo no están medidas para los mismos incrementos de ángulo.

De esta forma, para cada imagen de oreja, obtenemos 128 características adicionales (las 128 distancias medidas desde el centroide a los puntos de los dos contornos) que, añadidas a los ocho parámetros que caracterizaban la elipse deformada (longitud de

Comparación de dos orejas

los semiejes mayor y menor, coordenadas del centro, excentricidad de la elipse, ángulo de rotación y distorsión del ovoide, y error medio medido en píxeles), dan lugar a un vector de características de 136 componentes para cada imagen.

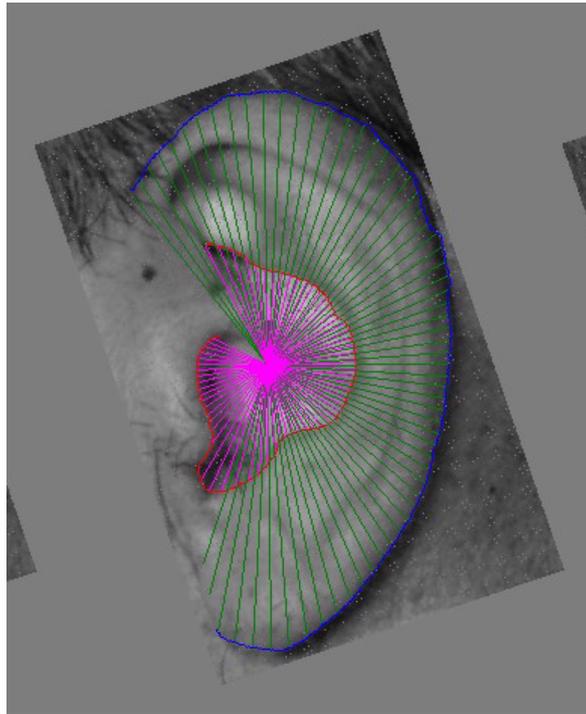


Figura 6-3: *Distancias medidas desde el centroide del contorno interno hasta los puntos del contorno interno (en fucsia) y a los del contorno externo (en verde).*

El la figura 6-3 podemos apreciar los puntos transformados correspondientes a los contornos externos (en azul) e interno (en rojo), así como las distancias medidas desde el centroide hasta los puntos de los contornos (en verde para el contorno externo y en fucsia para el contorno interno).

Para probar la tesis de que el vector de características, así obtenido, sirve para clasificar las imágenes de las orejas, como pertenecientes a un individuo determinado y, ante la imposibilidad de acceder a bases de datos con imágenes de orejas más o menos estándar, que nos permitieran comparar nuestros resultados con aquellos obtenidos en otros estudios realizados previamente sobre el mismo tema (referenciados en el capítulo

de Estado del Arte de esta misma tesis), hemos optado por generar nuestra propia base de datos de imágenes.

Siempre bajo las mismas condiciones de iluminación, hemos tomado imágenes de las orejas de 100 individuos, de diferentes edades (entre 19 y 65 años), raza y sexo. Para cada individuo se han tomado seis imágenes de la oreja derecha (que conforman el grupo de imágenes para el análisis) y una imagen de la oreja izquierda (que es la que se ha utilizado como imagen de control). Se ha utilizado una cámara de fotografía digital Nikon D100, con una distancia focal de 135mm para seis de las siete imágenes y de 200mm para la imagen que hemos llamado ZOOM. El individuo se ha colocado sentado, a una distancia de unos 2 metros de la cámara, de perfil, mirando a unos puntos de referencia previamente fijados.

Cinco de las imágenes recogidas corresponden a la misma oreja (oreja derecha), con el individuo mirando al frente (FRONTAL, `_front`), mirando arriba y abajo (ARRIBA, `_up`, ABAJO, `_down`) y mirando a izquierda y derecha (IZQUIERDA, `_left`, DERECHA, `_right`). La sexta imagen corresponde a la misma oreja, con el individuo mirando al frente, pero modificando la distancia focal de la cámara con la que se captura la imagen (ZOOM, `_zoom`). La última imagen corresponde a la otra oreja del individuo (oreja izquierda), mirando al frente y tomada con la misma distancia focal de la cámara que las cinco anteriores (POSTERIOR, `_back`).

La base de datos construída contiene información referente a 100 casos (100 individuos) y, para cada uno de ellos tenemos la medida de 141 variables, que son las siguientes: identificador del individuo, identificador de la imagen, sexo, raza, y edad del individuo, y las 136 componentes del vector de características determinado para la imagen en cuestión.

Resultados experimentales y análisis estadístico

7.1. Análisis Estadístico Básico de los Datos

7.1.1. Introducción

El objetivo de esta sección es explorar si, con los datos obtenidos de las diferentes vistas de las orejas, de los 100 individuos analizados, es posible establecer un método de clasificación, que permita identificar correctamente a cada individuo.

Vamos, en primer lugar, a introducir la notación que vamos a utilizar en este análisis:

- x_{ijk} : parámetro k , de la vista j , del individuo i
- N_k : Número de parámetros. En nuestro caso hemos tomado 64 medidas en el lóbulo interno de la oreja, 64 medidas en el lóbulo externo de la oreja, la excentricidad de la elipse asociada a la oreja y la distorsión del ovoide asociado a la orejas lo cual hace un total de 130 parámetros.
- N_j : Número de vistas de cada oreja. En nuestro caso tenemos las vistas FRONTAL - ARRIBA - ABAJO - ZOOM - POSTERIOR - IZQUIERDA y DERECHA, lo

cual hace un total de 7 vistas para cada oreja.

- N_i : Número de individuos. En nuestro caso hemos analizado 100 individuos
- $x_k = \frac{1}{N_i N_j} \sum_{i=0}^{N_i} \sum_{j=0}^{N_j} x_{ijk}$ (media total para todos los individuos y todas las vistas del parámetro k)
- $V_k^{total} = \frac{1}{N_i N_j} \sum_{i=0}^{N_i} \sum_{j=0}^{N_j} (x_{ijk} - x_k)^2$ (varianza total para todos los individuos y todas las vistas del parámetro k)
- $x_{ik} = \frac{1}{N_j} \sum_{j=0}^{N_j} x_{ijk}$ (media intra grupo del parámetro k para todas las vistas del individuo i)
- $V_{ik}^{intra_grupo} = \frac{1}{N_j} \sum_{j=0}^{N_j} (x_{ijk} - x_{ik})^2$ (varianza intra grupo del parámetro k para todas las vistas del individuo i)
- $V_k^{inter_grupo} = \frac{1}{N_i} \sum_{i=0}^{N_i} (x_{ik} - x_k)^2$ (varianza inter grupo del parámetro k para todas las medias intra grupos)

Un resultado clásico fundamental que vamos a utilizar a lo largo de esta sección es el siguiente:

Teorema 1 *La varianza total V_k^{total} se puede descomponer a partir de las varianzas intra-grupo e inter-grupo de la siguiente forma:*

$$V_k^{total} = V_k^{inter_grupo} + \frac{1}{N_i} \sum_{i=0}^{N_i} V_{ik}^{intra_grupo} \quad (7.1)$$

Demostración: La demostración es clásica y la ponemos aquí a título ilustrativo:

$$\begin{aligned} V_k^{total} &= \frac{1}{N_i N_j} \sum_{i=0}^{N_i} \sum_{j=0}^{N_j} (x_{ijk} - x_k)^2 = \frac{1}{N_i N_j} \sum_{i=0}^{N_i} \sum_{j=0}^{N_j} (x_{ijk} - x_{ik} + x_{ik} - x_k)^2 = \\ &= \frac{1}{N_i N_j} \sum_{i=0}^{N_i} \sum_{j=0}^{N_j} (x_{ijk} - x_{ik})^2 + \frac{1}{N_i N_j} \sum_{i=0}^{N_i} \sum_{j=0}^{N_j} (x_{ik} - x_k)^2 + \\ &+ 2 \frac{1}{N_i N_j} \sum_{i=0}^{N_i} \sum_{j=0}^{N_j} (x_{ijk} - x_{ik})(x_{ik} - x_k) = \end{aligned}$$

$$\begin{aligned}
& \frac{1}{N_i} \sum_{i=0}^{N_i} V_{ik}^{intra_grupo} + V_k^{inter_grupo} + 2 \frac{1}{N_i N_j} \sum_{i=0}^{N_i} \sum_{j=0}^{N_j} (x_{ijk} x_{ik} - x_{ijk} x_k - x_{ik}^2 + x_{ik} x_k) = \\
& \frac{1}{N_i} \sum_{i=0}^{N_i} V_{ik}^{intra_grupo} + V_k^{inter_grupo} + 2 \frac{1}{N_i} \sum_{i=0}^{N_i} (x_{ik} x_{ik} - x_{ik} x_k - x_{ik}^2 + x_{ik} x_k) = \\
& \frac{1}{N_i} \sum_{i=0}^{N_i} V_{ik}^{intra_grupo} + V_k^{inter_grupo} \text{ c.q.d.}
\end{aligned}$$

Una medida de la significación estadística del parámetro k viene dada por

$$s_k = \frac{\sum_{i=0}^{N_i} V_{ik}^{intra_grupo}}{N_i \cdot V_k^{total}} \quad (7.2)$$

s_k toma valores entre 0 y 1 y cuanto menor es s_k más significativo es el parámetro k , pues lo que interesa es que la varianza entre los grupos sea lo mayor posible y la varianza del parámetro, respecto a la media de su grupo, sea lo menor posible.

Para visualizar cuales de las distancias, consideradas como parámetros, resultan ser más significativas para el conjunto de vistas analizadas, hemos construido lo que representaría la oreja "media" de la base de datos considerada. Para cada parámetro, hemos obtenido su valor medio, considerando los valores que toma para los 100 individuos y para las 7 vistas de cada individuo, y hemos calculado el valor de s_k para ese parámetro. Representando luego el valor medio obtenido para cada parámetro como un punto, en la posición del contorno correspondiente, y con un nivel de gris, cuya intensidad es proporcional al valor de s_k , obtenemos una imagen como la de la derecha de la figura 7-1. Cuanto más oscuro es el nivel de gris del punto dibujado, más significativo resulta ser el parámetro correspondiente. Podemos apreciar que las zonas del contorno exterior más cercanas a la unión con la cabeza (lóbulo y parte superior de la oreja) son las que parecen ser menos significativas, lo que puede indicar que las distancias del centroide a esas zonas de la oreja, pueden variar mucho para las vistas de un mismo individuo.

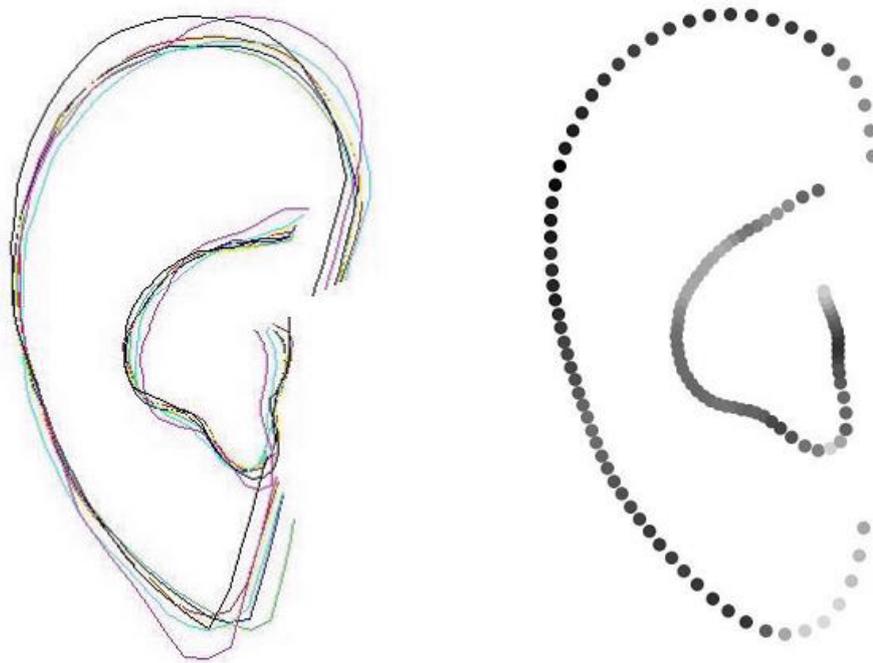


Figura 7-1: A la izquierda se representan las 7 vistas de la oreja de un individuo, una vez escaladas, rotadas y trasladadas, para que encajen correctamente. A la derecha se muestra lo que sería la oreja media, contando todos los individuos y todas las vistas. El nivel de gris de los círculos es proporcional a la significación estadística s_k . Cuanto más oscuro es el nivel de gris más significativo es el parámetro

7.1.2. Caracterización y comparación entre individuos

Para clasificar a los individuos, a partir de las medidas tomadas sobre las orejas, nos concentraremos en 2 aspectos concretos del problema de clasificación:

1. Selección de un número de parámetros mínimo, que permita discriminar correctamente los individuos. En lo que sigue denotaremos por k_s , $s = 1, \dots, N_s$ al conjunto de parámetros seleccionados. Cada individuo se identificará, a partir de las medias intra-grupo, para cada parámetro. Es decir, el individuo i será caracterizado por $\{x_{ik_s}\}_{s=1, \dots, N_s}$
2. Selección de la distancia apropiada para comparar una vista j , de un individuo i , con la oreja "media" de otro individuo i' . Por tanto, la distancia se escribirá como $d(j, i, i')$

Respecto a la selección de la distancia, hemos analizado diferentes posibilidades. Concretamente hemos realizado experimentos con las siguientes distancias :

1. La distancia euclídea:

$$d_E(j, i, i') = \sum_{s=0}^{N_s} (x_{ijk_s} - x_{i'k_s})^2 \quad (7.3)$$

2. Distancia normalizada a través de la varianza total

$$d_T(j, i, i') = \sum_{s=0}^{N_s} \frac{(x_{ijk_s} - x_{i'k_s})^2}{V_{k_s}^{total}} \quad (7.4)$$

en este caso escalamos el peso de cada variable en función de su varianza global

3. Distancia normalizada mejorada a través de la varianza inter-grupos

$$d_I(j, i, i') = \sum_{s=0}^{N_s} (1 - s_{k_s}) \frac{(x_{ijk_s} - x_{i'k_s})^2}{V_{k_s}^{total}} \quad (7.5)$$

este caso es un refinamiento del anterior, donde otorgamos más peso, en la distancia, a las variables más significativas.

Es importante resaltar aquí que, para aumentar la fiabilidad del análisis, algunas de las vistas se utilizan para seleccionar los parámetros y otras se utilizan para evaluar su capacidad discriminante. De esta forma, los resultados no se ven alterados por el hecho de que las vistas que se quieren clasificar hayan sido utilizadas para diseñar el método de clasificación, lo cual supondría una limitación importante a la hora de interpretar los resultados. Por ejemplo, una configuración sobre la que hemos trabajado bastante, es tomar las vistas FRONTAL - ARRIBA y ABAJO para realizar la selección de parámetros y, posteriormente, tomar la vista ZOOM para validar el análisis. Esta configuración resulta bastante lógica, pues todas las vistas corresponden a una situación donde se toma una fotografía fronto-paralela del individuo. Las vistas POSTERIOR, IZQUIERDA y DERECHA resultan más complejas de analizar porque, o bien corresponden a la otra oreja del individuo, o bien incluyen un fenómeno de distorsión de la forma de la oreja debido a la perspectiva.

7.1.3. Método de selección de parámetros

Una parte muy importante del análisis consiste en la selección de los parámetros que se utilizarán (k_s , $s = 1, \dots, N_s$) para caracterizar a los individuos (recuerdese que cada individuo i viene representado por las medias de las vistas de los parámetros seleccionados $\{x_{ik_s}\}_{s=1, \dots, N_s}$). El objetivo es que el número de parámetros sea el menor posible y, al mismo tiempo, el poder discriminante de estos parámetros sea máximo. Influenciados por la experiencia previa de nuestro grupo de investigación, hemos optado por un método para la selección de parámetros basado en la minimización de una energía que se adapte al caso de estudio. Para introducir esta energía primero

introduciremos algunas expresiones básicas que vamos a utilizar.

$$\begin{aligned} & \text{Error_Clasificación}(\{k_s\}_{s=1,\dots,N_s}) = \\ & = \{\text{N}^\circ \text{ vistas } j, \text{ de individuos } i: \exists i' \neq i \text{ tal que } d(j, i, i') < d(j, i, i)\} \end{aligned} \quad (7.6)$$

se entiende que una vista está mal clasificada si existe algún otro individuo (distinto a la persona a la que pertenece la oreja) para la cual, la distancia de la vista a la oreja, es menor, para dicho individuo, que para el individuo al que pertenece la oreja. Asociado a este concepto de error vamos a utilizar también como factor, para valorar el éxito del análisis, la siguiente expresión:

$$\begin{aligned} & \text{Exito_Clasificación_Acumulado}(\{k_s\}_{s=1,\dots,N_s}, M) = \\ & = \{\text{N}^\circ \text{ vistas } j, \text{ de individuos } i : d(j, i, i) < d(j, i, i') \forall i' \neq i \text{ salvo a lo sumo en } M \text{ individuos}\} \end{aligned} \quad (7.7)$$

es decir, el factor de éxito de clasificación acumulado mide, para un entero positivo M , cuantas vistas cumplen que, el número de individuos distintos al que pertenece la oreja, y cuya distancia a la vista sea menor que la distancia al propio individuo, sea menor o igual que M . En otras palabras, se mide la magnitud del error, en términos del número de individuos que resultarían elegidos antes que el individuo al que pertenece la oreja.

Para analizar el grado de dependencia de 2 parámetros k y k' utilizaremos la correlación que viene dada por la conocida expresión

$$c(k, k') = \frac{\frac{1}{N_i N_j} \sum_{i=0}^{N_i} \sum_{j=0}^{N_j} (x_{ijk} - x_k)(x_{ijk'} - x_{k'})}{\sqrt{V_k^{total}} \sqrt{V_{k'}^{total}}} \quad (7.8)$$

Finalmente la selección de parámetros se realizará apoyándonos en la minimización

de la siguiente función de error

$$Error(\{k_s\}_{s=1,\dots,N_s}) = w_C \cdot Error_Clasificación(\{k_s\}_{s=1,\dots,N_s}) + w_S \cdot \sum_{s=1}^{N_s} s_{k_s} + w_T \max_{\substack{s, s' \in \{1, \dots, N_s\} \\ s \neq s'}} c(k_s, k_{s'})$$

donde w_C , w_S , y w_T representan los pesos asociados a cada término. En los experimentos que se muestran en esta sección siempre se ha tomado $w_C = w_S = w_T = 1$. Básicamente, la función de error ha sido definida para tener en cuenta, simultáneamente, los aspectos más relevantes de la selección de parámetros, es decir : (1) que los parámetros clasifiquen a la población lo mejor posible, (2) que la significación estadística de los parámetros elegidos sea la mayor posible y (3) que los parámetros elegidos sean lo más independientes posible.

Encontrar el mínimo absoluto de la anterior función de error, inspeccionando todas las posibles combinaciones de parámetros, para un número prefijado de parámetros N_s , es un problema de un costo computacional enorme. Por ello, diseñaremos un método incremental de selección de parámetros que, si bien no asegura obtener el mínimo absoluto de la función de error, permite encontrar rápidamente una solución razonable al problema de selección de variables.

El método incremental que proponemos es muy sencillo: conocida una aproximación del mínimo para $N_s - 1$ parámetros, añadimos un nuevo parámetro k_{N_s} minimizando la función unidimensional de error simplificada:

$$Error(k_{N_s}) = w_C \cdot Error_Clasificación(\{k_s\}_{s=1,\dots,N_s}) + w_S \cdot s_{k_{N_s}} + w_T \max_{s \in \{1,\dots,N_s-1\}} c(k_s, k_{N_s})$$

lo cual, en nuestro caso, se puede hacer por simple inspección de todas las elecciones posibles de k_{N_s} . El algoritmo se detiene cuando se cumple que el error de clasificación

es 0 (es decir todas las vistas quedan bien clasificadas), o bien se ha alcanzado un número máximo de parámetros preestablecido.

7.1.4. Resultados experimentales del método de selección de parámetros propuesto

En primer lugar, para decidir cual de las distancias definidas en (7.3), (7.4) y (7.5) es la que mejor se adapta a nuestro caso de estudio, vamos a analizar su poder discriminante en el procedimiento de selección de parámetros. Nos situamos en el caso de elegir las vistas FRONTAL - ARRIBA y ABAJO para realizar la selección de parámetros. En la figura 7-2 se ilustra el número de vistas que quedan correctamente clasificadas en función del número de parámetros utilizando la distancia euclídea $d_E(j, i, i')$. Como puede observarse se requieren 12 parámetros para que todas las vistas queden correctamente clasificadas. En la figura 7-3, se muestra el mismo experimento para la distancia normalizada $d_T(j, i, i')$, y en la figura 7-4, para la distancia normalizada mejorada $d_I(j, i, i')$.

En la figura 7-5 se muestra una gráfica comparativa de los 3 experimentos. Como puede observarse con la distancia mejorada sólo se requieren 7 parámetros para clasificar correctamente todas las vistas, lo que sugiere que, esta distancia, es la mejor que se adapta al caso de estudio. Por tanto, en los experimentos que siguen siempre utilizaremos esta distancia.

En la figura 7-6 se ilustra la selección de parámetros cuando utilizamos 4 vistas (FRONTAL - ARRIBA - ABAJO y ZOOM). En este caso resultan necesarios 9 parámetros para clasificar correctamente las 400 vistas.

En la figura 7-7 se ilustra, gráficamente (puntos pintados en color verde), la posición de los parámetros de la oreja que han sido seleccionados para la clasificación, tanto utilizando 3 vistas como 4 vistas. Este gráfico resulta de gran interés, pues identifica

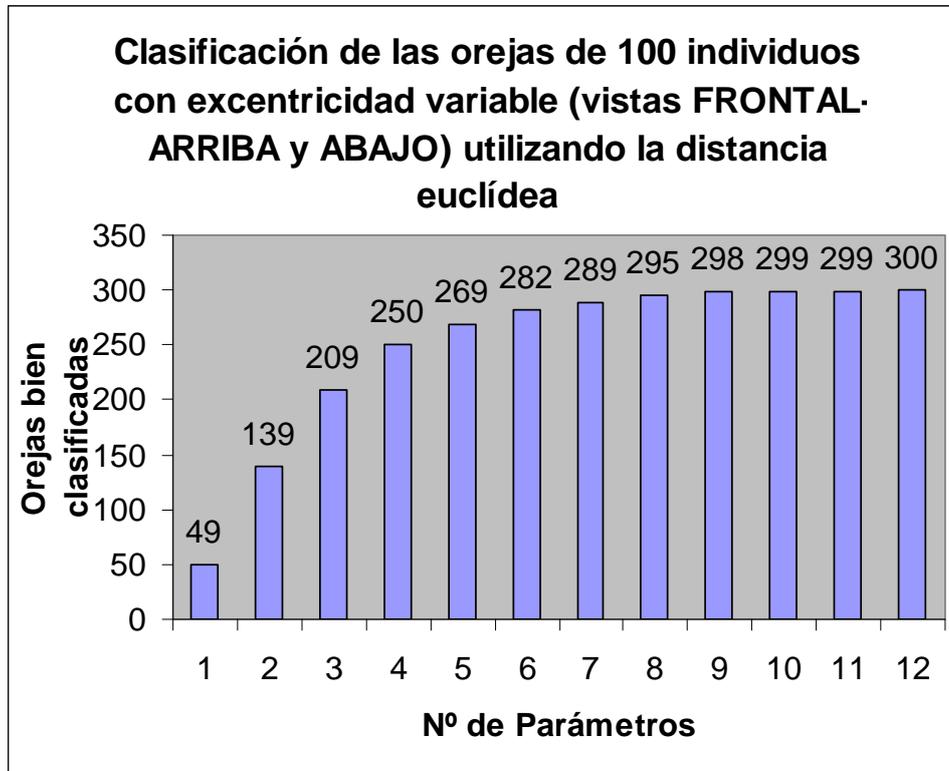


Figura 7-2: Resultado del proceso iterativo de selección de parámetros tomando las vistas FRONTAL-ARRIBA y ABAJO tomando la distancia $d_E(j, i, i')$

(para la población analizada) qué zonas de las orejas son más importantes, de cara a identificar al individuo.

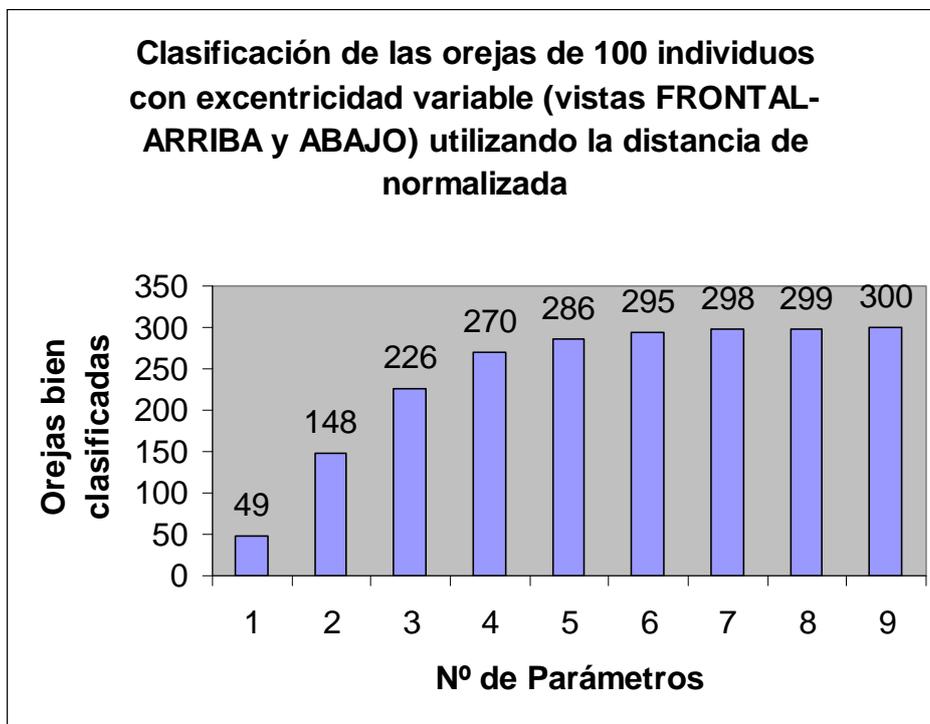


Figura 7-3: Resultado del proceso iterativo de selección de parámetros tomando las vistas FRONTAL-ARRIBA y ABAJO tomando la distancia $d_N(j, i, i')$

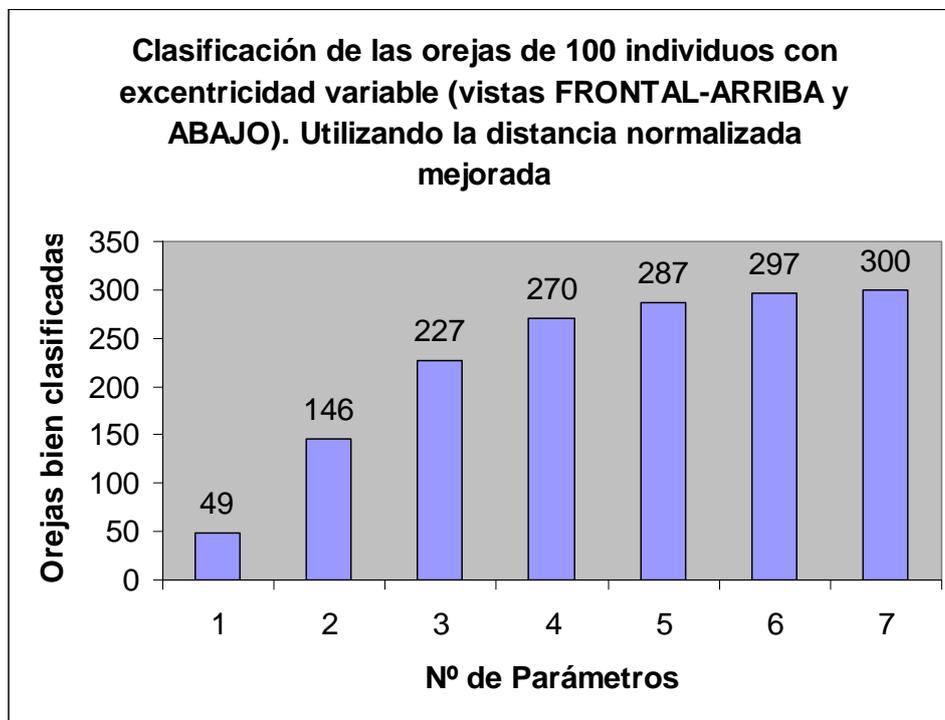


Figura 7-4: Resultado del proceso iterativo de selección de parámetros tomando las vistas FRONTAL-ARRIBA y ABAJO tomando la distancia $d_1(j, i, i')$

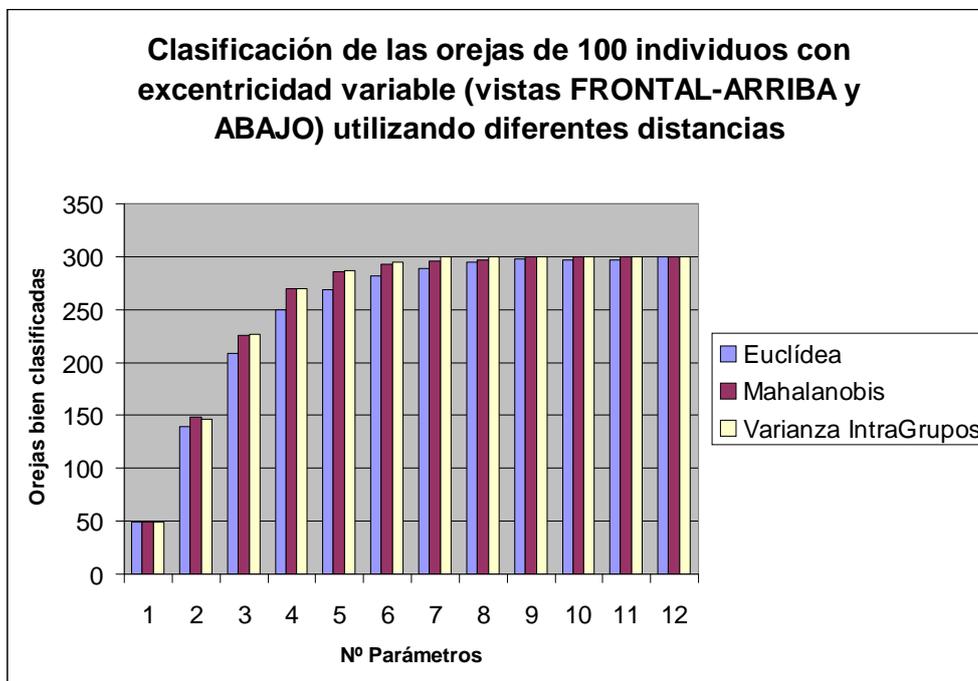


Figura 7-5: Gráfica comparativa de la capacidad de discriminación de las 3 distancias analizadas

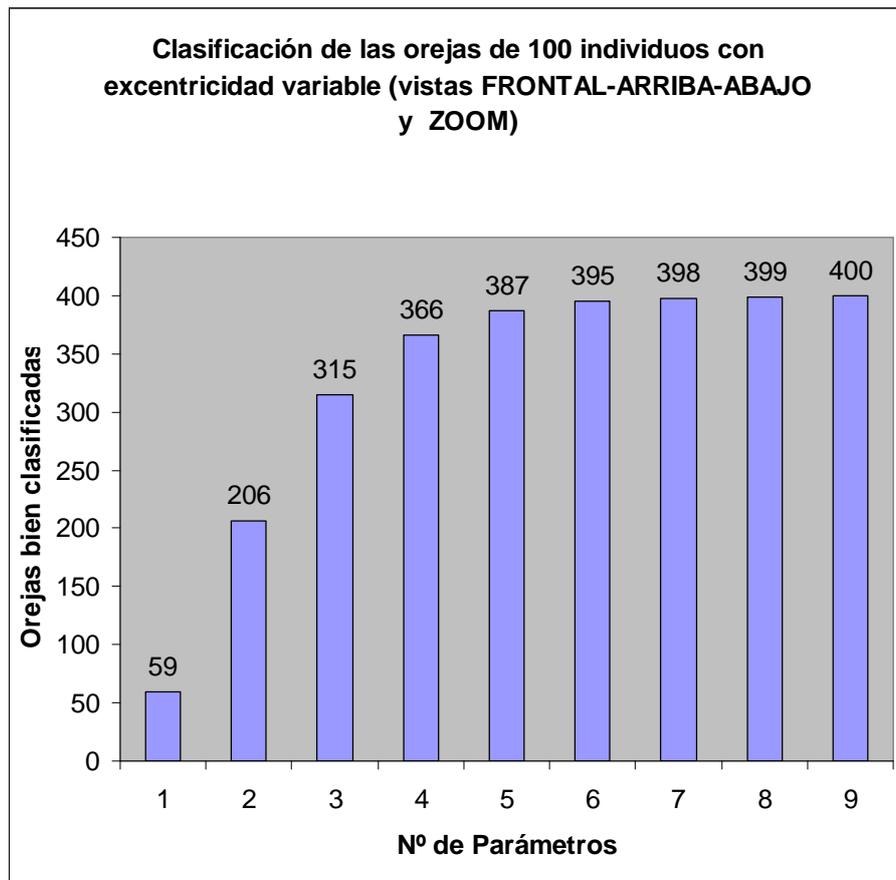


Figura 7-6: Resultado del proceso iterativo de selección de parámetros tomando las vistas FRONTAL-ARRIBA-ABAJO y ZOOM tomando la distancia $d_I(j, i, i')$

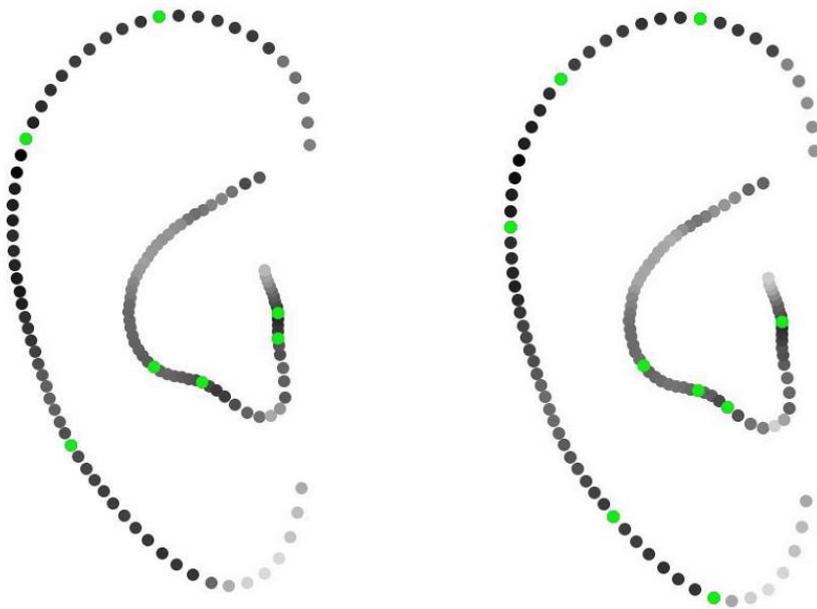


Figura 7-7: A la izquierda, en verde, los 7 parámetros seleccionados para clasificar las vistas *FRONTAL-ARRIBA-ABAJO*. A la derecha, los 9 parámetros seleccionados para clasificar las vistas *FRONTAL-ARRIBA-ABAJO* y *ZOOM*

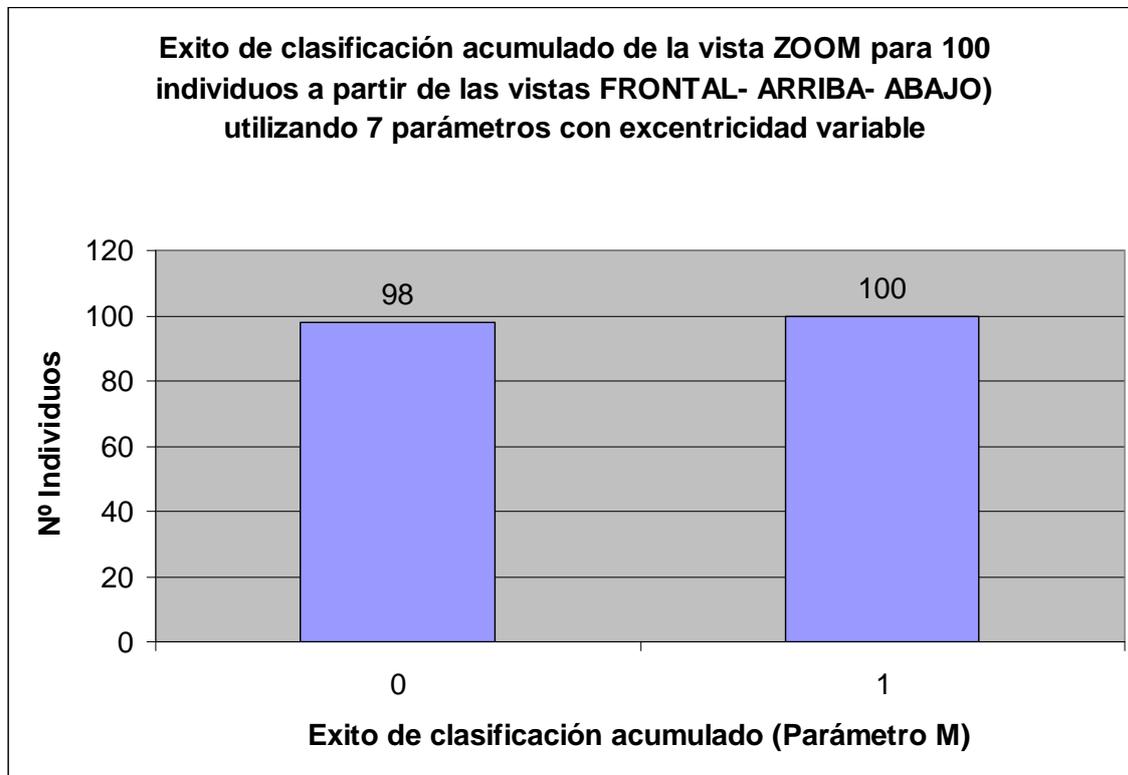


Figura 7-8: De las 100 vistas ZOOM analizadas, 98 han sido correctamente clasificadas y 2 han sido clasificadas como segunda opción, es decir que existe otro individuo más próximo a la vista.

7.1.5. Resultados experimentales de la clasificación de vistas

En este apartado, vamos a analizar si resulta posible clasificar correctamente vistas que no han sido utilizadas en el procedimiento de selección de parámetros. En primer lugar, vamos estudiar el caso en que, la selección de parámetros se realiza utilizando las vistas FRONTAL-ARRIBA y ABAJO y pretendemos, a partir de ellas, clasificar la vista ZOOM. En la figura 7-8 se ilustran los resultados del éxito de clasificación acumulado, definido en (7.7). Se puede observar que, de las 100 vistas ZOOM analizadas, 98 han sido correctamente clasificadas y 2 han sido clasificadas como segunda opción, es decir, que existe otro individuo más próximo a la vista.

A continuación pasamos a una situación más compleja donde, a partir de las vistas FRONTAL-ARRIBA-ABAJO y ZOOM, queremos clasificar la vista POSTERIOR. La situación resulta considerablemente más compleja debido a que, en este caso, la vista es de la otra oreja y, puede darse el caso, de que las orejas no sean perfectamente simétricas. Por otro lado, para sacar la vista posterior, la persona tiene que darse la vuelta, con lo cual, con frecuencia, puede darse que la vista se tome con una perspectiva distinta a las vistas iniciales, fronto-paralelas, de la otra oreja.

Los resultados se ilustran en la figura 7-9. Los resultados son considerablemente peores que en caso anterior. El 33 % de los individuos ha sido correctamente clasificados, 11 individuos han sido clasificados como segunda opción y, para llegar a un 76 % de éxitos, hay que permitir un margen de error de 19 individuos.

A continuación, a partir de las vistas FRONTAL-ARRIBA-ABAJO y ZOOM queremos clasificar la vista DERECHA. En este caso, la mayor dificultad encontrada es la distorsión de perspectiva que se produce al girar la cabeza. Los resultados se ilustran en la figura 7-10. El 30 % de los individuos ha sido correctamente clasificados y, para llegar a un 84 % de éxitos, hay que permitir un margen de error de 19 individuos.

A continuación, a partir de las vistas FRONTAL-ARRIBA-ABAJO y ZOOM queremos clasificar la vista IZQUIERDA. Al girar la cabeza a la izquierda, la distorsión de perspectiva que se produce en la oreja, es mayor que al girar a la derecha. Los resultados se ilustran en la figura 7-11. El 15 % de los individuos ha sido correctamente clasificados y, para llegar a un 66 % de éxitos, hay que permitir un margen de error de 20 individuos.

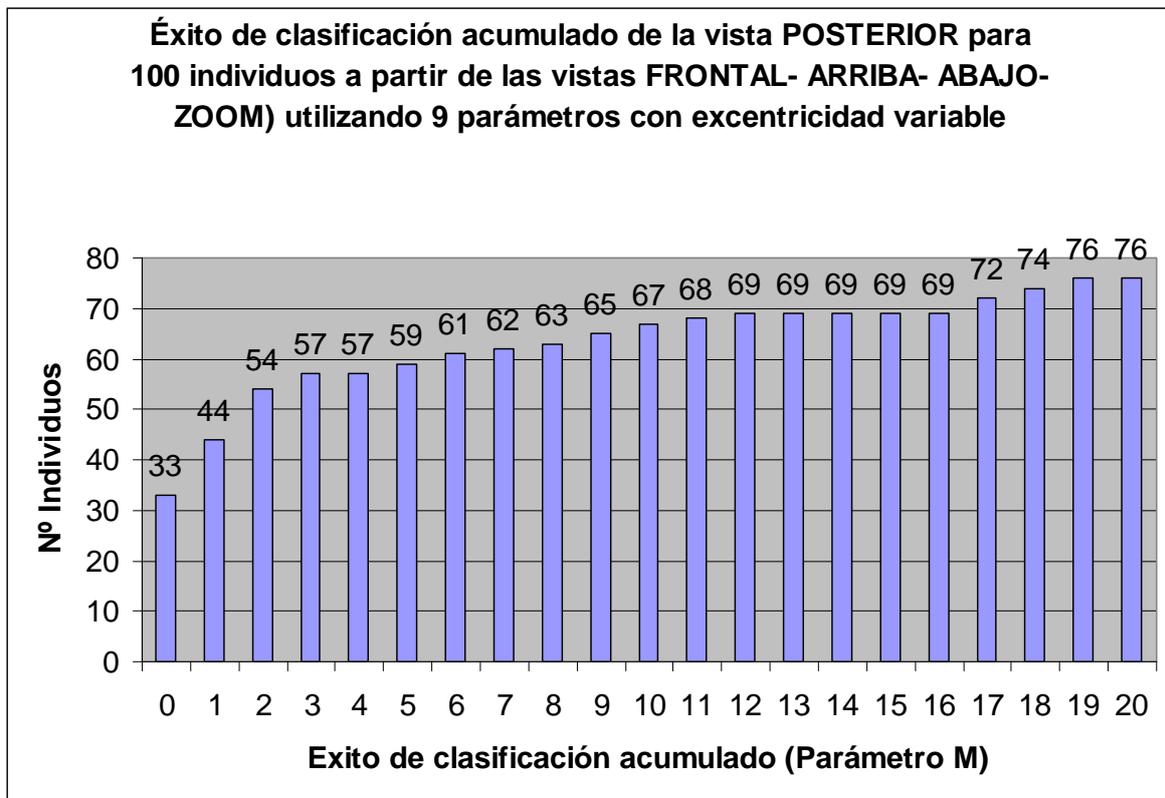


Figura 7-9: El 33 % de los individuos ha sido correctamente clasificados, 11 individuos han sido clasificados como segunda opción y, para llegar a un 76 % de éxitos, hay que permitir un margen de error de 19 individuos

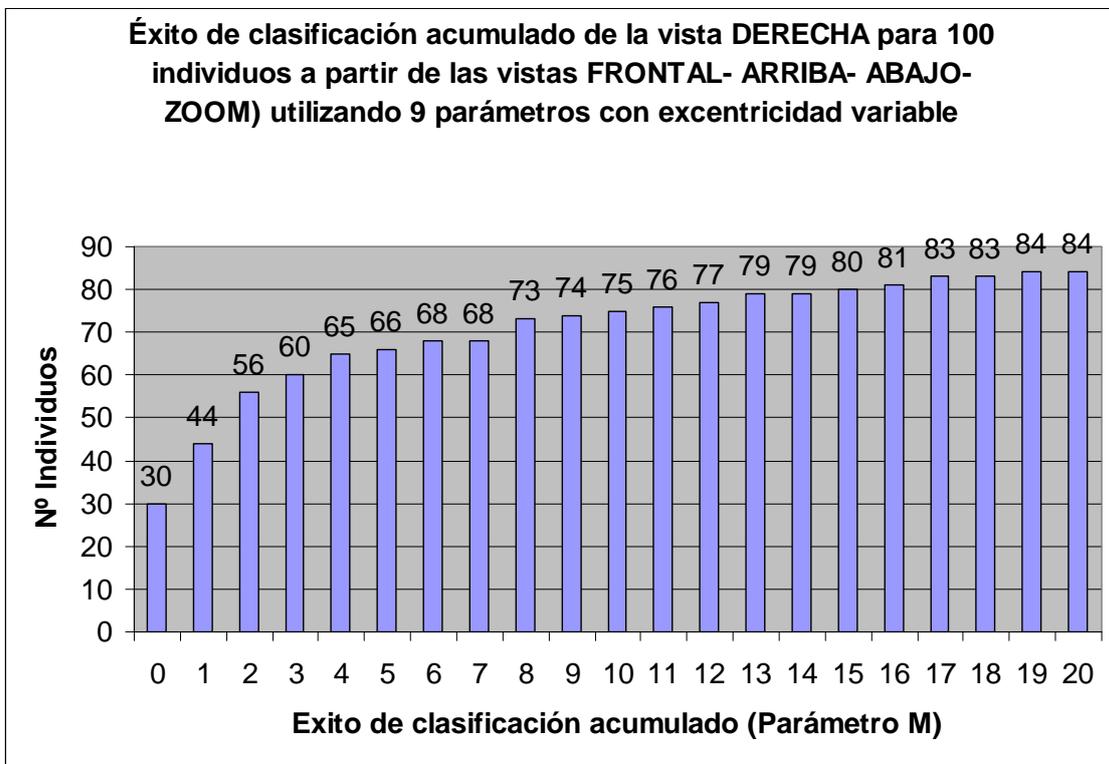


Figura 7-10: El 30 % de los individuos ha sido correctamente clasificados y, para llegar a un 84 % de éxitos, hay que permitir un margen de error de 19 individuos

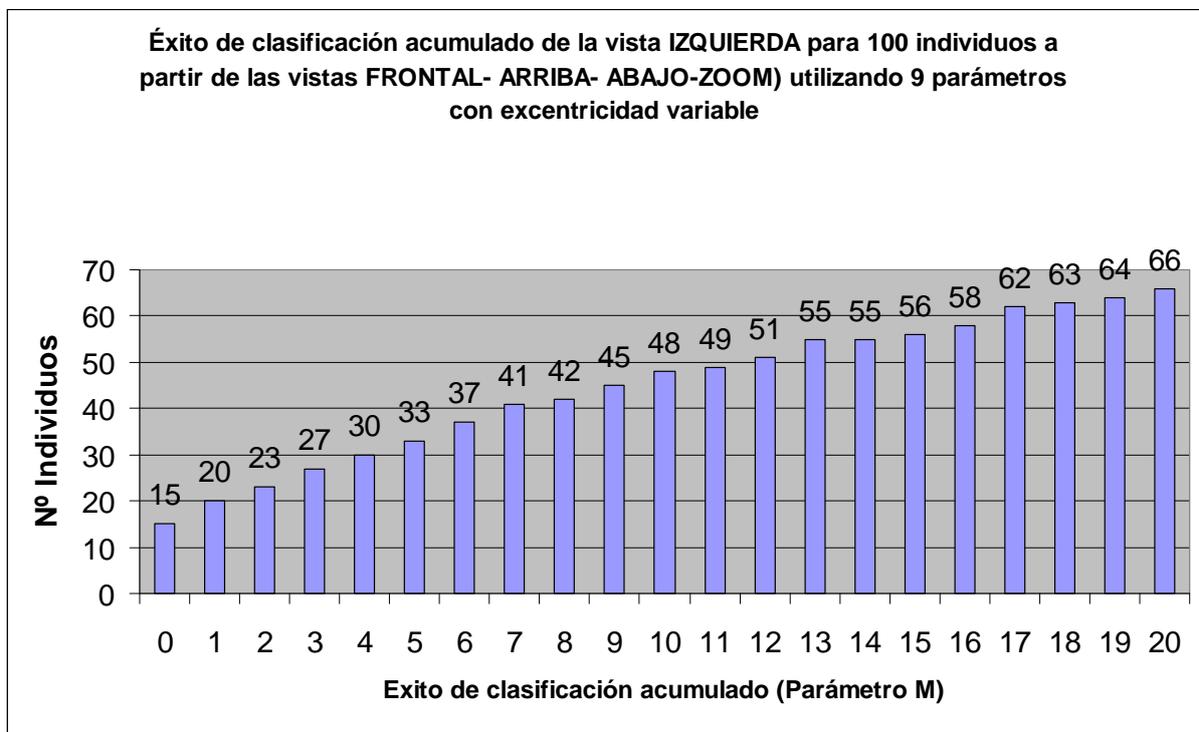


Figura 7-11: El 15 % de los individuos ha sido correctamente clasificados, para llegar a un 66 % de éxitos hay que permitir un margen de error de 20 individuos

7.1.6. Análisis de los datos normalizando la excentricidad de las orejas

En este apartado vamos a reproducir el análisis anterior, cuando tomamos las medidas de las orejas, normalizadas para que la excentricidad de la elipse asociada sea constante. Normalmente debemos esperar que, al normalizar la excentricidad, la clasificación de las vistas donde no hay deformación de perspectiva (vistas FRONTAL-ARRIBA-ABAJO y ZOOM) empeore un poco, pues al normalizar estamos perdiendo información pero, la clasificación de las vistas donde hay distorsión de perspectiva, como DERECHA e IZQUIERDA, deberían mejorar, ya que la normalización de la excentricidad corrige parcialmente la deformación de perspectiva. En la figura 7-12 se ilustra el proceso de clasificación de parámetros a partir de las vistas FRONTAL-ARRIBA y ABAJO. Como puede observarse resulta necesario utilizar 21 parámetros para conseguir clasificar correctamente todas las vistas.

En la figura 7-13 se ilustran las 7 vistas de la oreja de un individuo, después de la normalización de la excentricidad, y la oreja media, contando todos los individuos y todas las vistas. Cuanto más oscuro es el nivel de gris, mas significativa es la medida (s_k es más pequeño).

En la figura 7-14 se ilustra el proceso de clasificación de parámetros a partir de las vistas FRONTAL-ARRIBA, ABAJO y ZOOM.

En la figura 7-15 se ilustra la localización de los parámetros seleccionados, utilizando 3 vistas y utilizando 4 vistas.

Analizaremos ahora el éxito en la clasificación para vistas que no han sido utilizadas para la selección de los parámetros. En la figura 7-16 se muestra el éxito de clasificación de la vista ZOOM a partir de las vistas FRONTAL-ARRIBA y ABAJO. Como puede observarse el resultado es razonablemente bueno: 97 vistas han sido correctamente clasificadas y todas las vistas quedan bien clasificadas con un margen de error de 3

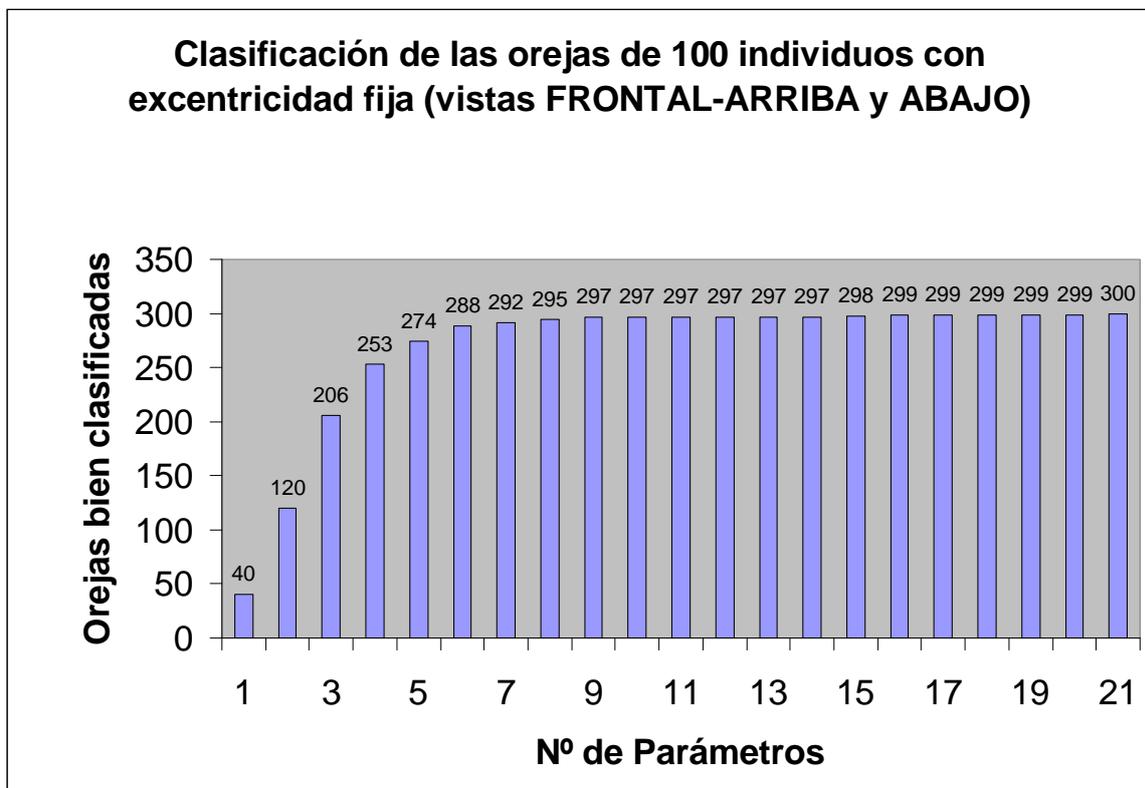


Figura 7-12: En este gráfico se ilustra el proceso de clasificación de parámetros a partir de las vistas FRONTAL-ARRIBA y ABAJO. Como puede observarse resulta necesario utilizar 21 parámetros para conseguir clasificar correctamente todas las vistas.

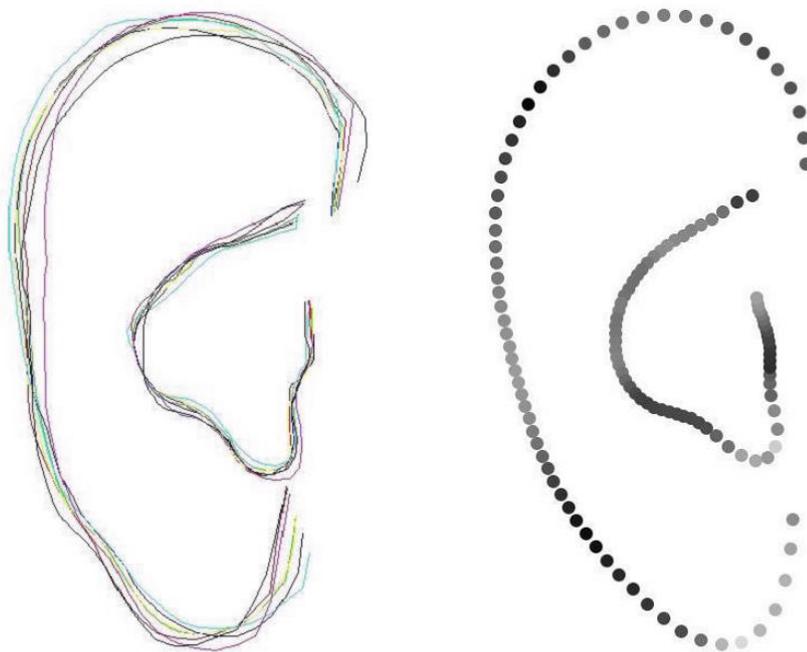


Figura 7-13: A la izquierda se ilustran las 7 vistas de la oreja de un individuo después de la normalización de la excentricidad. A la derecha se ilustra la oreja media contando todos los individuos y todas las vistas. Cuanto más oscuro es el nivel de gris, mas significativa es la medida (s_k es más pequeño)

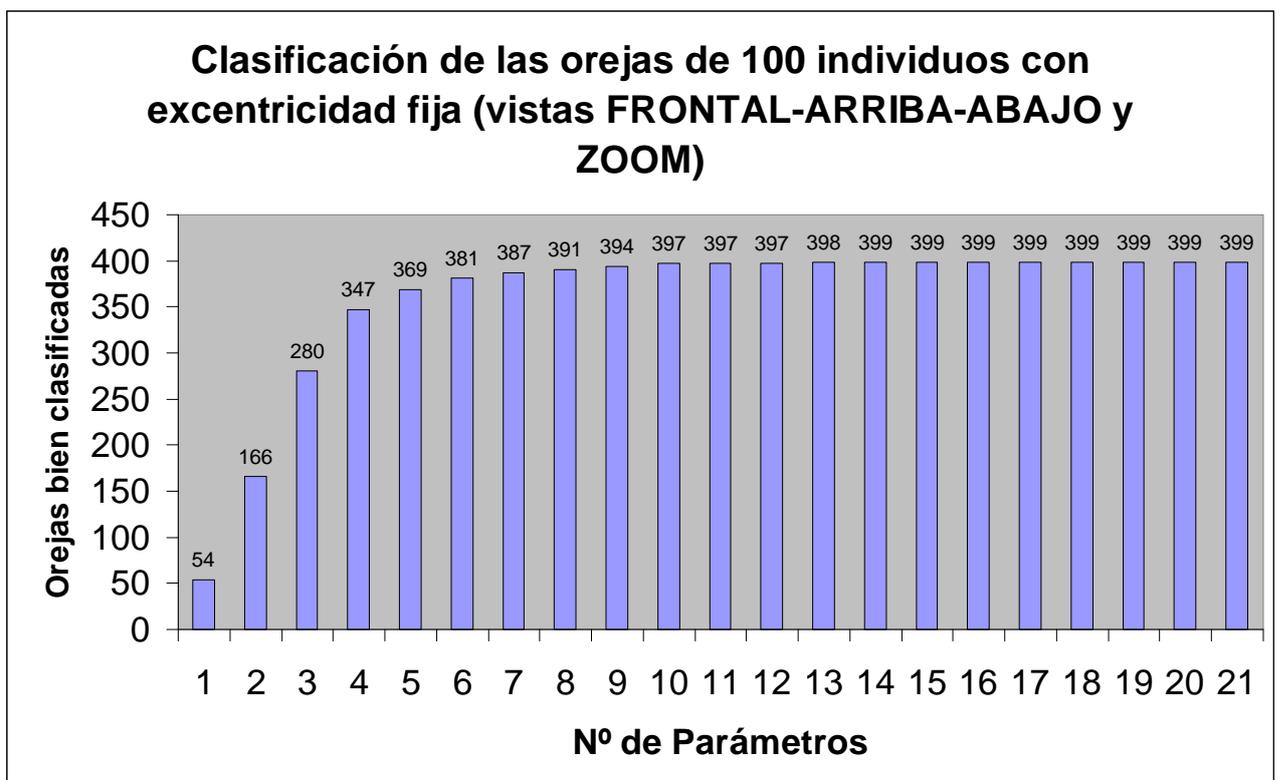


Figura 7-14: En este gráfico se ilustra el proceso de clasificación de parámetros a partir de las vistas FRONTAL-ARRIBA, ABAJO y ZOOM.

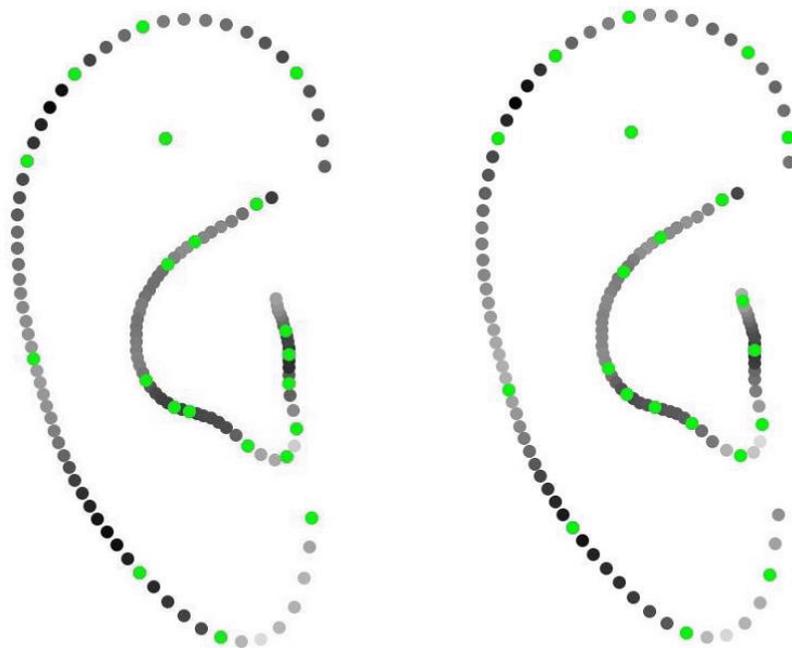


Figura 7-15: A la izquierda, en verde, parámetros seleccionados utilizando las vistas *FRONTAL-ARRIBA-ABAJO*. A la derecha lo mismo utilizando además la vista *ZOOM*. El círculo verde en el interior de la oreja significa que el parámetro de distorsión del ovoide también ha sido seleccionado

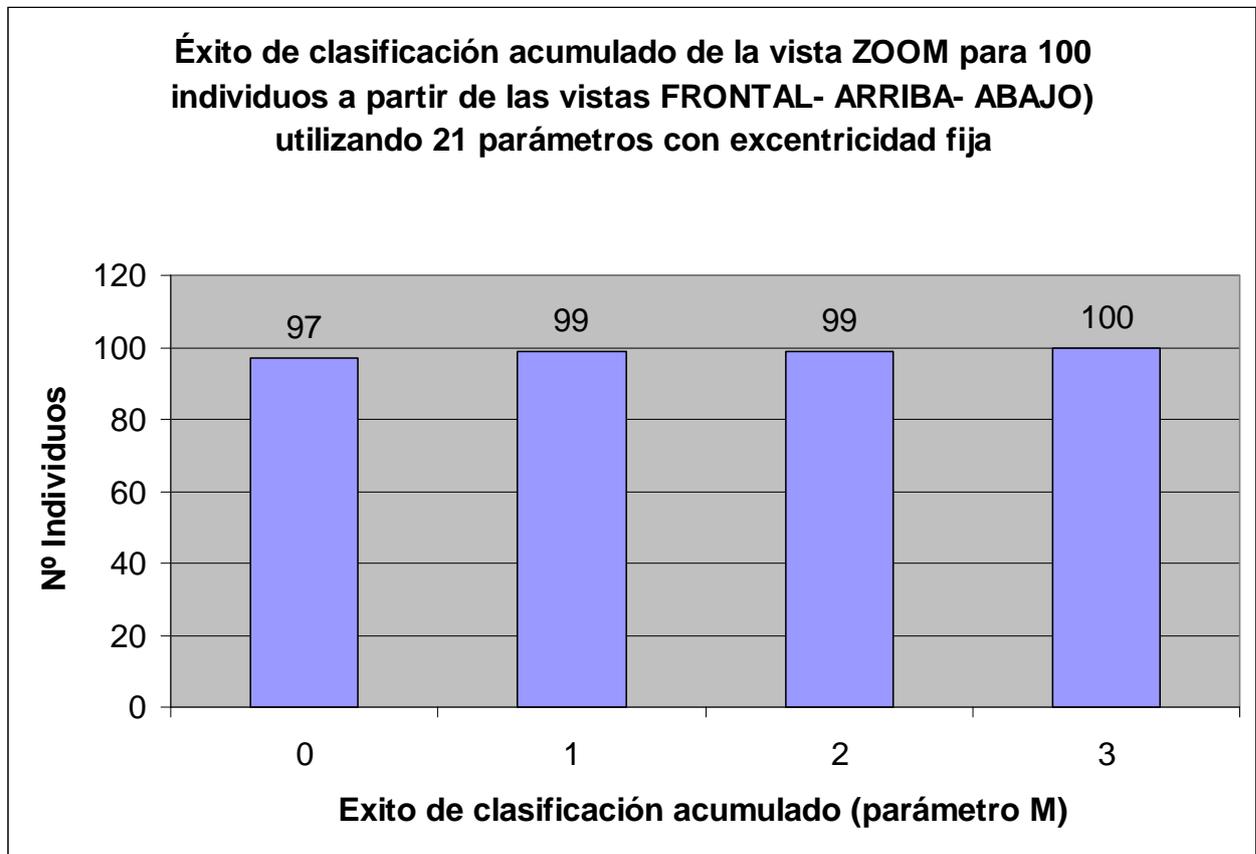


Figura 7-16: Ilustración del éxito de clasificación de la vista ZOOM a partir de las vistas FRONTAL-ARRIBA y ABAJO

individuos.

En la figura 7-17 se muestra el éxito de clasificación de la vista POSTERIOR a partir de las vistas FRONTAL-ARRIBA-ABAJO y ZOOM. 42 vistas han sido correctamente clasificadas y 92 vistas quedan bien clasificadas con un margen de error de 20 individuos.

En la figura 7-18 se muestra el éxito de clasificación de la vista DERECHA a partir de las vistas FRONTAL-ARRIBA-ABAJO y ZOOM. 38 vistas han sido correctamente clasificadas y 91 vistas quedan bien clasificadas con un margen de error de 20 individuos.

En la figura 7-19 se muestra el éxito de clasificación de la vista IZQUIERDA a partir

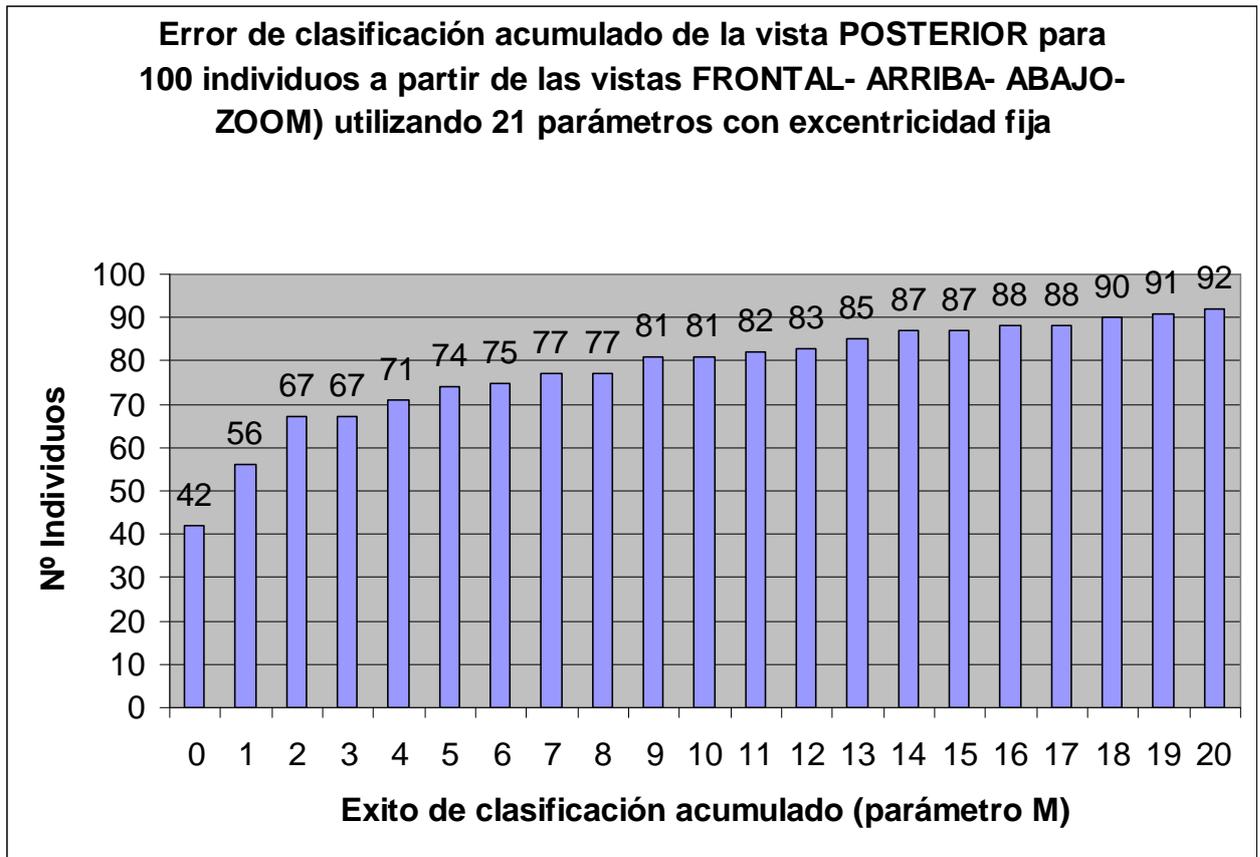


Figura 7-17: Exito de clasificación de la vista POSTERIOR a partir de las vistas FRONTAL-ARRIBA-ABAJO y ZOOM

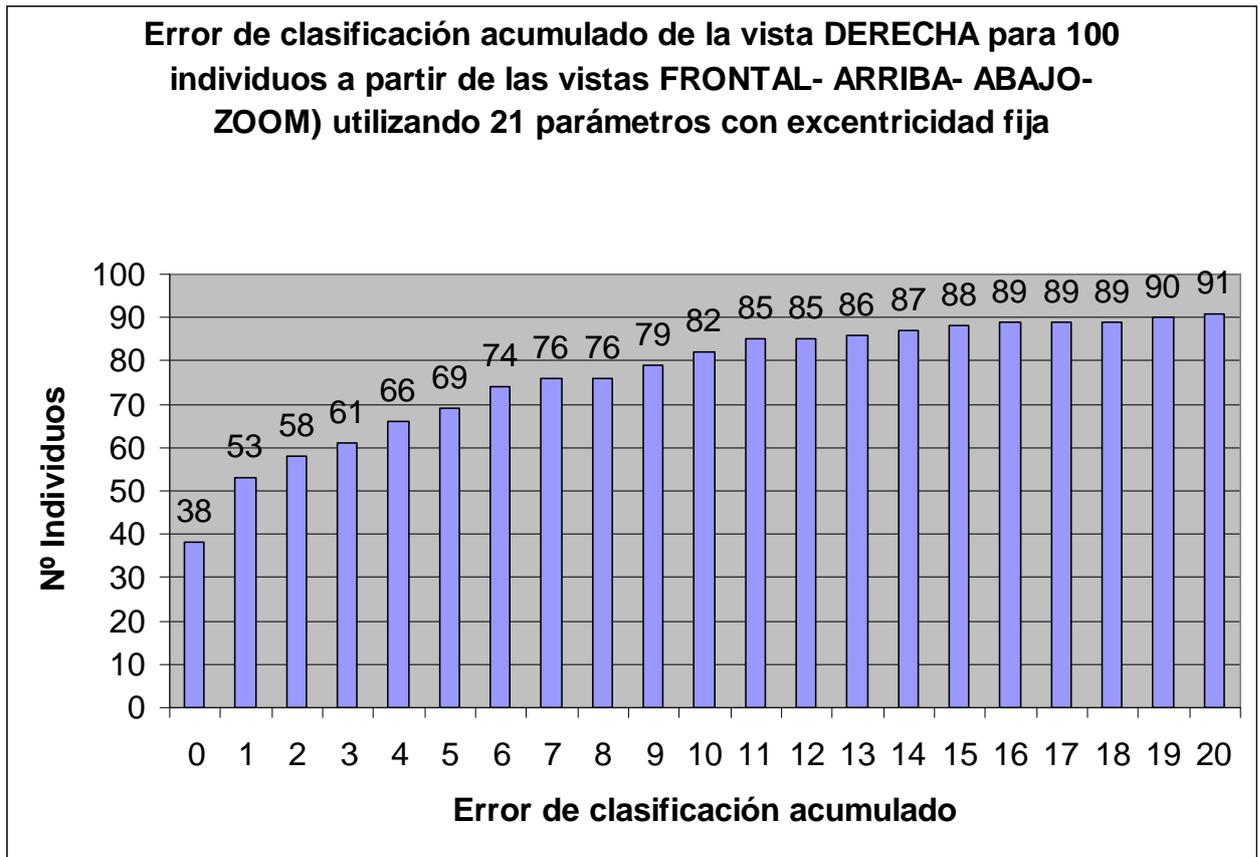


Figura 7-18: *Exito de clasificación de la vista DERECHA a partir de las vistas FRONTAL-ARRIBA-ABAJO y ZOOM*

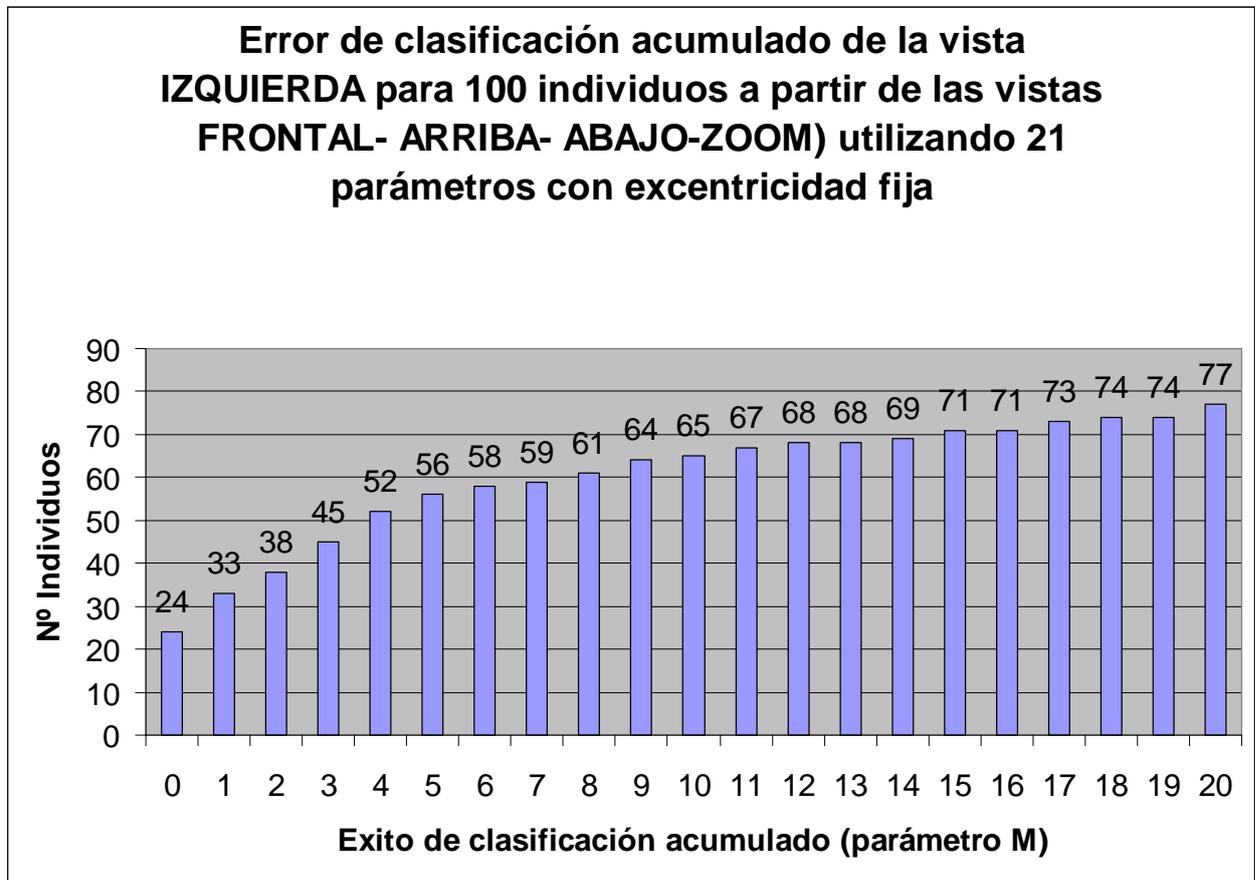


Figura 7-19: *Exito de clasificación de la vista IZQUIERDA, a partir de las vistas FRONTAL-ARRIBA-ABAJO y ZOOM*

de las vistas FRONTAL-ARRIBA-ABAJO y ZOOM. 24 vistas han sido correctamente clasificadas y 77 vistas quedan bien clasificadas con un margen de error de 20 individuos.

También hemos realizado un estudio, similar al explicado, utilizando la técnica del Análisis Discriminante Lineal Clásico. En la próxima sección comentamos el funcionamiento de esta técnica, así como los resultados obtenidos.

7.2. Análisis Discriminante Clásico

El análisis discriminante puede considerarse una técnica multivariante de clasificación de individuos en la que se presupone la existencia de dos o más grupos bien definidos a priori (por ejemplo, clientes solventes y no solventes; votantes de uno u otro partido; etc) y se persiguen los siguientes objetivos:

1. Analizar si existen diferencias entre los grupos en base a los valores que toman ciertas variables consideradas sobre los individuos de cada uno de los grupos y averiguar en qué sentido se dan dichas diferencias.
2. Elaborar procedimientos de clasificación sistemática de individuos de origen desconocido, en uno de los grupos analizados, en función de los valores que toman ciertas variables para esos individuos. Es decir, encontrar una combinación lineal de las variables independientes que mejor permite diferenciar (discriminar) a los grupos, (función discriminante), que se utilizará para clasificar nuevos casos.

Estos dos objetivos dan lugar a dos ramas dentro del Análisis Discriminante: el *Análisis Discriminante Descriptivo* y el *Análisis Discriminante Predictivo*, respectivamente.

En nuestro caso, los grupos definidos a priori (que son 100) son los que vienen determinados por cada uno de los individuos de los que se tomaron fotos de la oreja con distinta orientación. Nuestro objetivo inicial es, al igual que en la sección anterior, tratar de determinar qué variables o parámetros medidos, caracterizan mejor las diferencias existentes entre todos estos grupos. El punto de partida de nuestro análisis es el conjunto de las variables formado por la excentricidad y la distorsión del ovoide que mejor aproxima al borde de la oreja, así como un total de 64 distancias medidas, desde el centroide del grupo de puntos que se considera como contorno interno, a estos mismos puntos, y otras 64 distancias medidas, desde el mismo centroide, a los puntos que se consideran el contorno externo. Consideramos, por tanto, un total de 130 variables.

Nuestro objetivo será, entonces, identificar, en la medida de lo posible, las características que definen a los 100 individuos, utilizando para ello la información que nos proporcionan las 130 variables, y después, determinar cuáles de esas variables son las que más contribuyen a diferenciar entre uno y otro grupo. Luego trataremos de obtener las funciones discriminantes que, incluyendo sólo las variables seleccionadas previamente, nos permitan realizar el proceso de clasificación de las imágenes como pertenecientes o no a un determinado individuo. Para realizar este Análisis Discriminante hemos utilizado un paquete estadístico para ordenador, concretamente, la versión 14.0 del SPSS.

7.2.1. Método de selección de variables

En el cálculo de las funciones discriminantes antes comentado, pueden considerarse todas las variables (X_1, \dots, X_p) incluidas simultáneamente, sin considerar la capacidad discriminatoria de cada una de ellas. Sin embargo, si se utiliza un método de selección de variables para definir las funciones discriminantes, las variables (X_1, \dots, X_p) se irán incluyendo en el modelo una a una, según sea su capacidad discriminatoria, de la misma forma que lo hicimos en el estudio de la sección anterior. De esta forma, es muy probable que, finalmente, no todas las variables originales, 130 en nuestro caso, se utilicen para construir las funciones discriminantes. Únicamente se incluirán aquellas que más contribuyan a separar los grupos, con respecto a un cierto criterio, y se desestimarán aquellas cuya contribución sea escasa, bien porque no son relevantes, o porque la información que aportan ya esté recogida por alguna (o varias) de las variables ya incluidas en el modelo y resultan redundantes.

Entre los criterios para decidir qué variables entran o no en la definición de la función discriminante, cabría citar, entre otros, los siguientes: minimizar el estadístico Lambda de Wilks, minimizar la distancia de Mahalanobis entre los centroides de cada grupo, o maximizar el estadístico V de Rao. El más habitual, y el que empleamos en este

trabajo, es el de minimizar el estadístico Lambda de Wilks. Este estadístico se define, como un cociente entre la varianza “dentro de grupos” y la varianza “total”, y se basa en un resultado muy conocido en Estadística (ver 7.1), que descompone la variabilidad total de una variable como una suma de dos términos. De esta forma, el valor del cociente definido por el estadístico Lambda de Wilks representaría la proporción de la variabilidad total de las puntuaciones discriminantes debida a la diferencia “dentro de grupos” o, dicho de otra forma, la proporción de la variabilidad no explicada por las diferencias “entre grupos”, de forma similar a como lo hacía el estadístico s_k en el análisis estadístico de los datos de la sección anterior.

Teniendo en cuenta esta interpretación, parece obvio que el criterio para seleccionar una variable será minimizar este estadístico, es decir, conseguir que la variabilidad no explicada por las diferencias “entre grupos” sea mínima, o lo que es lo mismo, que las diferencias observadas entre los valores de las funciones discriminantes se deban fundamentalmente a diferencias existentes entre los grupos y no a diferencias internas dentro de ellos. Esto significaría que los grupos son muy diferentes entre sí, están muy separados, y ése es, recordémoslo, nuestro objetivo final: discriminar lo más posible entre los grupos. Además, con este criterio se consigue, simultáneamente, que los grupos sean internamente muy homogéneos, algo que, por ejemplo, no se tiene en cuenta si se utiliza como criterio de selección el de maximizar la distancia de Mahalanobis entre grupos.

Una vez establecido el criterio para decidir qué variable entra en el modelo en cada etapa, existen varias opciones sobre cómo realizar la selección en las distintas etapas. El método que se utiliza en este trabajo, con el fin de incluir en el modelo únicamente aquellas variables con mayor poder discriminatorio de entre las 130 variables disponibles, conocido como selección paso a paso (Stepwise en inglés) procede de la siguiente forma:

- En un principio se selecciona la variable que proporciona la máxima discrimi-

nación entre los grupos con el criterio elegido. Esta selección se hace contrastando, para cada variable, la hipótesis de igualdad de medias en todos los grupos, para lo cual se utiliza un test de la F que, en este caso, es una función del estadístico Λ de Wilks de la forma:

$$F = \frac{(n - g)}{(g - 1)} \left(\frac{1 - \Lambda}{\Lambda} \right)$$

donde n es el número de individuos (700 en nuestro caso) y g es el número de grupos (100 en nuestro caso). La primera variable seleccionada será la que minimice el Λ de Wilks, o equivalentemente, la que proporcione un valor del estadístico F mayor y significativo.

- En el segundo paso, la variable seleccionada se empareja con cada una de las $p - 1$ variables restantes, y se elige aquella pareja que produce la mayor discriminación (el mayor valor estadísticamente significativo del correspondiente estadístico F). En este caso, en que una variable se añade a un modelo que ya contiene otra(s) variable(s), el valor de F tiene la siguiente expresión:

$$F = \frac{(n - g - s)}{(g - 1)} \left(\frac{1 - \frac{\Lambda_{s+1}}{\Lambda_s}}{\frac{\Lambda_{s+1}}{\Lambda_s}} \right) \quad (7.9)$$

donde s es el número de variables previamente incluidas y Λ_s y Λ_{s+1} son los valores de la Lambda de Wilks antes y después de añadir la nueva variable, respectivamente. El estadístico F en 7.9, que se distribuye según una F con $(g - 1, n - s - g + 1)$ grados de libertad, representa el incremento producido en la discriminación tras la incorporación de la nueva variable con respecto al total alcanzado con la(s) variable(s) ya incluida(s) previamente. Por tanto, la variable elegida será la que tenga un mayor valor de F , es decir, la que más incrementa la capacidad discriminante.

- La tercera y posteriores variables se seleccionarían de forma similar, pero comprobando, tras cada etapa, si las variables seleccionadas previamente siguen siendo significativas o si, por el contrario, alguna de ellas puede excluirse. En ciertos

casos, una variable puede ser muy útil al principio del proceso de selección pero, después de haber incluido otras tantas variables, puede perder capacidad discriminadora y dejar de ser significativa porque, por ejemplo, la información que contiene resulta redundante. Esta variable sería entonces eliminada del modelo antes de proseguir con la selección de la siguiente variable. Para decidir si una variable se elimina se utiliza nuevamente un test de la F de la forma 7.9 que, en este caso, representaría el descenso que se produce en la discriminación debido a la supresión de la variable que sale.

Como vemos, con este método de selección paso a paso, en cada etapa puede entrar, y también salir, una variable, dependiendo del valor del correspondiente estadístico F asociado al estadístico Lambda de Wilks en esa etapa o, en general, al estadístico que se utilice como criterio de selección. El proceso de entrada/salida de variables termina cuando ninguna de las que están fuera del modelo cumple el criterio de entrada o cuando la variable seleccionada para entrar es una de las que se acaban de eliminar o, finalmente, cuando ninguna de las variables del modelo cumple el criterio de salida.

En este modelo de Análisis Discriminante no tenemos un control directo sobre los parámetros que podemos ir incluyendo cada vez, sino sólo sobre los criterios estadísticos que deben cumplir los parámetros para que sean incluidos o excluidos del modelo. De esta manera, para que los resultados obtenidos con este método fueran comparables a los obtenidos en la sección anterior, en la que estudiábamos la capacidad clasificadora del modelo obtenido incluyendo un parámetro adicional cada vez, hemos utilizado como criterio de entrada y salida de variables, el de minimizar el valor del estadístico Lambda de Wilks, fijando un valor mínimo para la entrada de variables y un valor máximo para la salida de las mismas. Hemos garantizado siempre que el valor F-mínimo para entrar fuera mayor que el valor F-máximo para salir, evitando así que una variable pudiera estar entrando y saliendo del modelo anterior, de forma indefinida.

Fijando los valores de F-mínimo y F-máximo, y partiendo de un modelo que in-

cluyera sólo un parámetro, hemos ido forzando la inclusión de un nuevo parámetro adicional cada vez, y comprobando la capacidad clasificadora de las funciones discriminantes obtenidas con estos parámetros seleccionados. En algunos casos, no hemos logrado que se incluyera, de una vez a la siguiente, un único parámetro adicional, debido a que, en el paso correspondiente, había más de un parámetro que cumplía el criterio estadístico de entrada. Esto es lo que provoca que, en las gráficas que pasamos a comentar, existan algunos "huecos".

Comenzamos, como en el estudio anterior, situándonos en el caso de utilizar las vistas FRONTAL-ARRIBA-ABAJO, con excentricidad variable, para la selección de parámetros. En la figura 7-20 se ilustra el número de vistas que quedan correctamente clasificadas, en función del número de parámetros manejados. Como podemos apreciar, se requieren 8 parámetros para clasificar correctamente todas las vistas. No aparecen barras sobre el 3 y el 4, indicando que, partiendo del modelo que incluía 2 parámetros, cuando hemos modificado los valores de la F-mínima y F-máxima para tratar de incluir un tercero; puesto que el proceso de entrada/salida de variables termina cuando ninguna de las que están fuera del modelo cumple el criterio de entrada, no se ha incluido sólo un tercero, sino 3 parámetros más, pasando de 2 a 5 parámetros en el modelo, lo que provoca los huecos en los modelos que incluirían 3 y 4 parámetros.

En la figura 7-21 se ilustra la selección de parámetros cuando utilizamos las vistas FRONTAL, ARRIBA, ABAJO y ZOOM, con excentricidad variable. En este caso resultan necesarios 13 parámetros para clasificar correctamente las 400 vistas.

7.2.2. Obtención de las funciones discriminantes

Técnicamente, podemos decir que el análisis discriminante tratará de encontrar funciones de todas estas 130 variables (si se consideraran todas incluidas simultáneamente) o de algunas de ellas (si se considera el proceso de selección paso a paso), cuyos valores separen o discriminen lo más posible a los 100 grupos existentes. Estas fun-

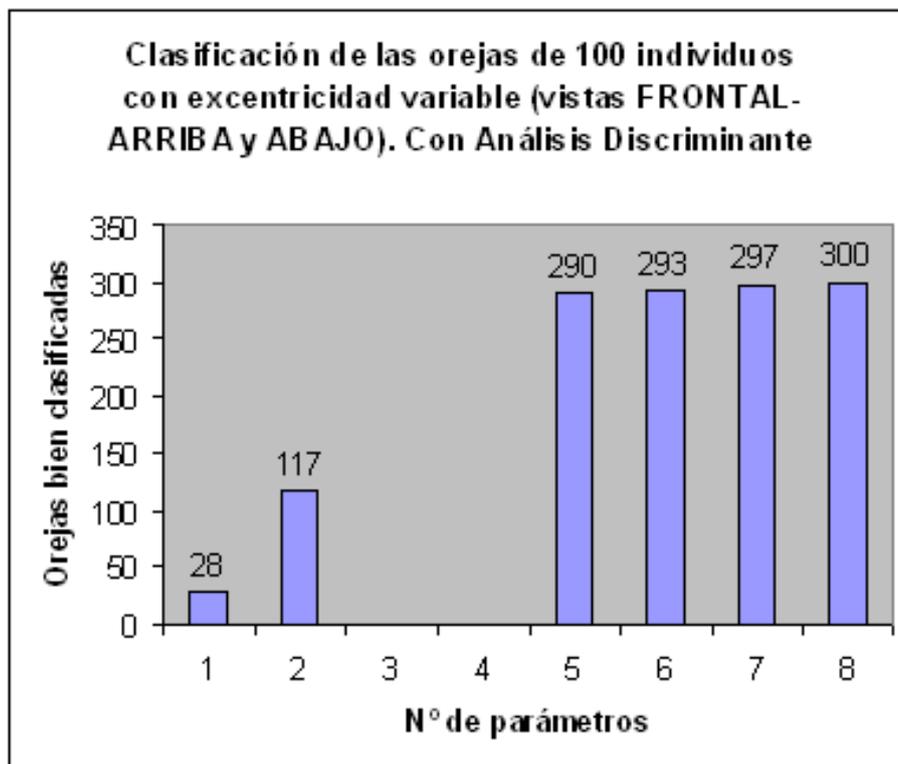


Figura 7-20: Resultado de repetir el proceso de selección de parámetros, forzando la inclusión de un parámetro adicional cada vez, tomando las vistas FRONTAL, ARRIBA y ABAJO, con excentricidad variable.

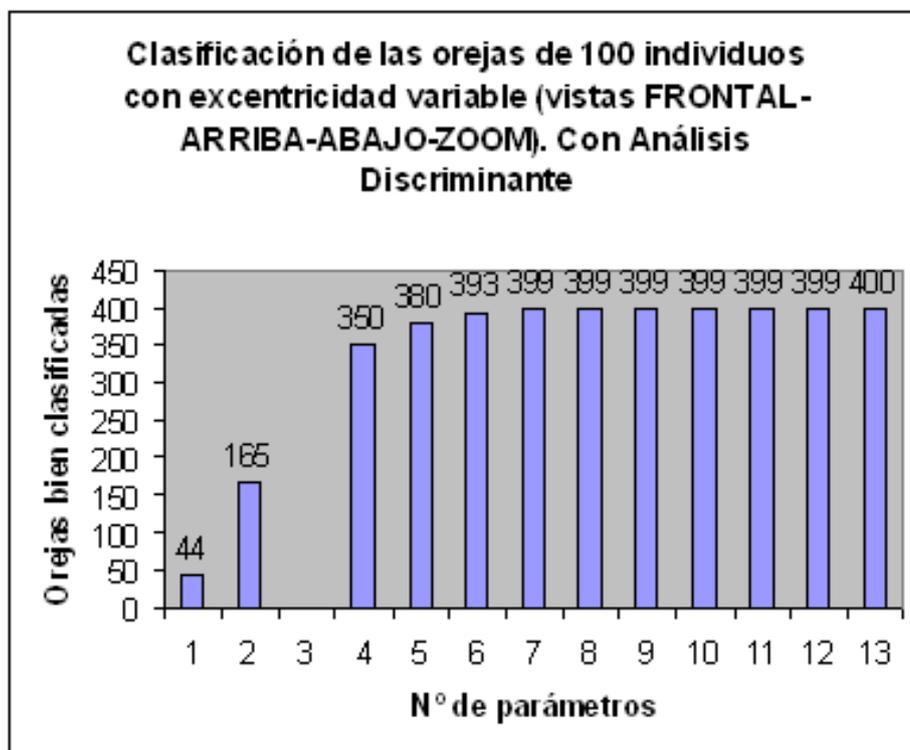


Figura 7-21: Resultado de repetir el proceso de selección de parámetros, forzando la inclusión de un parámetro adicional cada vez, tomando las vistas FRONTAL, ARRIBA, ABAJO y ZOOM, con excentricidad variable.

ciones, denominadas funciones o ejes discriminantes, serán combinaciones lineales de las variables originales de la forma:

$$\mathbf{Y} = a_0 + a_1X_1 + a_2X_2 + \dots + a_pX_p$$

donde p es el número de variables explicativas, en nuestro caso $p = 130$, y los coeficientes $\{a_0, a_1, \dots, a_p\}$ se eligen de tal forma que se consiga la máxima separación entre los 100 grupos existentes, es decir, tratando de que los valores que toman estas funciones discriminantes \mathbf{Y} en los 100 grupos sean lo más diferentes posibles.

Estadísticamente, este criterio equivale a maximizar la varianza “entre grupos” frente a la varianza “dentro de grupos”. Por tanto, los coeficientes $\{a_0, a_1, \dots, a_p\}$ se elegirán de tal forma que se consiga maximizar el valor del cociente:

$$\lambda = \frac{\text{Varianza entre grupos}}{\text{Varianza dentro de grupos}}$$

Si la varianza “entre grupos” es grande, es decir, si hay grandes diferencias entre los valores que toma la función \mathbf{Y} en los distintos grupos, pero la varianza “dentro de grupos” es pequeña, es decir, los valores de \mathbf{Y} para las imágenes de un mismo individuo son muy similares, entonces diremos que la función discriminante separa bien a los grupos, que serán, internamente muy homogéneos y, a la vez, muy diferentes entre sí.

Aunque no vamos a entrar aquí en desarrollos formales sobre la obtención de los coeficientes que definen las funciones discriminantes, sí es necesario plantear algunas cuestiones importantes de cara al análisis e interpretación de los resultados obtenidos. En primer lugar, hay que señalar que, el número de funciones que pueden obtenerse, es el mínimo entre el número de variables explicativas disponibles y el número de grupos menos uno. En nuestro caso, al ser 100 los grupos en los que se clasifica la variable numérica que nos ocupa (identificador del individuo), y 130 las variables explicativas, se

pueden definir únicamente 99 funciones discriminantes. Como además hemos utilizado el método de selección de parámetros paso a paso, serán menos las variables consideradas y , por tanto, menos las funciones discriminantes calculadas. Estas funciones se obtienen de forma sucesiva en función de su capacidad discriminatoria. Así, la primera función discriminante, que será de la forma:

$$\mathbf{Y}_1 = a_{01} + a_{11}X_1 + a_{21}X_2 + \dots + a_{p1}X_p \quad (7.10)$$

será la que tenga mayor poder discriminatorio, es decir, la que mejor separe los 100 grupos. La segunda función, que vendrá definida por:

$$\mathbf{Y}_2 = a_{02} + a_{12}X_1 + a_{22}X_2 + \dots + a_{p2}X_p \quad (7.11)$$

será la siguiente en capacidad discriminatoria y, además, estará incorrelacionada con la función anterior \mathbf{Y}_1 , y así sucesivamente. A los valores de estas funciones para cada uno de los sujetos de la población [las 700 imágenes tomadas (7 por individuo), en nuestro caso] se les denomina puntuaciones discriminantes. Es decir, al final del análisis discriminante obtendremos 700 valores (puntuaciones discriminantes) de cada función discriminante, uno por cada imagen.

En ciertas ocasiones, la capacidad discriminatoria de la primera función, \mathbf{Y}_1 , es tan grande, que la información añadida por las demás funciones apenas es relevante y se ignoran, ya que su contribución a la separación entre los grupos no es significativa. Para determinar cuántas funciones discriminantes son significativas existen varios criterios, entre los cuales, los más utilizados son el del porcentaje relativo, el coeficiente de correlación canónica eta (η) y el estadístico Lambda de Wilks (Λ), que es el que hemos utilizado en nuestro trabajo.

El estadístico Λ de Wilks es también una medida de las diferencias entre los grupos, debidas a las funciones discriminantes, que se utiliza para medir de forma secuencial el poder discriminatorio de cada una de las funciones que se van construyendo, em-

pezando siempre por la primera, que es la de mayor capacidad discriminadora. Así, en cada etapa, se plantea la siguiente cuestión: ¿merece la pena incluir una nueva variable o parámetro como eje discriminante? La respuesta será afirmativa si aún existen diferencias significativas entre los grupos, y será negativa si la separación conseguida ya es suficiente.

Aunque no se presente aquí la expresión de Λ , es importante saber que sus valores varían entre 0 y 1: valores próximos a 0 indican mucha discriminación, es decir los grupos están muy separados, mientras que valores cercanos a 1 representan escasa discriminación o poca diferencia entre los grupos. No obstante, es preferible apoyar esta interpretación del valor de Λ en una prueba estadística que permita decidir sobre la significación del valor de este estadístico. Para ello se suele recurrir al estadístico V de Barlett, que es una función de Λ y tiene una distribución asintótica Chi-cuadrado conocida y tabulada. Valores de este estadístico tras haber introducido k ejes discriminantes (V_k) grandes, superiores al valor crítico correspondiente, rechazan la hipótesis de igualdad entre grupos y, por tanto, indican la pertinencia de obtener una función discriminante más. Por el contrario, valores pequeños de V_k indican que las funciones extraídas son ya suficientes y que las restantes no contribuyen a aumentar la discriminación.

Examinando todos estos coeficientes que acabamos de ver podemos determinar cuántas funciones discriminantes son sustancialmente significativas y qué utilidad tiene cada una de ellas para explicar las diferencias entre los grupos. En el paquete estadístico que hemos utilizado para nuestro trabajo, estos valores de muestran al usuario en varias tablas, una de las cuales es la tabla de autovalores, como la mostrada en 7-22, que corresponde a la clasificación de las vistas ARRIBA, ABAJO y FRONTAL, con excentricidad variable, incluyendo 8 parámetros (que es cuando quedan completamente clasificadas las vistas).

En la tabla 7-22 vemos que la capacidad discriminante de la primera función, con respecto al total acumulado por las 8 funciones consideradas, es del 25.1%, mientras

que el porcentaje relativo correspondiente a la segunda función es del 23.9%, y así sucesivamente. Estos valores confirman que la importancia relativa de la primera función (la que más discrimina) es superior a las demás.

Los valores del coeficiente de correlación canónica *eta* indican qué función cumple mejor el objetivo de separar los grupos. Los valores de *eta* al cuadrado admiten otra interpretación complementaria en términos de varianza. Así, la proporción de varianza explicada por la primera función es $0,987^2 \times 100 = 97,42$ por ciento. Esto significa que con una única función discriminante (construida con las 8 variables seleccionadas previamente) recogemos el 97.4% de la varianza debida a las diferencias entre los grupos. La segunda función explicaría el $0,987^2 \times 100 = 97,42$ por ciento de la varianza restante (el 2.58 por ciento no explicado por la primera función) y así sucesivamente. Por tanto, la varianza explicada por las dos primeras funciones es del $97,42 + (97,42 \times 0,0258) = 99,933$ por ciento de la variación total de la variable dependiente (el individuo). Es decir, con las primeras dos funciones discriminantes se explica el 99.93% de las diferencias existentes entre los individuos.

Los valores del estadístico Λ de Wilks, que también devuelve el proceso en una tabla (ver 7-23), sirven para contrastar secuencialmente la significación de cada función discriminante, e indica cuáles de las funciones discriminantes sería conveniente retener para interpretar las diferencias entre los grupos (las que sean significativas). En el ejemplo que estamos comentando, las 8 funciones discriminantes resultan ser significativas (Sig.= 0,000), por lo que hay que considerarlas todas.

7.2.3. Interpretación de las funciones discriminantes

Una vez obtenidas las funciones discriminantes, nuestro objetivo es establecer la contribución relativa de las distintas variables a la discriminación, o lo que es lo mismo, determinar cuáles son las variables que más contribuyen a discriminar entre un grupo y otro (de entre las variables incluidas en la función discriminante). En nuestro caso,

Autovalores

Función	Autovalor	% de varianza	% acumulado	Correlación canónica
1	39,059(a)	25,1	25,1	,987
2	37,250(a)	23,9	49,0	,987
3	27,471(a)	17,7	66,7	,982
4	16,600(a)	10,7	77,4	,971
5	12,373(a)	8,0	85,3	,962
6	10,138(a)	6,5	91,8	,954
7	9,435(a)	6,1	97,9	,951
8	3,291(a)	2,1	100,0	,876

a Se han empleado las 8 primeras funciones discriminantes canónicas en el análisis.

Figura 7-22: Autovalores de las funciones discriminantes utilizadas

Lambda de Wilks

Contraste de las funciones	Lambda de Wilks	Chi-cuadrado	gl	Sig.
1 a la 8	,000	5477,342	792	,000
2 a la 8	,000	4573,207	686	,000
3 a la 8	,000	3680,390	582	,000
4 a la 8	,000	2859,910	480	,000
5 a la 8	,000	2157,274	380	,000
6 a la 8	,002	1521,933	282	,000
7 a la 8	,022	931,402	186	,000
8	,233	356,828	92	,000

Figura 7-23: Valores del estadístico Lambda de Wilks para las funciones discriminantes utilizadas

puesto que hemos ido incluyendo variables hasta conseguir clasificar correctamente todas las vistas, hemos de suponer que todas las variables finalmente incluidas en el modelo contribuyen de alguna forma a discriminar entre los individuos, pero se trataría de averiguar con que fuerza contribuye cada una de ellas a esa discriminación. Es decir, se trata de averiguar cuáles son las variables que más determinan la pertenencia de una imagen determinada a uno u otro individuo.

Entre los criterios más utilizados a este respecto cabría destacar los siguientes: posición de los centroides de grupo, coeficientes de correlación o de estructura y los coeficientes estandarizados (que es la que utilizamos en el trabajo).

Una primera forma de establecer la relación entre las variables explicativas (X_1, \dots, X_p) y las funciones discriminantes, con el fin de determinar la contribución de cada variable a la discriminación, sería atender a los propios coeficientes (a_{1i}, \dots, a_{pi}) que definen a las funciones discriminantes de las ecuaciones 7.10 y 7.11. Por ejemplo, si en la función \mathbf{Y}_1 definida en 7.10, el coeficiente a_{11} que acompaña a la variable X_1 fuera mucho mayor que los de las otras variables, podríamos pensar que esta variable X_1 va a contribuir al valor final de \mathbf{Y}_1 en mayor medida que las otras variables, y por lo tanto es la más importante. Sin embargo, esto no es del todo cierto, ya que las variables (X_1, \dots, X_p) suelen medir magnitudes muy diferentes y/o venir expresadas en unidades distintas, lo cual dificulta su comparación. Piénsese, en nuestro caso, en variables como cualquiera de las distancias al contorno externo, que toma valores del orden de las centenas, comparadas con la variable que mide la excentricidad de la elipse, que toma valores entre 0 y 1. Para solventar este problema se utilizan, en lugar de los propios coeficientes (a_{1i}, \dots, a_{pi}) , otros coeficientes denominados estandarizados, que son los que definirían a las funciones discriminantes \mathbf{Y}_1 e \mathbf{Y}_2 si éstas se calcularan a partir de las variables (X_1, \dots, X_p) estandarizadas. Las variables estandarizadas ya no dependen de unidades de medida y tienen todas la misma media, igual a 0, y la misma varianza, igual a 1, lo cual facilita su comparación. Por tanto, la magnitud de un coeficiente estandarizado, ignorando su signo, indicará la importancia relativa de la variable correspondiente en

el cálculo de la función discriminante y nos permitirá determinar qué variables, de las incluidas en las funciones discriminantes, contribuyen más a la discriminación: cuanto mayor sea el coeficiente, mayor será la contribución de esa variable.

En la imagen 7-24 aparecen señaladas con puntos verdes las variables incluidas en el modelo, cuando trabajamos con excentricidad variable, considerando 3 vistas (FRONTAL, ARRIBA, ABAJO) y 4 vistas (FRONTAL, ARRIBA, ABAJO y ZOOM) respectivamente. De las variables incluidas en el primero de los modelos (3 vistas), tenemos en 7-25 la tabla de los coeficientes estandarizados correspondientes.

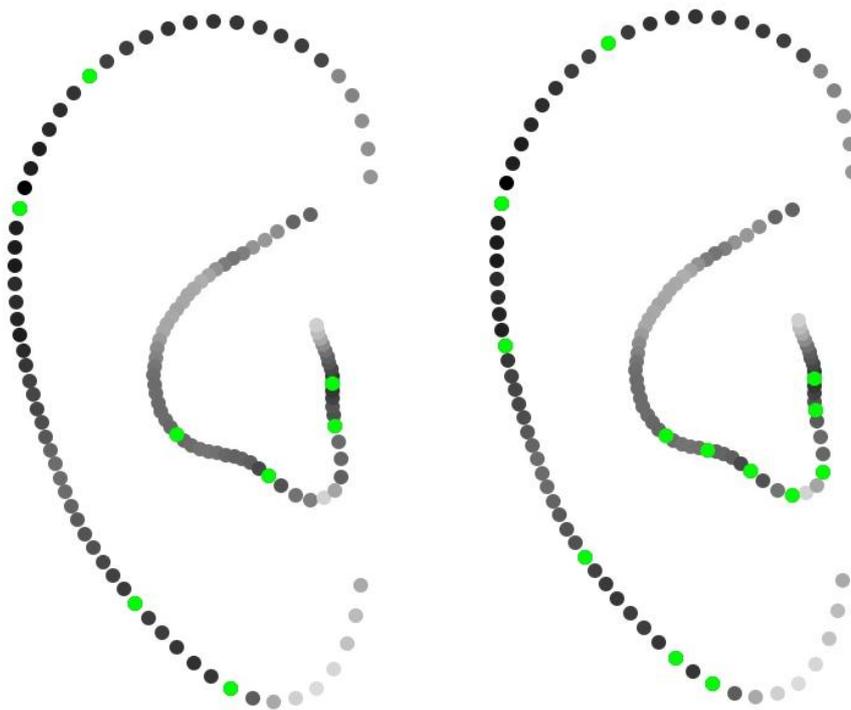


Figura 7-24: A la izquierda, en verde, parámetros seleccionados utilizando las vistas *FRONTAL-ARRIBA-ABAJO*. A la derecha lo mismo utilizando además la vista *ZOOM*.

Sin embargo, los coeficientes estandarizados adolecen también de ciertos problemas. Por ejemplo, si dos variables aportan la misma discriminación y están muy correlaciona-

Coeficientes estandarizados de las funciones discriminantes canónicas

	Función							
	1	2	3	4	5	6	7	8
dext16	1,611	,344	-1,348	-,825	,056	-,133	-,173	,601
dext23	-1,149	-1,347	,381	1,141	,491	,087	,043	-,270
dext49	-,917	,225	,191	-,837	-,148	,857	-,957	,227
dext55	-,444	,556	,173	,392	,065	-,824	1,155	,357
dint29	,516	,059	,258	,555	-,792	-,077	-,400	,339
dint40	,565	,258	,304	,714	,557	,401	,080	-,191
dint49	-,147	,052	,061	-,126	-,220	,830	,969	,226
dint54	,323	,034	,886	-,537	,572	-,534	-,264	-,129

Figura 7-25: En la tabla se incluyen los coeficientes estandarizados por los que se multiplicaría cada una de las 8 variables incluidas en el modelo, para dar lugar a cada una de las 8 funciones discriminantes consideradas.

dos entre sí, sus coeficientes estandarizados son más bajos que si se usara sólo una de ellas. Las correlaciones se calculan para todas las variables y se devuelven en una matriz de estructura. Los coeficientes de estructura nos permiten determinar qué variables correlacionan más con una función discriminante concreta, tanto si estas se incluyen en la definición de dicha función como si no, y no están afectados por las relaciones con las restantes variables. En la matriz de estructura se señala, con un asterisco, para cada variable, cuál es la función con la que dicha variable está más correlacionada. Las variables con coeficientes de estructura más elevados (de entre las que correlacionan con cada función discriminante), serán los que den significado a cada función discriminante.

7.2.4. Evaluación del procedimiento de clasificación

En los apartados precedentes hemos comentado, básicamente, cómo construir o estimar las funciones discriminantes. Si nuestro objetivo hubiera sido averiguar en qué difieren principalmente los grupos, con lo visto hasta ahora tendríamos más que sufi-

ciente. Sin embargo, nuestro interés en las funciones discriminantes radica en su capacidad para clasificar nuevos casos. Ahora bien, debemos tener claro que, la clasificación de casos, es algo muy distinto a la estimación de la función discriminante. De hecho, una función discriminante perfectamente estimada podría no pasar de una pobre capacidad clasificadora. Aquí radica la principal diferencia con el método propuesto en el estudio anterior; los dos aspectos: selección y clasificación, en este caso, se realizan por separado.

Una vez obtenida la función discriminante podemos utilizarla, en primer lugar, para efectuar una clasificación de los mismos casos utilizados para obtener la función: esto permitirá comprobar el grado de eficacia de la función desde el punto de vista de la clasificación. Esto es lo que hemos hecho para determinar las variables a incluir en el modelo, en los casos de excentricidad fija y variable, considerando 3 y 4 vistas respectivamente (ARRIBA-ABAJO-FRONTAL y ARRIBA-ABAJO-FRONTAL-ZOOM).

Si los resultados son satisfactorios, la función discriminante podrá utilizarse, en segundo lugar, para clasificar futuros casos de los que, conociendo el valor de las variables independientes, se desconozca el grupo al que pertenecen.

El proceso de clasificación se realiza asignando, en primer lugar, a cada caso a clasificar, una puntuación d_i , haciendo uso de las funciones discriminantes calculadas, y una *probabilidad a priori* de pertenecer a cada grupo $P(g_k)$. Esta probabilidad a priori puede ser la misma para todos los grupos (independientemente del tamaño del grupo), con lo que se consigue que el tamaño de los grupos no influya en la clasificación; o bien, proporcional al tamaño del grupo. En este trabajo hemos utilizado la primera opción.

Luego se obtiene la probabilidad condicional $P(d_i / g_k)$ que representa la probabilidad de obtener un valor d_i particular si el caso observado pertenece a una grupo dado g_k . Por último se calcula una *probabilidad a posteriori* $P(g_k / d_i)$, es decir, la probabilidad de que el caso pertenezca a cada uno de los grupos, dado que, a ese caso, le

corresponde una determinada puntuación discriminante. Esta probabilidad a posteriori se calcula utilizando el Teorema de Bayes:

$$P(g_k / d_i) = \frac{P(d_i/g_k) P(g_k)}{\sum P(d_i/g_k) P(g_k)}$$

El sumatorio del denominador posee tantos términos como grupos. Con esta regla de clasificación, los casos nuevos son clasificados en el grupo al que corresponde una mayor probabilidad a posteriori.

Un problema habitual de los modelos estadísticos es que el modelo estimado siempre se ajusta lo más perfectamente posible a los datos de la muestra concreta utilizada. Esto, obviamente, constituye un pequeño inconveniente, pues la estructura de la muestra puede presentar ligeras divergencias respecto de la estructura real de la población. Para evitar este efecto de sobreajuste muestral puede llevarse a cabo una validación cruzada, que consiste en:

1. Seleccionar, de la muestra original, un subconjunto aleatorio de casos (muestra de validación)
2. Estimar las funciones discriminantes con los casos restantes (muestra de entrenamiento)
3. Utilizar esas funciones discriminantes para clasificar los casos de la muestra de validación.

La validación cruzada consiste, por tanto, en clasificar casos con una función que no incluye información sobre ellos. Puede llevarse a cabo una sola vez o repetirse varias veces. Si la muestra original es grande, puede bastar con un solo intento, utilizando una muestra de validación del 10 % al 20 % de los casos. Para realizarla debe crearse, primero, una variable (variable de selección) que distinga entre los casos que serán utilizados como muestra de entrenamiento y los que serán utilizados como muestra de

validación. Si se efectúa un Análisis Discriminante de esta manera, se obtiene como resultado las matrices de confusión, donde se indican las tasas de aciertos (en porcentajes) correspondientes a los casos seleccionados (muestra de entrenamiento) y a los no seleccionados (muestra de validación).

7.2.5. Resultados experimentales de la clasificación de vistas

En nuestro trabajo, hemos realizado varios casos de validación cruzada, que pasamos a comentar a continuación y, en cada uno de ellos, dado que la muestra original era bastante grande (700 vistas en total) hemos realizado un solo intento.

El primer estudio lo hemos hecho considerando excentricidad variable y partiendo de 400 vistas, divididas en dos grupos: una muestra de entrenamiento formada por las vistas ARRIBA, ABAJO y FRONTAL (300 en total), con las que se obtienen las funciones discriminantes, y una muestra de validación formada por las vistas ZOOM (100 vistas en total), que son las se pretenden clasificar. Los resultados obtenidos se muestran en la figura 7-26. Podemos apreciar que, con 8 parámetros (que eran los necesarios para clasificar bien el 100% de las vistas FRONTAL-ARRIBA-ABAJO, ver 7-20), se clasifican correctamente el 98% de las vistas ZOOM. Si seguimos añadiendo parámetros al modelo, se mejora un poco el porcentaje de clasificación, hasta el 99% pero, añadiendo hasta 13 parámetros (los que eran necesarios para clasificar bien el 100% de las vistas FRONTAL-ARRIBA-ABAJO-ZOOM cuando no se realizaba validación cruzada, ver 7-21) no logramos que esta clasificación mejore más.

A continuación pasamos a la situación más compleja donde partimos de 500 vistas, divididas en dos grupos: una muestra de entrenamiento formada por las vistas ARRIBA, ABAJO, FRONTAL y ZOOM (400 en total), con las que se obtienen las funciones discriminantes, y una muestra de validación formada por las vistas POSTERIOR (100 vistas en total), que son las se pretenden clasificar. Los resultados obtenidos se muestran en la figura 7-27. Podemos apreciar que, con 13 parámetros (que eran los necesarios

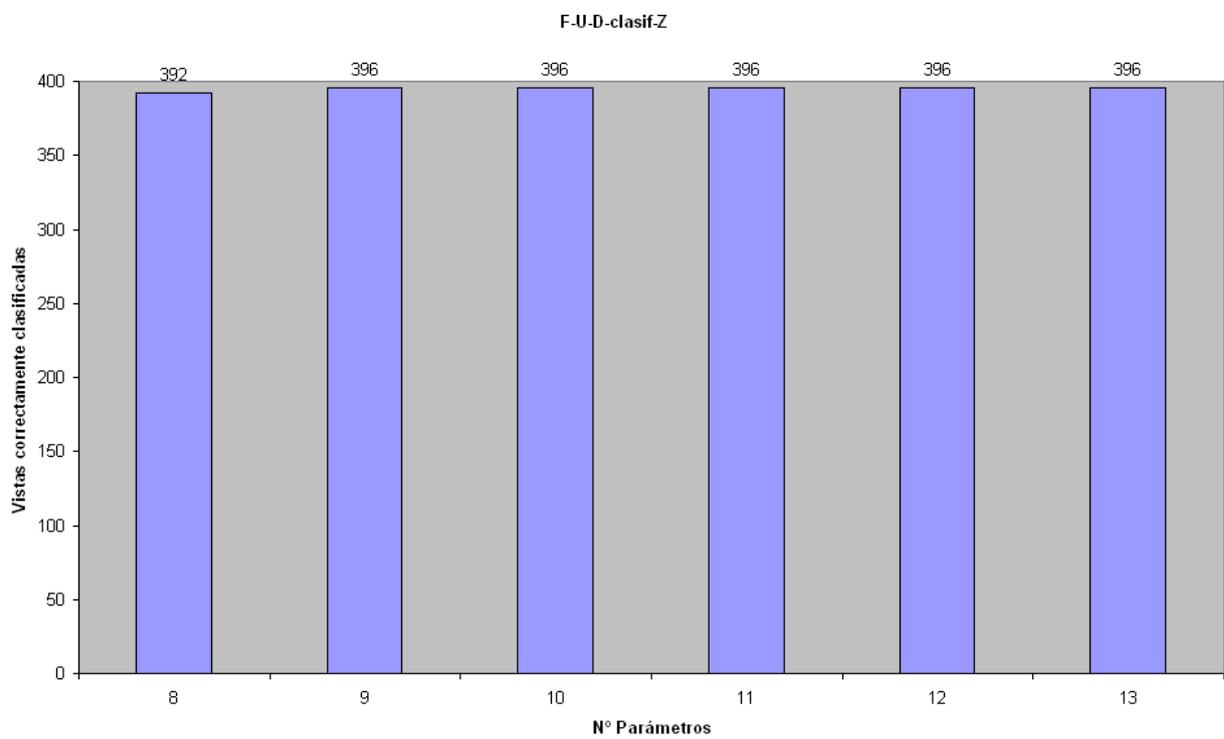


Figura 7-26: *Exito de clasificación de las vistas ZOOM a partir de las vistas FRONTAL-ARRIBA y ABAJO, con excentricidad variable, utilizando validación cruzada*

para clasificar bien el 100% de las vistas FRONTAL-ARRIBA-ABAJO-ZOOM, ver 7-21), se clasifican correctamente el 54% de las vistas POSTERIOR.

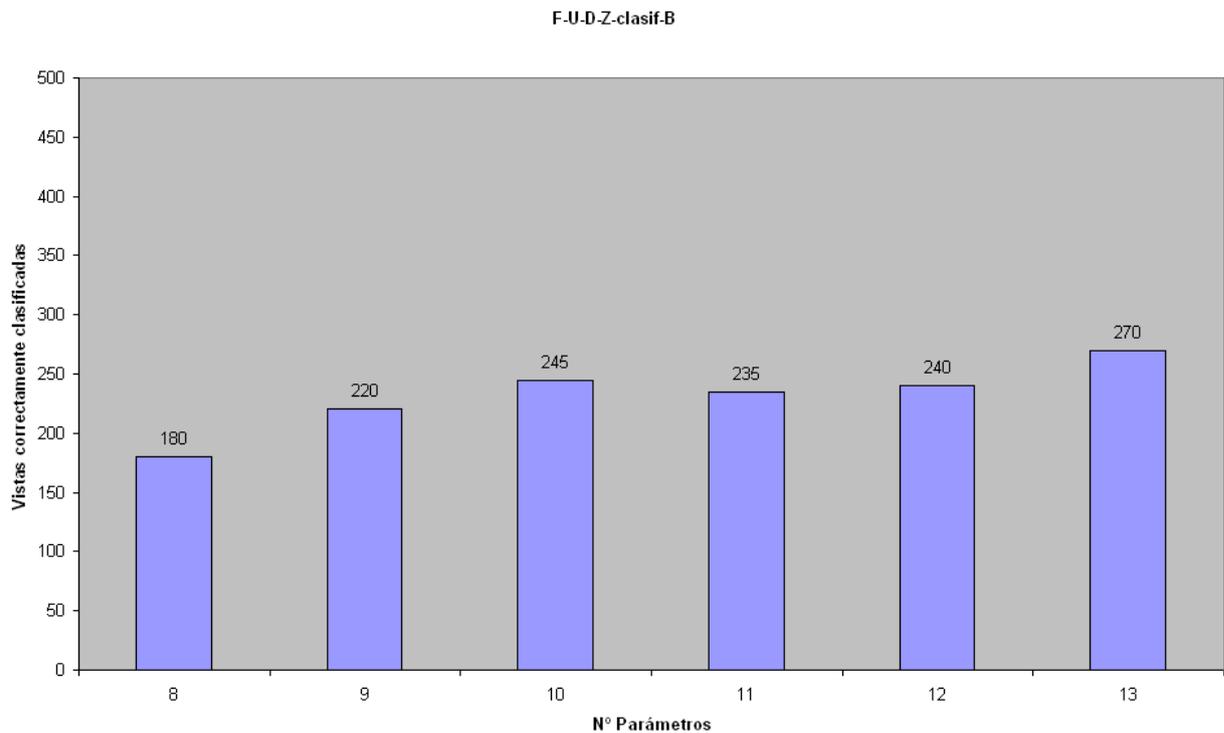


Figura 7-27: Exito de clasificación de las vistas POSTERIOR a partir de las vistas FRONTAL-ARRIBA-ABAJO y ZOOM, con excentricidad variable, utilizando validación cruzada

En el siguiente paso partimos también de 500 vistas, divididas en dos grupos: la misma muestra de entrenamiento que en el caso precedente, formada por las vistas ARRIBA, ABAJO, FRONTAL y ZOOM (400 en total), con las que se obtienen las funciones discriminantes, y una muestra de validación formada por las vistas IZQUIERDA (100 vistas en total), que son las se pretenden clasificar. Los resultados obtenidos se muestran en la figura 7-28. Podemos apreciar que, con 13 parámetros (que eran los necesarios para clasificar bien el 100% de las vistas FRONTAL-ARRIBA-ABAJO-ZOOM, ver 7-21), se clasifican correctamente el 31% de las vistas IZQUIERDA.

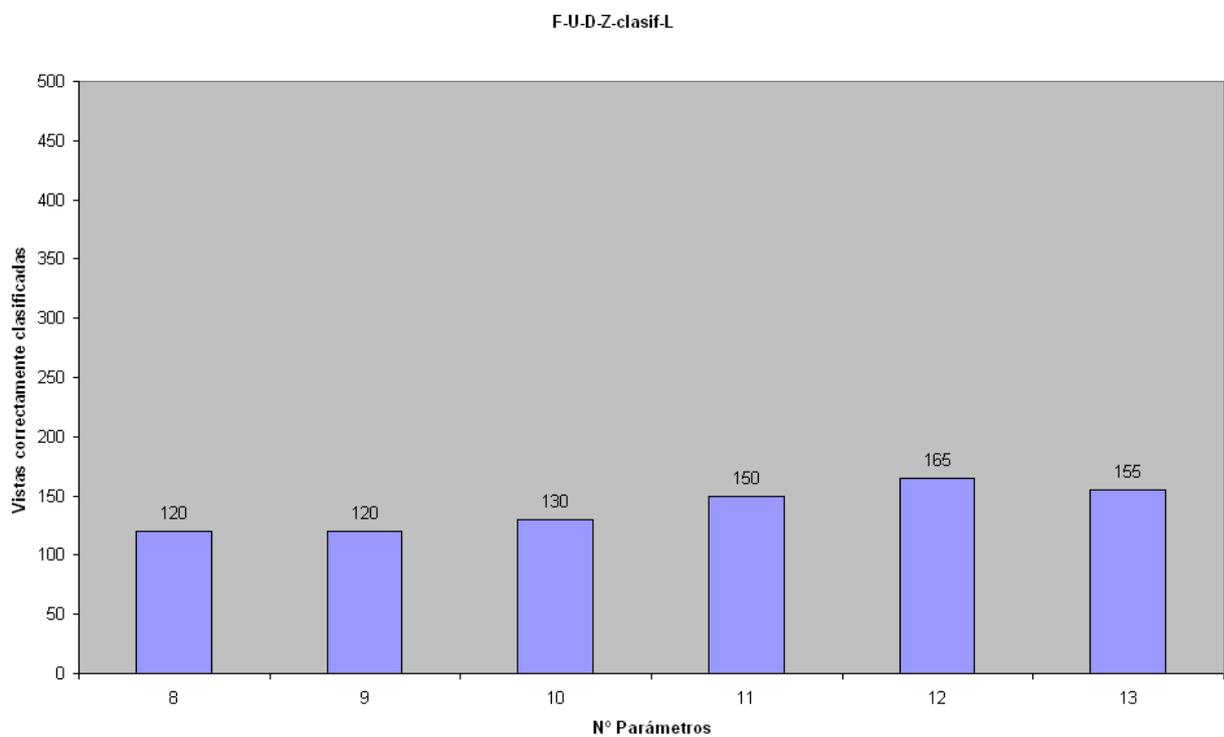


Figura 7-28: Exito de clasificación de las vistas IZQUIERDA a partir de las vistas FRONTAL-ARRIBA-ABAJO y ZOOM, con excentricidad variable, utilizando validación cruzada

El último paso de clasificación lo hacemos partiendo también de 500 vistas, divididas en dos grupos: la misma muestra de entrenamiento que en los casos precedentes, formada por las vistas ARRIBA, ABAJO, FRONTAL y ZOOM (400 en total), con las que se obtienen las funciones discriminantes, y una muestra de validación formada por las vistas DERECHA (100 vistas en total), que son las se pretenden clasificar. Los resultados obtenidos se muestran en la figura 7-29. Podemos apreciar que, con 13 parámetros (que eran los necesarios para clasificar bien el 100 % de las vistas FRONTAL-ARRIBA-ABAJO-ZOOM, ver 7-21), se clasifican correctamente el 43 % de las vistas DERECHA.

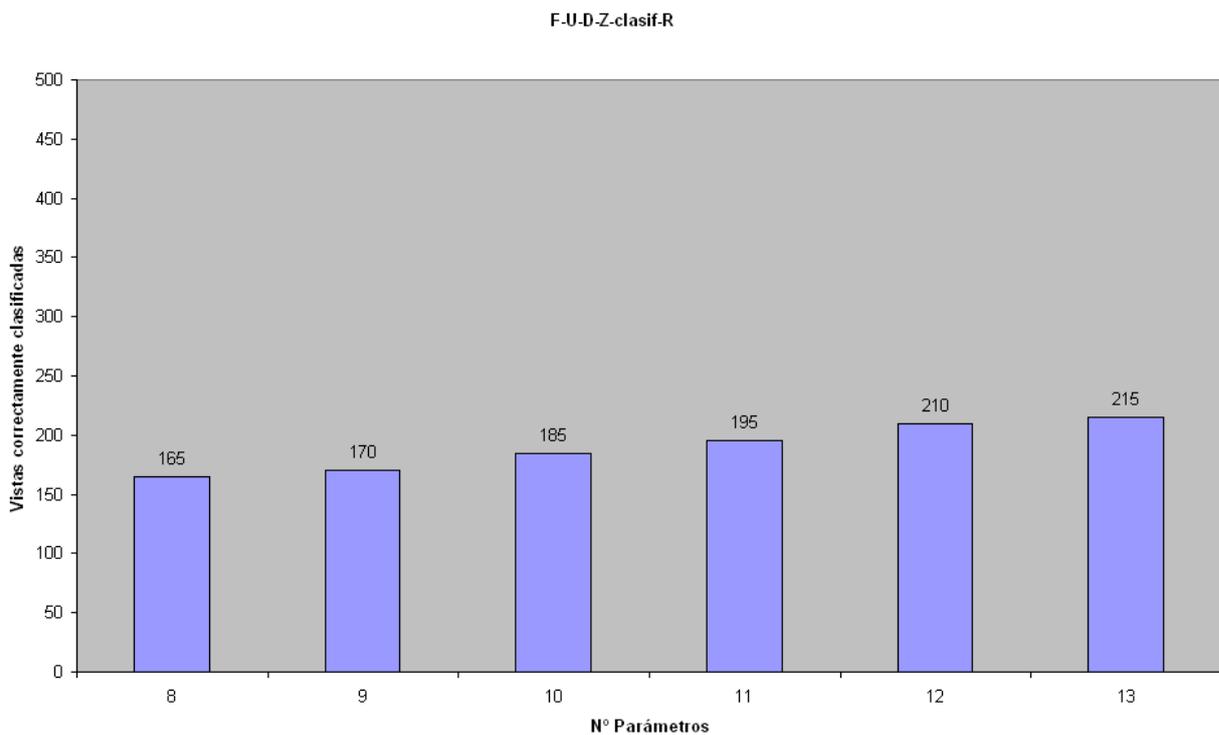


Figura 7-29: *Exito de clasificación de las vistas DERECHA a partir de las vistas FRONTAL-ARRIBA-ABAJO y ZOOM, con excentricidad variable, utilizando validación cruzada*

En la figura 7-30 podemos apreciar la evolución del porcentaje de clasificación

correcta de vistas, considerando siempre excentricidad variable, cuando el número de parámetros incluidos en el modelo varía entre 8 y 13. En el eje de Vistas Utilizadas, representamos por F-U-D y F-U-D-Z los casos de clasificación de 3 y 4 vistas, cuando no consideramos validación cruzada. El resto corresponde a cada uno de los casos de clasificación con validación cruzada que acabamos de explicar.

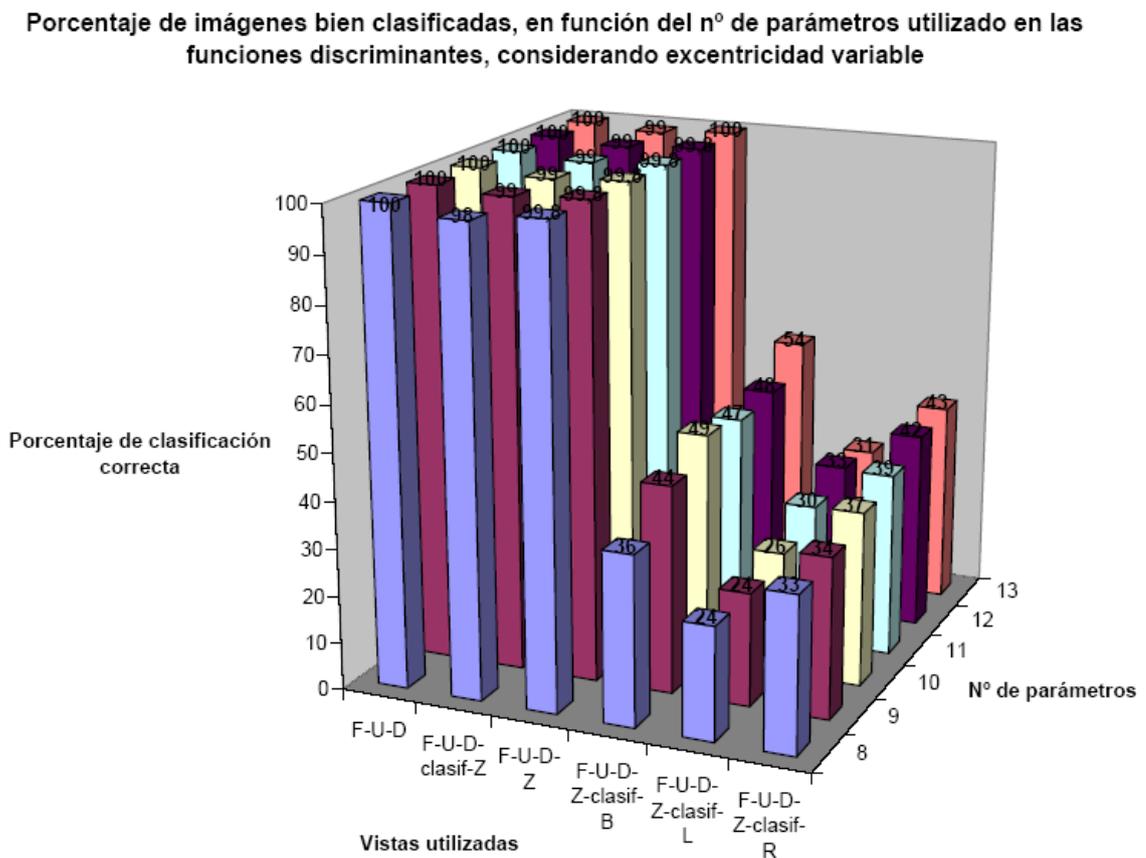


Figura 7-30: Porcentaje de vistas correctamente clasificadas, en función del número de parámetros considerados en las funciones discriminantes. F-D-U y F-U-D-Z corresponden a los casos donde no se considera validación cruzada y se clasifican las vistas FRONTAL-ARRIBA-ABAJO y FRONTAL-ARRIBA-ABAJO-ZOOM, respectivamente. El resto de los casos consideran validación cruzada. La excentricidad se considera variable en todos los casos.

7.2.6. Análisis de los datos normalizando la excentricidad de las orejas

En este apartado vamos a reproducir todas las fases del análisis anterior, cuando tomamos las medidas de las orejas, normalizadas para que la excentricidad de la elipse asociada sea constante.

En cuanto a la selección de variables, si pretendemos determinar qué parámetros son necesarios para clasificar correctamente las vistas *FRONTAL*, *ARRIBA* y *ABAJO*, obtenemos la gráfica de la figura 7-31. Los huecos en la gráfica están motivados por la misma causa que comentamos, en la fase de selección de parámetros, cuando considerábamos excentricidad variable. Podemos ver que se hacen necesarios 10 parámetros para clasificar las 300 vistas correctamente. Si pretendemos determinar los parámetros necesarios para clasificar correctamente las vistas *FRONTAL*, *ARRIBA*, *ABAJO* y *ZOOM*, obtenemos la gráfica de la figura 7-32. Son necesarios 11 parámetros para clasificar correctamente todas las vistas.

En la imagen 7-33 aparecen señaladas con puntos verdes las variables incluidas en el modelo, cuando trabajamos con excentricidad fija, considerando 3 vistas (*FRONTAL*, *ARRIBA*, *ABAJO*) y 4 vistas (*FRONTAL*, *ARRIBA*, *ABAJO* y *ZOOM*) respectivamente.

En lo que se refiere a la clasificación de las vistas, utilizando validación cruzada, los resultados obtenidos considerando excentricidad fija son los siguientes:

Empezamos partiendo de 400 vistas, divididas en dos grupos: una muestra de entrenamiento formada por las vistas *ARRIBA*, *ABAJO* y *FRONTAL* (300 en total), con las que se obtienen las funciones discriminantes, y una muestra de validación formada por las vistas *ZOOM* (100 vistas en total), que son las se pretenden clasificar. Los resultados obtenidos se muestran en la figura 7-34. Podemos apreciar que, con 10 parámetros (que eran los necesarios para clasificar bien el 100 % de las vistas *FRONTAL-ARRIBA-*

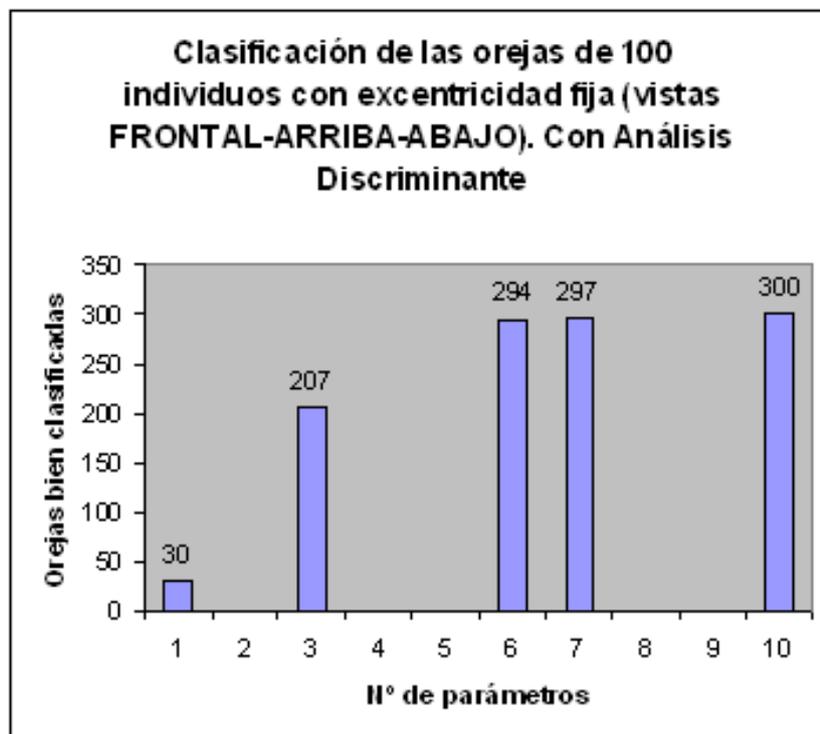


Figura 7-31: Resultado de repetir el proceso de selección de parámetros, forzando la inclusión de un parámetro adicional cada vez, tomando las vistas *FRONTAL*, *ARRIBA* y *ABAJO*, con excentricidad fija.

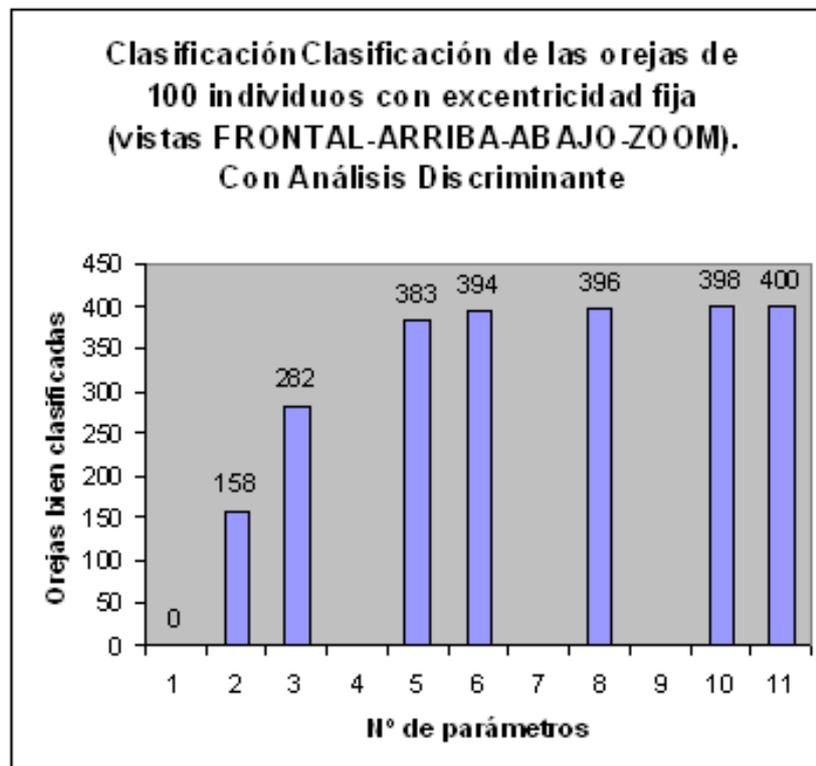


Figura 7-32: Resultado de repetir el proceso de selección de parámetros, forzando la inclusión de un parámetro adicional cada vez, tomando las vistas FRONTAL, ARRIBA, ABAJO y ZOOM, con excentricidad fija.

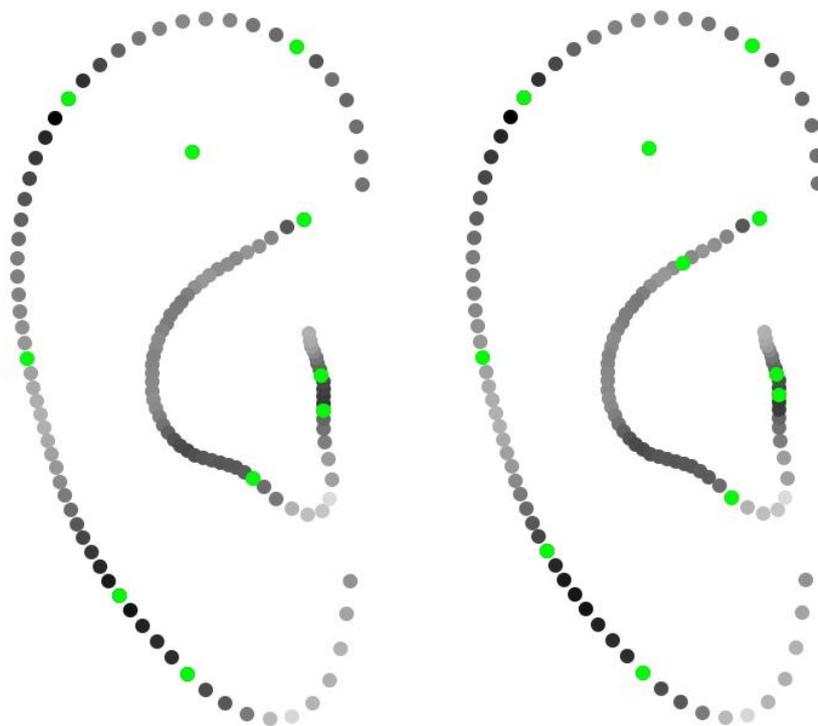


Figura 7-33: A la izquierda, en verde, parámetros seleccionados utilizando las vistas *FRONTAL-ARRIBA-ABAJO*. A la derecha lo mismo utilizando además la vista *ZOOM*. El círculo verde en el interior de la oreja significa que el parámetro de *distorsión del ovoide* también ha sido seleccionado

ABAJO, ver 7-31), se clasifican correctamente el 99 % de las vistas ZOOM. Si seguimos añadiendo parámetros al modelo, se mantiene el mismo porcentaje de clasificación, el 99 %, con 11 parámetros (los que eran necesarios para clasificar bien el 100 % de las vistas FRONTAL-ARRIBA-ABAJO-ZOOM cuando no se realizaba validación cruzada, ver 7-32) y se consigue el 100 % de clasificación correcta, si consideramos 13 parámetros.

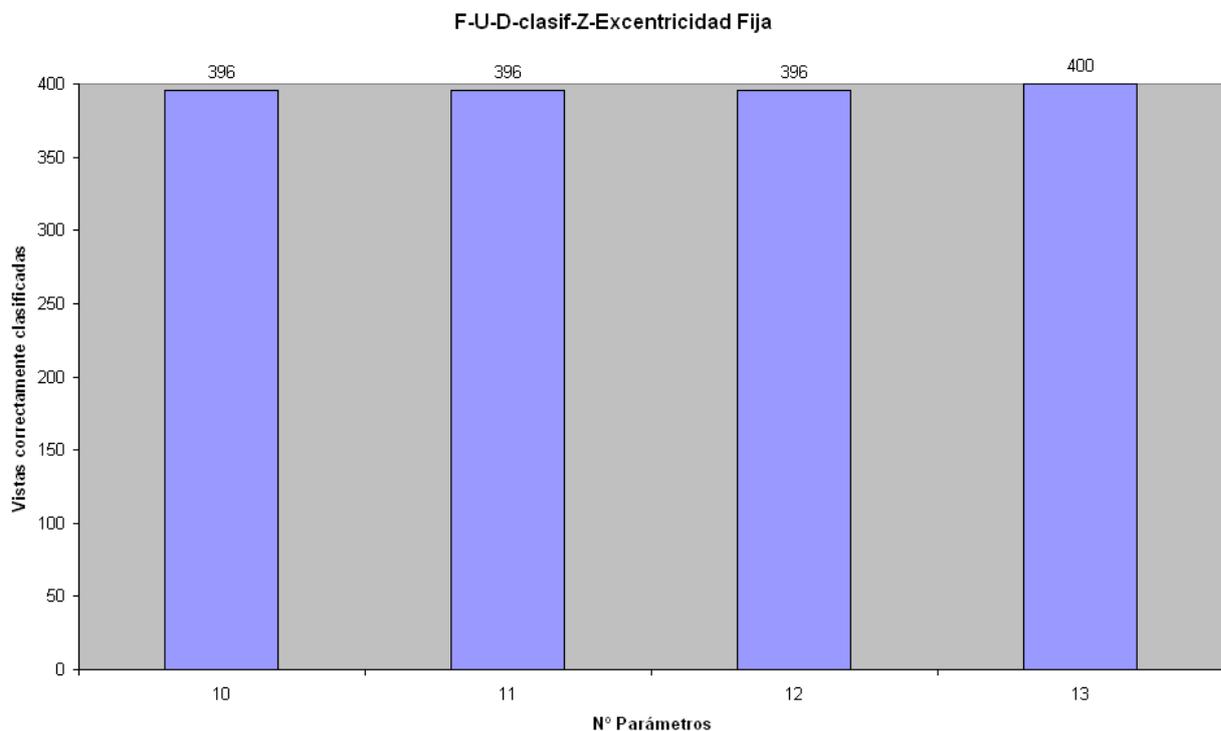


Figura 7-34: *Éxito de clasificación de las vistas ZOOM a partir de las vistas FRONTAL-ARRIBA y ABAJO, con excentricidad fija, utilizando validación cruzada*

A continuación pasamos a la situación más compleja con 500 vistas, divididas en dos grupos: una muestra de entrenamiento formada por las vistas ARRIBA, ABAJO, FRONTAL y ZOOM (400 en total), y una muestra de validación formada por las vistas POSTERIOR (100 vistas en total). Los resultados obtenidos se muestran en la figura 7-35. Podemos apreciar que, con 11 parámetros (los que eran necesarios para clasificar bien el 100 % de las vistas FRONTAL-ARRIBA-ABAJO-ZOOM, ver 7-32), se clasifican

correctamente el 47% de las vistas POSTERIOR.

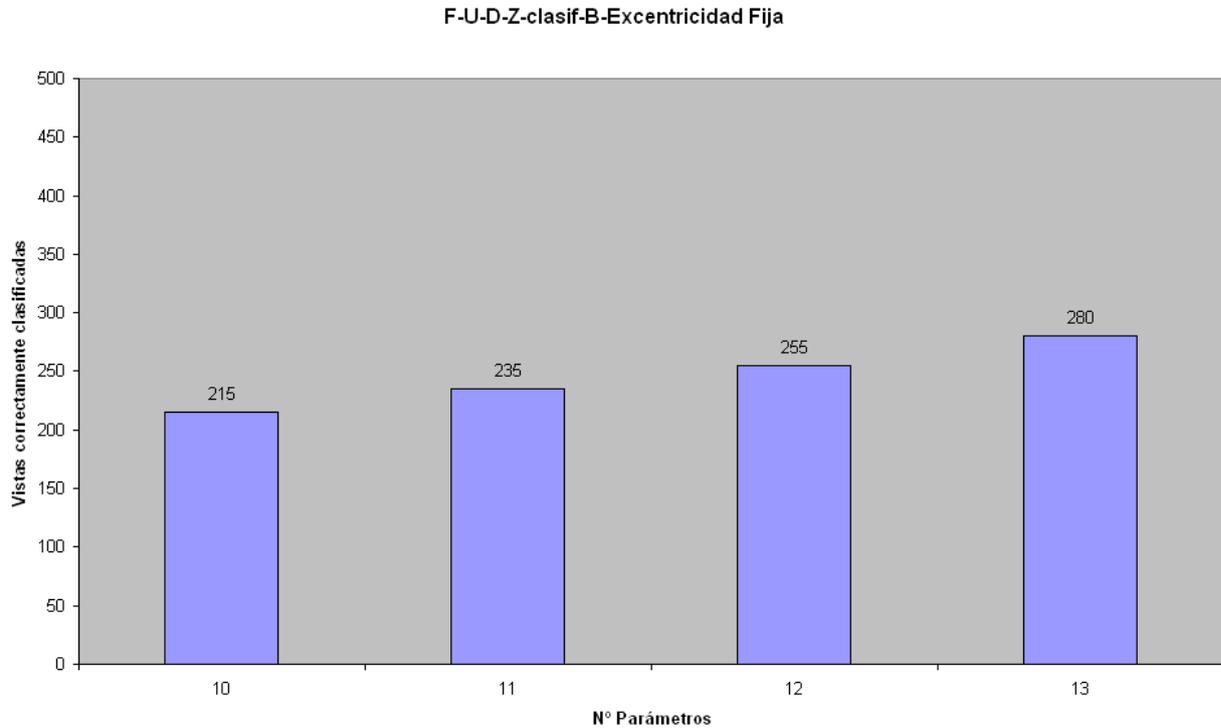


Figura 7-35: *Exito de clasificación de las vistas POSTERIOR a partir de las vistas FRONTAL-ARRIBA-ABAJO y ZOOM, con excentricidad fija, utilizando validación cruzada*

En el siguiente paso partimos también de 500 vistas, divididas en dos grupos: la misma muestra de entrenamiento que en el caso precedente, formada por las vistas ARRIBA, ABAJO, FRONTAL y ZOOM (400 en total), y una muestra de validación formada por las vistas IZQUIERDA (100 vistas en total). Los resultados obtenidos se muestran en la figura 7-36. Podemos apreciar que, con 11 parámetros (los que eran necesarios para clasificar bien el 100% de las vistas FRONTAL-ARRIBA-ABAJO-ZOOM, ver 7-32), se clasifican correctamente el 37% de las vistas IZQUIERDA.

El último paso de clasificación lo hacemos partiendo también de 500 vistas, divididas en dos grupos: la misma muestra de entrenamiento que en los casos precedentes,

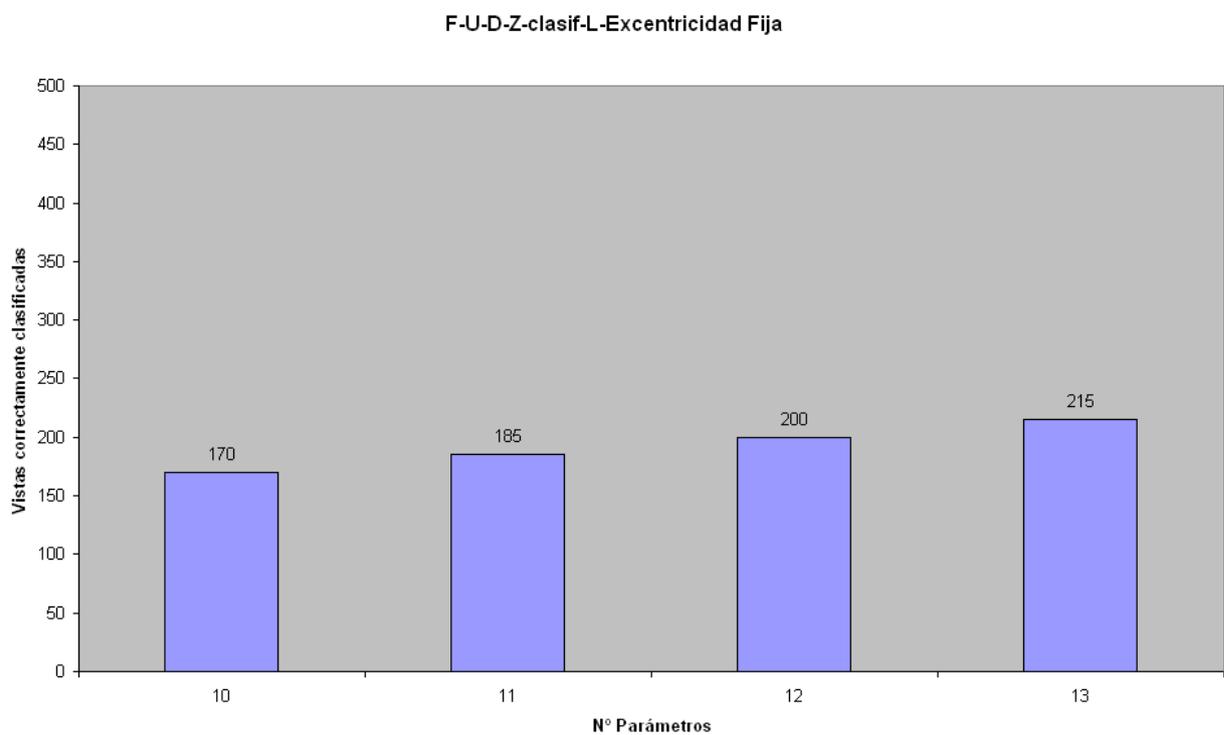


Figura 7-36: *Exito de clasificación de las vistas IZQUIERDA a partir de las vistas FRONTAL-ARRIBA-ABAJO y ZOOM, con excentricidad fija, utilizando validación cruzada*

formada por las vistas ARRIBA, ABAJO, FRONTAL y ZOOM (400 en total), y una muestra de validación formada por las vistas DERECHA (100 vistas en total). Los resultados obtenidos se muestran en la figura 7-37. Podemos apreciar que, con 11 parámetros (los que eran necesarios para clasificar bien el 100 % de las vistas FRONTAL-ARRIBA-ABAJO-ZOOM, ver 7-32), se clasifican correctamente el 39 % de las vistas DERECHA.

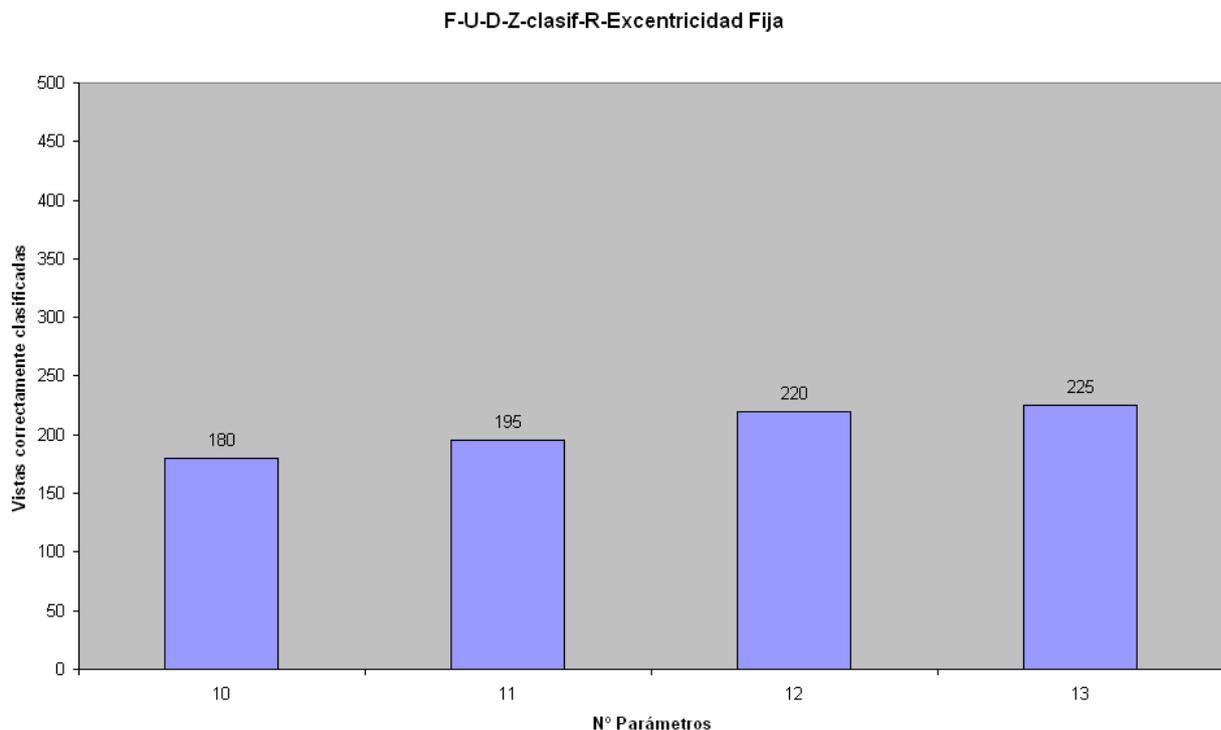


Figura 7-37: *Exito de clasificación de las vistas DERECHA a partir de las vistas FRONTAL-ARRIBA-ABAJO y ZOOM, con excentricidad fija, utilizando validación cruzada*

En la figura 7-38 podemos apreciar la evolución del porcentaje de clasificación correcta de vistas, considerando siempre excentricidad fija, cuando el número de parámetros incluidos en el modelo varía entre 10 y 13. En el eje de Vistas Utilizadas, representamos por F-U-D y F-U-D-Z los casos de clasificación de 3 y 4 vistas, cuando

no consideramos validación cruzada. El resto corresponde a cada uno de los casos de clasificación con validación cruzada que acabamos de explicar.

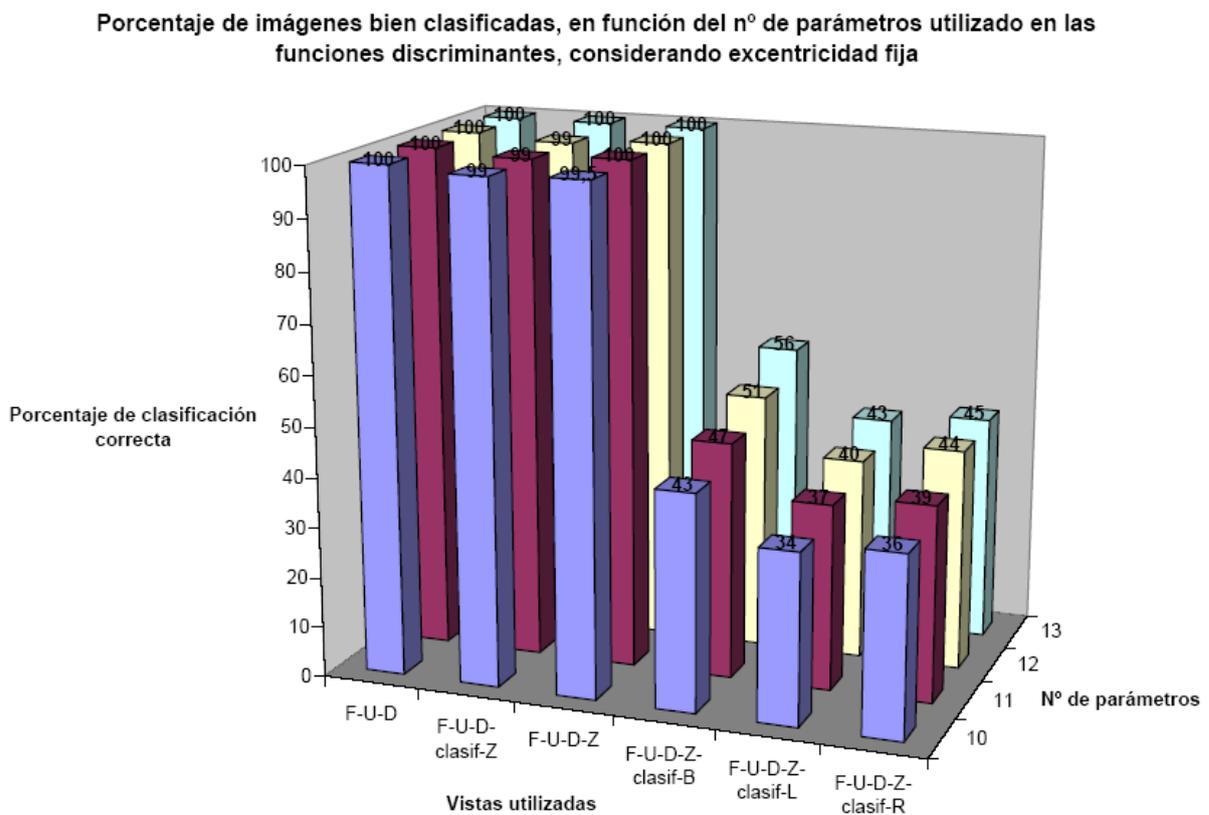


Figura 7-38: Porcentaje de vistas correctamente clasificadas, en función del número de parámetros considerados en las funciones discriminantes. F-D-U y F-U-D-Z corresponden a los casos donde no se considera validación cruzada y se clasifican las vistas *FRONTAL-ARRIBA-ABAJO* y *FRONTAL-ARRIBA-ABAJO-ZOOM*, respectivamente. El resto de los casos consideran validación cruzada. La excentricidad se considera fija en todos los casos.

7.3. Comparación de resultados obtenidos con ambos métodos

Como podremos comprobar, a pesar de que el objetivo perseguido es el mismo, y el estadístico involucrado en la selección de parámetros también es similar, los resultados obtenidos no coinciden exactamente en ambos casos (si bien son bastantes parecidos), lo que pudiera deberse a varios motivos. El principal, puede ser la diferencia de funcionamiento del proceso seguido en ambos casos. En el Análisis Estadístico inicial, las variables son seleccionadas comparando, en cada paso, la capacidad clasificadora del conjunto de variables introducidas en el modelo hasta ese paso. Sólo se continúa añadiendo variables nuevas si, con las que hay incluidas en un paso determinado, no se consigue todavía clasificar todas las imágenes. En el Análisis Discriminante, por el contrario, los dos procesos (selección y clasificación) se realizan por separado. Por una parte, se seleccionan las variables más discriminantes en general, atendiendo más a criterios estadísticos que de cara a la clasificación. Por otro lado, se realiza el proceso de clasificación, con el conjunto de variables previamente seleccionado, utilizando para ello las funciones discriminantes (combinaciones lineales de las variables seleccionadas).

Para la comparación de los resultados obtenidos por ambos métodos, mantendremos el mismo orden seguido en los estudios anteriores. Analizaremos primero los resultados, considerando la excentricidad variable, pasando luego a la situación donde consideramos la excentricidad de las orejas normalizada. Las tablas siguientes resumen la situación en ambos casos, para el Análisis Discriminante.

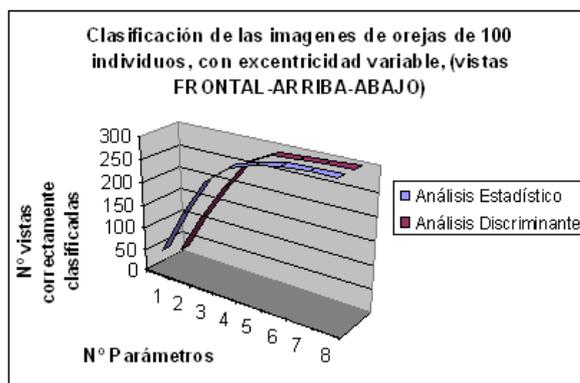
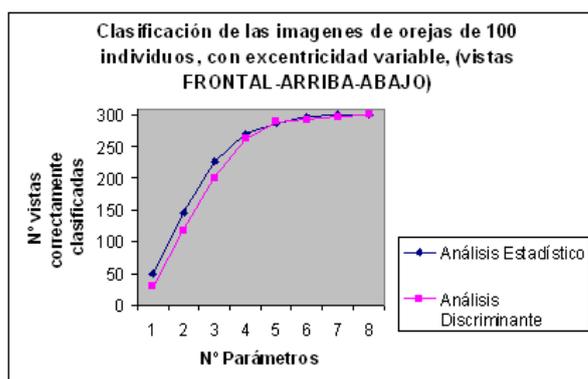
Análisis Discriminante.- Excentricidad Variable						
	8 parámetros	9 parámetros	10 parámetros	11 parámetros	12 parámetros	13 parámetros
F-U-D	100%	100%	100%	100%	100%	100%
F-U-D- clasif. Z	98%	99%	99%	99%	99%	99%
F-U-D-Z	99,80%	99,80%	99,80%	99,80%	99,80%	100%
F-U-D-Z- clasif B	36%	44%	49%	47%	48%	54%
F-U-D-Z- clasif L	24%	24%	26%	30%	33%	31%
F-U-D-Z- clasif R	33%	34%	37%	39%	42%	43%

Análisis Discriminante.- Excentricidad Fija				
	10 parámetros	11 parámetros	12 parámetros	13 parámetros
F-U-D	100%	100%	100%	100%
F-U-D- clasif. Z	99%	99%	99%	100%
F-U-D-Z	99,50%	100,00%	100,00%	100,00%
F-U-D-Z- clasif B	43%	47%	51%	56%
F-U-D-Z- clasif L	34%	37%	40%	43%
F-U-D-Z- clasif R	36%	39%	44%	45%

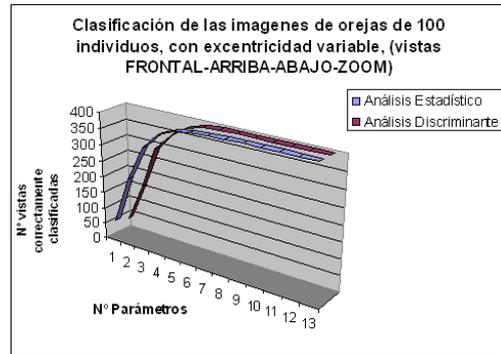
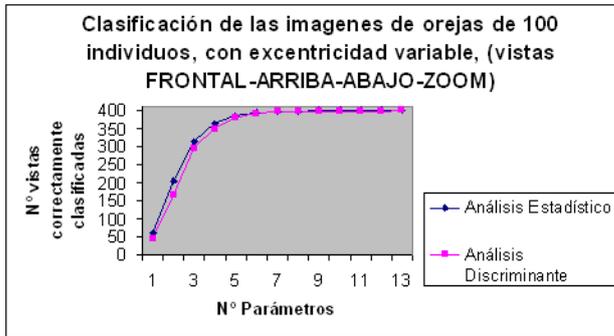
7.3.1. Excentricidad Variable:

Comenzaremos situándonos en el caso de intentar clasificar las mismas vistas que utilizamos para obtener los parámetros. Estudiaremos dos posibilidades, según consideremos 3 ó 4 vistas:

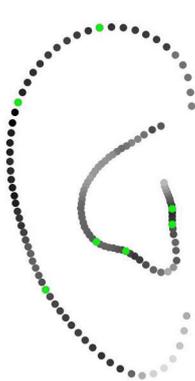
- Si las vistas utilizadas son FRONTAL-ARRIBA-ABAJO, e intentamos clasificar las mismas vistas. En el primer estudio, utilizando la distancia normalizada mejorada (caso más favorable), necesitábamos 7 parámetros para clasificar correctamente todas las vistas, mientras que, con el Análisis Discriminante se hacen necesarios 8 parámetros.



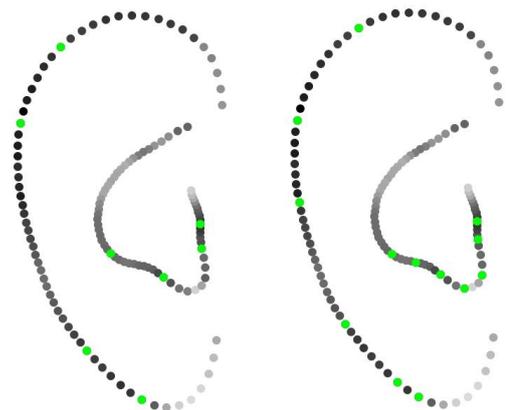
- Si las vistas utilizadas son FRONTAL-ARRIBA-ABAJO-ZOOM, en el primer estudio necesitamos 9 parámetros para la correcta clasificación del 100% de las vistas; mientras que, utilizando el Análisis Discriminante se necesitan 13 parámetros para lograr lo mismo.



Teniendo en cuenta lo anterior, podemos decir que, cuando se utilizan sólo vistas fronto-paralelas (3 ó 4 vistas), para tratar de clasificar esas mismas vistas, funciona mejor el Análisis Estadístico inicial que el Análisis Discriminante Clásico, ya que se requiere un menor número de parámetros para clasificarlas todas correctamente. Comparando las variables seleccionadas en el modelo del Análisis Discriminante con las obtenidas en el estudio inicial equivalente, podemos apreciar que, aunque las variables no coinciden exactamente en ambos casos, las consideradas más discriminantes por los dos métodos, se sitúan por las mismas zonas de la oreja, indicando que, efectivamente, hay zonas de la oreja que aportan más información que otras, de cara a identificar a los individuos.



Parámetros para clasificación de 3 y 4 vistas con Análisis Estadístico inicial



Parámetros para clasificación de 3 y 4 vistas con Análisis Discriminante

Pasamos ahora a comparar los resultados obtenidos cuando dividimos las vistas en dos grupos: un grupo de entrenamiento (con el que se determinan los parámetros a utilizar) y otro de validación (formado por las vistas que se pretenden clasificar). Estudiaremos cuatro posibilidades:

- Utilización de las vistas **FRONTAL-ARRIBA-ABAJO** como grupo de entrenamiento para clasificar las vistas **ZOOM**: En este caso, usando el Análisis Discriminante, con 8 parámetros (los necesarios para clasificar bien el 100 % de las 3 vistas de entrenamiento) se clasificaban correctamente el 98 % de las vistas **ZOOM**, y este porcentaje se aumentaba hasta el 99 %, con 13 parámetros (los necesarios para clasificar correctamente el 100 % de las 4 vistas implicadas, cuando no se realizaba validación cruzada). En esta misma situación, usando el Análisis Estadístico inicial, se obtenía una clasificación correcta del 98 % de las vistas (correspondería al éxito de clasificación acumulado para el parámetro $M=0$ en la figura 7-8). Los resultados obtenidos con ambos métodos, son equivalentes.
- Utilización de las vistas **FRONTAL-ARRIBA-ABAJO-ZOOM** como grupo de entrenamiento para clasificar las vistas **POSTERIOR**: En este caso, usando el Análisis Discriminante, con 13 parámetros (los necesarios para clasificar bien el 100 % de las 4 vistas de entrenamiento) se clasificaban correctamente el 54 % de las vistas **POSTERIOR**. En esta misma situación, usando el Análisis Estadístico inicial, se obtenía una clasificación correcta del 33 % de las vistas (correspondería al éxito de clasificación acumulado para el parámetro $M=0$ en la figura 7-9). Los resultados obtenidos con el Análisis Discriminante son mejores.
- Utilización de las vistas **FRONTAL-ARRIBA-ABAJO-ZOOM** como grupo de entrenamiento para clasificar las vistas **IZQUIERDA**: En este caso, usando el Análisis Discriminante, con 13 parámetros se clasificaban correctamente el 31 % de las vistas **IZQUIERDA**. En esta misma situación, usando el Análisis Estadístico inicial, se obtenía una clasificación correcta del 15 % de las vistas (correspondería al

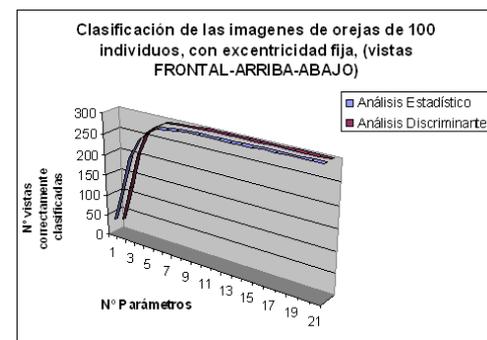
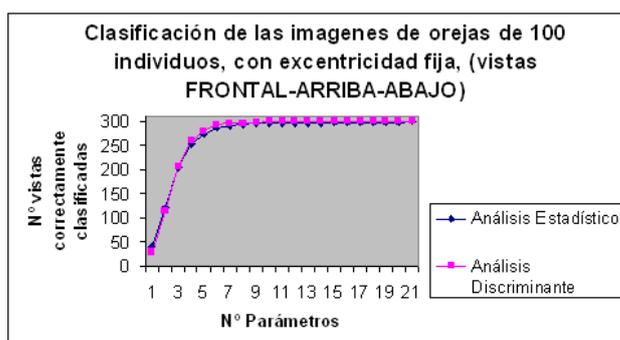
éxito de clasificación acumulado para el parámetro $M=0$ en la figura 7-11). Los resultados obtenidos con el Análisis Discriminante son mejores.

- Utilización de las vistas FRONTAL-ARRIBA-ABAJO-ZOOM como grupo de entrenamiento para clasificar las vistas DERECHA: En este caso, usando el Análisis Discriminante, con 13 parámetros se clasificaban correctamente el 43 % de las vistas DERECHA. En esta misma situación, usando el Análisis Estadístico inicial, se obtenía una clasificación correcta del 30 % de las vistas (correspondería al éxito de clasificación acumulado para el parámetro $M=0$ en la figura 7-10). Los resultados obtenidos con el Análisis Discriminante son mejores.

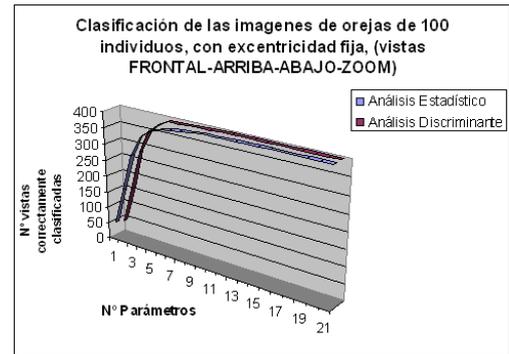
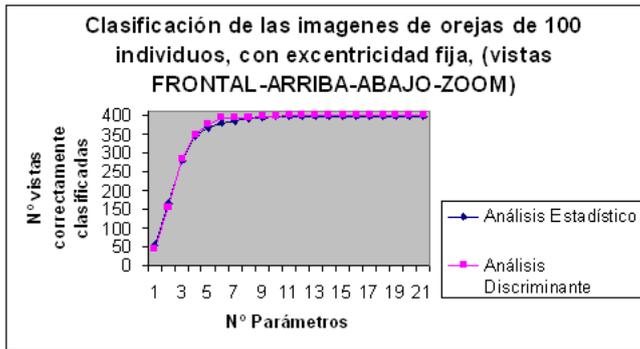
7.3.2. Excentricidad Normalizada:

Comenzaremos situándonos en el caso de intentar clasificar las mismas vistas que utilizamos para obtener los parámetros. Estudiaremos dos posibilidades, según consideremos 3 ó 4 vistas:

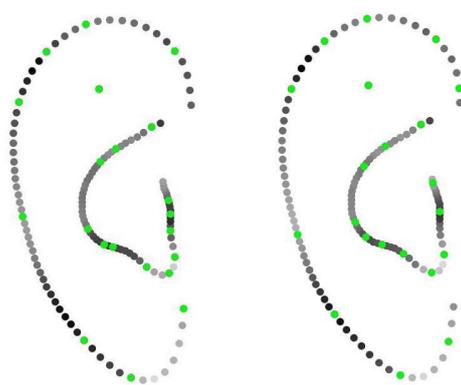
- Si las vistas utilizadas son FRONTAL-ARRIBA-ABAJO, e intentamos clasificar las mismas vistas. En el primer estudio, utilizando la distancia normalizada mejorada (caso más favorable), necesitábamos 21 parámetros para clasificar correctamente todas las vistas, mientras que, con el Análisis Discriminante se hacen necesarios sólo 10 parámetros.



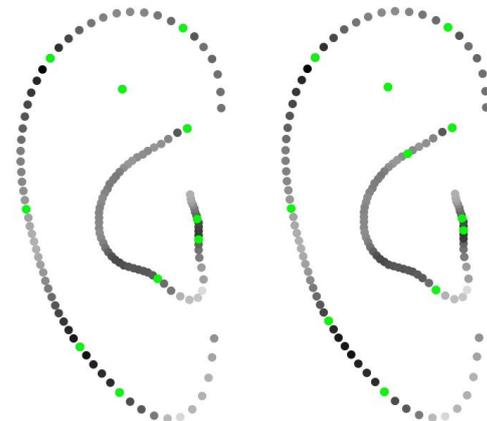
- Si las vistas utilizadas son FRONTAL-ARRIBA-ABAJO-ZOOM, en el primer estudio necesitamos también 21 parámetros para la correcta clasificación del 100 % de las vistas; mientras que, utilizando el Análisis Discriminante se necesitan 11 parámetros para lograr lo mismo.



Podemos afirmar por tanto, que, cuando se utilizan sólo vistas fronto-paralelas (3 ó 4 vistas), para tratar de clasificar esas mismas vistas y, contrariamente a lo que ocurría en el apartado anterior, funciona mejor el Análisis Discriminante Clásico que el Análisis Estadístico inicial, ya que se requiere un número considerablemente menor de parámetros para clasificarlas todas correctamente. También, en este caso, las variables consideradas más discriminantes se sitúan cercanas en las mismas zonas de la oreja.



Parámetros para clasificación de 3 y 4 vistas con Análisis Estadístico inicial



Parámetros para clasificación de 3 y 4 vistas con Análisis Discriminante

Pasamos ahora a comparar los resultados obtenidos cuando dividimos las vistas en dos grupos: un grupo de entrenamiento (con el que se determinan los parámetros a utilizar) y otro de validación (formado por las vistas que se pretenden clasificar). Estudiaremos las mismas cuatro posibilidades que antes:

- Utilización de las vistas **FRONTAL-ARRIBA-ABAJO** como grupo de entrenamiento para clasificar las vistas **ZOOM**: En este caso, usando el Análisis Discriminante, con 10 parámetros (los necesarios para clasificar bien el 100% de las 3 vistas de entrenamiento) se clasificaban correctamente el 99% de las vistas **ZOOM**, y este porcentaje se mantenía en el 99%, con 11 parámetros (los necesarios para clasificar correctamente el 100% de las 4 vistas implicadas, cuando no se realizaba validación cruzada) y se alcanzaba el 100% de clasificación correcta con 13 parámetros. En esta misma situación, usando el Análisis Estadístico inicial, se obtenía una clasificación correcta del 97% de las vistas (correspondería al éxito de clasificación acumulado para el parámetro $M=0$ en la figura 7-16). Los resultados obtenidos con ambos métodos, son equivalentes.
- Utilización de las vistas **FRONTAL-ARRIBA-ABAJO-ZOOM** como grupo de entrenamiento para clasificar las vistas **POSTERIOR**: En este caso, usando el Análisis Discriminante, con 11 parámetros (los necesarios para clasificar bien el 100% de las 4 vistas de entrenamiento) se clasificaban correctamente el 47% de las vistas **POSTERIOR**. En esta misma situación, usando el Análisis Estadístico inicial, se obtenía una clasificación correcta del 42% de las vistas (correspondería al éxito de clasificación acumulado para el parámetro $M=0$ en la figura 7-17). Los resultados obtenidos con el Análisis Discriminante son ligeramente mejores, pero bastante aproximados.
- Utilización de las vistas **FRONTAL-ARRIBA-ABAJO-ZOOM** como grupo de entrenamiento para clasificar las vistas **IZQUIERDA**: En este caso, usando el Análisis Discriminante, con 11 parámetros se clasificaban correctamente el 37% de las

vistas IZQUIERDA. En esta misma situación, usando el Análisis Estadístico inicial, se obtenía una clasificación correcta del 24 % de las vistas (correspondería al éxito de clasificación acumulado para el parámetro $M=0$ en la figura 7-19). con el Análisis Discriminante son ligeramente mejores, pero bastante aproximados.

- Utilización de las vistas FRONTAL-ARRIBA-ABAJO-ZOOM como grupo de entrenamiento para clasificar las vistas DERECHA: En este caso, usando el Análisis Discriminante, con 13 parámetros se clasificaban correctamente el 39 % de las vistas DERECHA. En esta misma situación, usando el Análisis Estadístico inicial, se obtenía una clasificación correcta del 38 % de las vistas (correspondería al éxito de clasificación acumulado para el parámetro $M=0$ en la figura 7-18). Los resultados obtenidos con ambos métodos, son equivalentes.

Conclusiones

Consideramos que la oreja humana es una fuente adecuada de información para la identificación pasiva de un individuo en muchas aplicaciones diferentes. Dada la creciente necesidad actual de establecer medidas de seguridad y control de acceso en varios lugares públicos (edificios oficiales, aeropuertos, estaciones, etc...), la biometría de la oreja parece ser una buena solución, ya que: (1) las orejas son una parte del cuerpo que está generalmente visible y (2) se puede tomar imágenes de ella fácilmente, incluso sin el conocimiento ni consentimiento de la persona examinada.

La biometría de la oreja también se puede utilizar para acentuar la efectividad de otras biometrías ampliamente utilizadas (como voz, geometría de la mano o identificación de rostros), en la implementación de sistemas biométricos multimodales (híbridos), que están recibiendo mucha atención últimamente.

En cualquier caso, se hace necesario un método robusto de extracción de características, a partir de las imágenes tomadas de la oreja, que se pueda usar para determinar la identidad de algunos individuos.

En este trabajo proponemos un método de identificación de individuos, basado en la extracción de ciertas características geométricas de las imágenes de sus orejas. El método propuesto no es totalmente automático y requiere de la intervención inicial del usuario, para facilitar una primera aproximación al contorno de la oreja, aspecto este

susceptible de mejora en el futuro.

Partiendo de esta aproximación, mejoramos la localización del contorno externo de la oreja en la imagen, usando para ello una adaptación del modelo de contornos activos (snakes) que se ajusta a los requerimientos del contorno de la oreja (contorno abierto), ya que el modelo original de los snakes está definido para trabajar con curvas cerradas. La aproximación final al contorno de la oreja, así obtenido, resulta ser muy buena, como demuestran las imágenes mostradas en el capítulo dedicado a la explicación de esta parte del método.

Se usa entonces un modelo de ovoide para ajustar este contorno final estimado. En una primera aproximación, se obtienen los parámetros de un ovoide que se adapta al contorno de la oreja, que luego se irán refinando, basándonos en un criterio de distancia euclídea. La ventaja de este modelo de ovoide, aparte de ser bastante simple y ajustarse bastante bien al contorno externo de la oreja, es que, sus parámetros (algunos de los cuales son invariantes euclídeos), nos permiten alinear los contornos extraídos de varias imágenes de orejas para compararlos entre sí, independientemente de su tamaño y localización.

Como el objetivo final es poder comparar diferentes imágenes de orejas, se hace necesario un proceso de normalización de coordenadas. Para ello nos basamos en los parámetros del ovoide que hemos asociado al contorno y aplicamos una transformación para llevar el centro de la elipse deformada al origen de coordenadas y colocarla en vertical. El hecho de localizar el centro del nuevo sistema de coordenadas en el origen, hace que nuestro método sea invariante a rotaciones, traslaciones y escalados. Una vez hecho esto, existen dos opciones: comprimir/expandir las elipses deformadas obtenidas hasta alcanzar una excentricidad fija dada, o bien, dejar la excentricidad sin fijar, con valores variables. En el estudio hemos considerado las dos opciones.

Luego, para cada imagen, calculamos el centroide y obtenemos las distancias de este centroide a los puntos del contorno externo e interno de la oreja. Estas distancias,

Conclusiones

junto con los parámetros del mejor ovoide que ajusta cada contorno, dan lugar a un vector de características que identifica la imagen. La clasificación de imágenes como pertenecientes a un individuo se realiza, entonces, basándonos en la simple comparación entre el vector de características de la imagen a clasificar y los vectores de características presentes en la base de datos.

Para la implementación del método hemos generado una base de datos con imágenes correspondientes a 100 individuos, 7 imágenes por individuo, tomadas todas en las mismas condiciones de iluminación. Para todas ellas hemos calculado el conjunto de distancias antes mencionado y, junto con los parámetros del mejor ovoide que ajusta cada contorno, se han sometido a un estudio estadístico (análisis discriminante) y matemático para extraer, de entre todas estas medidas, el conjunto de características que mejor discrimina las imágenes. Los resultados obtenidos nos permiten señalar (para la población analizada) qué zonas de las orejas son más importantes, de cara a identificar al individuo, y otras zonas que presentan menos diferencias entre ellos.

Los resultados obtenidos hasta ahora son prometedores, tal y como se explica detalladamente en el apartado de Resultados Experimentales.

De las 130 variables iniciales de estudio, se concluye que la capacidad discriminante de un pequeño subconjunto de ellas (entre 7 y 21) resulta suficiente para identificar al individuo al que corresponde la imagen de la oreja. Además estas variables señalan unas zonas concretas de la oreja que son las de mayor poder diferenciador.

El método propuesto es bastante bueno para clasificar las vistas que corresponden a una situación donde se toma una fotografía fronto-paralela del individuo, incluso cuando se varía la distancia del individuo al objetivo de la cámara (ZOOM). En esta situación, tanto con el estudio propuesto por nosotros, como por el análisis discriminante clásico, se obtiene un porcentaje de clasificación correcta cercano al 100%. Las vistas POSTERIOR, IZQUIERDA y DERECHA resultan más complejas de analizar porque, o bien corresponden a la otra oreja del individuo, o bien incluyen un fenómeno

de distorsión de la forma de la oreja debido a la perspectiva. En estos casos, se obtienen mejores resultados cuando tomamos las medidas de las orejas, normalizadas para que la excentricidad de la elipse asociada sea constante, ya que la normalización de la excentricidad corrige parcialmente la deformación de perspectiva. Para estas vistas con perspectiva se obtienen mejores resultados con el análisis discriminante clásico que con nuestro estudio estadístico, estando siempre la clasificación correcta por debajo del 60 %.

Futuras Líneas de Desarrollo

Durante el desarrollo de esta tesis se nos han planteado otros trabajos que se podrían desarrollar para mejorar o completar los métodos propuestos pero que, por cuestiones de tiempo y esfuerzo, quedaba claro que estaban fuera del ámbito de esta tesis como tal, pero que sería interesante tenerlas en cuenta en futuros trabajos. Entre estas futuras líneas de desarrollo podemos destacar:

9.0.3. Automatización de la aproximación inicial:

Todo el desarrollo del método implementado parte de una aproximación inicial a los contornos externo e interno de la oreja que, actualmente, se introduce de forma manual por el usuario. Esto conlleva, como hemos podido constatar con la realización de la base de datos, mucho esfuerzo. Una posible mejora del método consistiría en intentar eliminar esta intervención del usuario y automatizar el proceso de obtención de ese polígono inicial de aproximación a los contornos. Para obtener algunos puntos cercanos a los contornos de la oreja podrían intentar utilizarse, como base, algunos de los métodos conocidos de tratamiento y segmentación de imágenes. De hecho, algunos investigadores han intentado extraer los contornos de la oreja (el contorno externo básicamente) utilizando el detector de bordes de Canny, al parecer, con buenos resultados [31].

9.0.4. Optimización de la localización del centroide:

El centroide, calculado a partir de los puntos del contorno interno, es un punto de referencia clave en el desarrollo del algoritmo. Partiendo de él se obtienen todas las medidas de las distancias a los dos contornos de la oreja, que luego servirán para clasificar las imágenes y para determinar las zonas de la oreja que más discriminan entre los individuos. Una pequeña variación en cualquiera de los puntos del contorno interno puede influir en la localización de este centroide y, por tanto, afectar al resto de medidas. Futuros trabajos deberían ir encaminados a optimizar la localización de este punto de referencia.

9.0.5. Implementación de Métodos de Reconocimiento y no sólo de Identificación:

La *Identificación* se usa para verificar que las medidas biométricas extraídas de un individuo encajan suficientemente con las medidas biométricas previamente adquiridas para ese mismo sujeto.

Vamos a representar por s al sujeto en el momento del registro o inscripción de usuarios, y por s' al mismo sujeto en el momento de la *Identificación*.

Consideremos $Gs = f(s)$ como una función que extrae algunos biométricos del sujeto s , y $d(Gs, Gs')$ como la medida, previamente calculada, de la distancia entre los valores de esta función evaluada, para el mismo sujeto, en las dos situaciones anteriormente descritas.

En este contexto, la *Identificación* consistiría en determinar si $d(Gs, Gs') \leq T$, siendo T el valor de un umbral de aceptación predeterminado.

Esta parte es la que hemos implementado en nuestro trabajo, de manera que la identificación o clasificación de imágenes como pertenecientes a un individuo deter-

minado, se realiza por la simple comparación entre el vector de características de la imagen de entrada y todos los vectores de características incluidos en la base de datos.

En el mismo contexto anterior, el *Reconocimiento* es un proceso más complicado que la *Identificación*, ya que el sistema tiene que determinar si la identidad de un sujeto se puede verificar entre la de los sujetos previamente registrados.

Si representamos por $I = \{G_0, G_1, \dots, G_n\}$ al conjunto de identidades previamente registradas en el sistema, entonces el problema de *Reconocimiento* de un sujeto s se reduce a encontrar el elemento G_i del conjunto $\{G_i / G_i \in I \ \& \ d(Gs', G_i) \leq T\}$ tal que $d(Gs', G_i)$ sea la menor de todas.

Debe permitirse una cierta tolerancia en el criterio de encaje, ya que las medidas biométricas, del mismo sujeto, pueden sufrir una pequeña variación a lo largo del tiempo. Esta tolerancia se define en términos de la FRR (Tasa de Falso Rechazo) y la FAR (Tasa de Falsa Aceptación) mostrada por el sistema:

Falso Rechazo: Un sujeto registrado es rechazado, porque el sistema no encuentra sus medidas biométricas actuales lo suficientemente similares a ninguna de las medidas biométricas almacenadas en la base de datos.

Falsa Aceptación: Un sujeto "impostor" se acepta como legítimo, porque el sistema encuentra las medidas biométricas del sujeto "impostor" lo suficientemente parecidas a las medidas biométricas de alguno de los sujetos registrados en la base de datos.

Idealmente un sistema de reconocimiento debería tener valores muy bajos de ambas tasas (FRR y FAR).

La implementación, partiendo de nuestro algoritmo, de este proceso de *Reconocimiento* y la medida de las tasas de Falsa Aceptación y Falso Rechazo del sistema resultante, quedan como una propuesta de trabajo futuro.

Bibliografía

- [1] Alvarez, L., Gonzalez, E., Mazorra, L.: *Fitting ear contour using an ovoid model*. Proceedings of 39th IEEE International Carnahan Conference on Security Technology: 145-148 (2005).
- [2] Jain, A. K., Ross, A. and Prabhakar, S.: *An Introduction to Biometric Recognition*. IEEE Transactions on Circuits and Systems for Video Technology: Special Issue on Image and Video-Based Biometrics, 14 (1): 4-20 (2004).
- [3] Ratha, N. K., Senior, A., Bolle, R. M.: *Automated Biometrics*. Proceedings of International Conference on Advances in Pattern Recognition, (2001)
- [4] Moreno, B. Sánchez, A., Vélez, J.F.: *On the Use of Outer Ear Images for Personal Identification in Security Applications*. IEEE 33rd Annual International Carnahan Conference on Security Technology: 469-476 (1999)
- [5] Choras, M.: *Ear Biometrics Based on Geometrical Feature Extraction*. Electronic Letters on Computer Vision and Image Analysis, 5: 84-95 (2005).
- [6] Víctor, B. Bowyer, K. and Sarkar, S.: *An Evaluation of Face and Ear Biometrics*. Proceedings of 16th International Conference on Pattern Recognition, 429-432 (2002).

- [7] Hoogstrate, A. J., Van den Heuvel, H., Huyben, E.: *Ear Identification Based on Surveillance Camera's Images*. From http://www.forensic_evidence.com/site/ID/IDearCamera.html. Version updated May 31 (2000). Retrieved Jun 10 (2004).
- [8] Bromba, GmbH: *Bioidentification Frequently Asked Questions*. From <http://www.bromba.com/faq/biofaq.thm>. Version updated Sept 12 (2003). Retrieved Oct 28 (2003).
- [9] Chang, K., Bowyer, K. and Barnabas, V.: *Comparison and Combination of Ear and Face Images in Appearance-Based Biometrics*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 25 (9):1160-1165 (2003).
- [10] Hurley, D., Nixon, M. and Carter, J.: *Force Field Energy Functionals for Image Feature Extraction*. Image and Vision Computing Journal, 20: 429-432 (2002).
- [11] Hurley, D., Nixon, M. and Carter, J.: *Automatic Ear Recognition by Force Field Transformations*. IEEE Colloquium on Visual Biometrics (2000).
- [12] Burge, M. and Burger, W.: *Ear Biometrics*. BIOMETRICS: Personal Identification in a Networked Society. A.Jain, R. Bolle and S. Pankanti editors, Kluwer Academic: 273-286 (1998).
- [13] Burge, M. and Burger, W.: *Ear Biometrics in Computer Vision*. Proceedings of 15th International Conference on Pattern Recognition, 2 : 822-826 (2000).
- [14] Burge, M. and Burger, W.: *Ear Biometrics for Machine Vision*. 21st Workshop of the Austrian Association for Pattern Recognition (1997).
- [15] Bamber, D.: *Prisoners to appeal as unique earprint evidence is discredited*. Telegraph Newspaper (U.K.). From <http://portal.telegraph.co.uk/news/main.jtml?xml=/news/2001/12/02/nearp02.xml>. Version updated Feb 12 (2001). Retrieved Oct 3 (2003).

- [16] Carreira Perpiñan, M.A.: *Compression Neural Networks for Feature Extraction. Application to Human Recognition from Ear Images*. MSC thesis, Faculty of Informatics, Technical University of Madrid, Spain (1995).
- [17] Curiel López de Arcante, A. M. and Granell Navarro, J.: *La huella de la oreja como método de identificación*. Acta Otorrinolaringol. Esp, 57: 329-332 (2006).
- [18] Iannarelli, A. : *Ear Identification*. Paramont Publishing (1989).
- [19] Morgan, J.: *Court Holds Earprint Identification Not Generally Accepted in Scientific Community*. From http://www.forensic_evidence.com/site/ID/ID-Kunze.html. Version updated (1999). Retrieved Sept 9 (2003).
- [20] Yuizono, T., Wang, Y., Satoh, K. and Nakayanna, S. : *Study on Individual Recognition for Ear Image by Using Genetic Local Search*. Proceedings of Congress Evolutionary Computation: 237-242 (2002).
- [21] Bhanu, B. and Chen, H. : *Human Ear Recognition in 3D*. Proceedings of Workshop Multimodal User Authentication: 91-98 (2003).
- [22] Pun, K. and Moon, Y. : *Recent Advances in Ear Biometrics*. Proceedings of 6th International Conference on Automatic Face and Gesture Recognition: 164-169 (2004).
- [23] Kass, M., Witkin, A. and Terzopoulos, D. : *Snakes: Active Contours Models*. Proceedings of the 1st International Conference on Computer Vision: 259-268 (1987).
- [24] Caselles, V., Catta, F. and Dibos, F.: *A Geometric Model for Active Contours in Image Processing*. Numer. Math, 66 (1993).
- [25] Malladi, R. and Sethian, J.A.: *An $O(N \log M)$ Algorithm for Shape Modelling*. Proceedings of the National Academy of Sciences, 23 : 9389-9392 (1996).
- [26] Osher, S., Sethian, J. A.: *Fronts Propagating with Curvature Dependent Speed: Algorithms based on Hamilton-Jacobi Formulation*. J. Comps. Phys, 79: 12-49 (1998).

- [27] Kass, M., Witkin, A. and Terzopoulos, D. : *Snakes: Active Contours Models*. International Journal on Computer Vision,1: 321-331 (1988).
- [28] Cohen, L. D., Cohen, I. : *Finite Element Methods for Active Contour Models and Ballons for 2D and 3D Images*. IEEE Trans. Pattern Analysis and Machine Intelligence, 15: 1131-1147 (1993).
- [29] Blake, A. and Yuille, A. editores: *Active Vision*. MIT Press (1992).
- [30] Caselles, V., Kimmel, R. and Sapiro, G. : *Geodesic Active Contours*. International Journal on Computer Vision, 22 (1): 61-79 (1997).
- [31] Ansari, S. and Gupta, P.: *Localization of Ear using Outer Helix Curve of the Ear*. (2007).
- [32] Sana, A., Gupta, P. and Purkait, R.: *Ear Biometrics: A New Approach*. Proceedings of 6th International Conference on Advances in Pattern Recognition (2007).
- [33] Bry, X.: *Analyses factorielles multiples*. Ed. ECONOMICA (1996).
- [34] Pérez, C.: *Técnicas de Análisis Multivariante de Datos: Aplicaciones con SPSS*. Pearson Educación S.A. (2004).
- [35] Baker, D. E.: *A geometric method for determining shape of bird eggs*. Auk 119 (4): 1179-1186, (2002)
- [36] Monus, F. and Barta, Z. : *Repeatability analysis of egg shape in a wild tree sparrow (passer montanus) population: A sensitive method for egg shape description*. Acta Zoologica Academiae Scientiarum Hungaricae 51 (2): 151-162 (2005)
- [37] Preston, F. W.: *The shapes of bird's eggs*. Auk 70:160-182 (1953)
- [38] Todd, P.H. and Smart, I.H.M.: *The shape of bird's eggs*. Journal Theor. Biol. 106: 239-243, (1984)

-
- [39] *Cap. 23: Análisis Discriminante*. From <http://www2.uca.es/serv/ai/formacion/spss/Pantalla/23d>
Retrieved Apr 21 (2008)
- [40] Chen, H. and Bhanu, B.: *Human Ear Recognition in 3D*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 29 (4) (2007)
- [41] Yan, P. and Bowyer K.W.: *An Automatic 3D Ear Recognition System*. Proceedings of the 3rd International Symposium 3D Data Processing, Visualization and Transmission, 213-218 (2006)